# Preparatory material for the first SEMIC workshop on Personal Data Spaces

This document contains recommended pre-reading material for participants of the first SEMIC workshop regarding personal data spaces. In the first part of this document, the legal (European) framework around personal data spaces is outlined. The second part of this document focuses on the core principles of MyData and SOLID, since they are key players of the emerging personal data space landscape and this workshop.

## 1. Legal Framework

### 1. Data Governance Act

The European Data Governance act will support the establishment and development of common European data spaces in strategic domains such as health, environment, energy, agriculture, mobility, finance, manufacturing, public administration and skills, by involving both private and public players. By doing so, the European Commission seeks to "*increase trust in data sharing, to make it easier for citizens and businesses to make their data available for the benefit of society*", across sectors and borders.

These considerations (chapters 3 and 4) concern the technical means to exchange personal data and non-personal data. This includes the provision of consent, the storage and the reuse of such data. In this context, Solid and MyData are two important elements for Europe to define its means.

### 2. Interoperability Act

Since the publication of the Data Governance Act, many initiatives have emerged or gained visibility with regards to the data spaces, their creation or expansion. Ensuring the flow of data "*across sectors and borders*"[1] requires more than a minimum level of interoperability between the systems at stake. This is one reason for which the European Commission recently adopted the Interoperable Europe Act proposal and its accompanying communication to strengthen cross-border interoperability and cooperation in the public sector across the EU.

The Act proposes to introduce a **structured and co-owned EU cooperation framework** for public administrations with the following pillars:

1. **An Interoperable Europe Board** - that is co-owned by the Member States and the EU and supported by public and private actors – for the development of a common strategic agenda for cross-border interoperability, the support in operational implementing interoperability solutions, and progress monitoring.
2. **Mandatory interoperability assessments** to evaluate the impact of changes in IT systems related to cross-border interoperability in the EU.
3. An '**Interoperable Europe Portal**' as a community platform and one-stop-shop for shared and reusable interoperability solutions

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

4. **Innovation and support measures**, including regulatory sandboxes and GovTech cooperation, to promote policy experimentation, developing skills and the scaling up of interoperability solutions for reuse.

The SEMIC Action has been working on interoperability with a broad community of practitioners for over a decade. It actively supports its community on various technologies developed for tackling the exchanges and reuses of data, such as Solid, in an interoperable manner. Together with the Joint Research Centre of the European Commission and MyData, the SEMIC Action and Solid are bringing together various actors involved in the management and exchange of personal data as part of the Data Governance Act. In the following section, we describe these two initiatives.

# 2. MyData

MyData is a non-profit organisation proposing a human-centric approach to personal data management. It gathers a large community of "*organisations and individuals who have committed to the MyData principles for ethical personal data management*"[2]. This three-pager declaration is reproduced below. It can also be accessed in any European language from here.

Furthermore, companies can show leadership by providing human-centric solutions based on the MyData principles that empower individuals to manage their personal data. Once proven, these companies could receive an award that recognises them as a MyData operator. https://www.mydata.org/participate/declaration/

## 2.1. Declaration of the MyData principles

As the importance of personal data in society continues to expand, it becomes increasingly urgent to make sure that individuals are in a position to know and control their personal data, but also to gain personal knowledge from it and to claim their share of the benefits.

Today, the balance of power is massively tilted towards organisations, who alone have the power to collect, trade and make decisions based on personal data, whereas individuals can only hope, if they work hard, to gain some control over what happens with their data. The shifts and principles that MyData lays out in its Declaration aim at restoring balance and moving towards a human-centric vision of personal data. MyData believes they are the conditions for a just, sustainable and prosperous digital society whose foundations are:

- Trust and confidence, that rest on balanced and fair relationships between people, as well as between people and organisations;
- Self-determination, that is achieved, not only by legal protection, but also by proactive actions to share the power of data with individuals;
- Maximising the collective benefits of personal data, by fairly sharing them between organisations, individuals and society.

[2] https://www.mydata.org/participate/declaration/

## MyData shifts: what needs to change

Our overriding goal is to empower individuals to use their personal data to their own ends, and to securely share them under their own terms. MyData will apply and practice this human-centric approach to our own services, and we will build tools and share knowledge to help others do the same.

### 2.1.1.1. From formal to actionable rights

In many countries, individuals have enjoyed legal data protection for decades, yet their rights have remained mostly formal: little known, hard to enforce, and often obscured by corporate practices. We want true transparency and truly informed consent to become the new normal for when people and organisations interact. We intend access and redress, portability, and the right to be forgotten, to become "one-click rights": rights that are as simple and efficient to use as today's and tomorrow's best online services.

### 2.1.1.2. From data protection to data empowerment

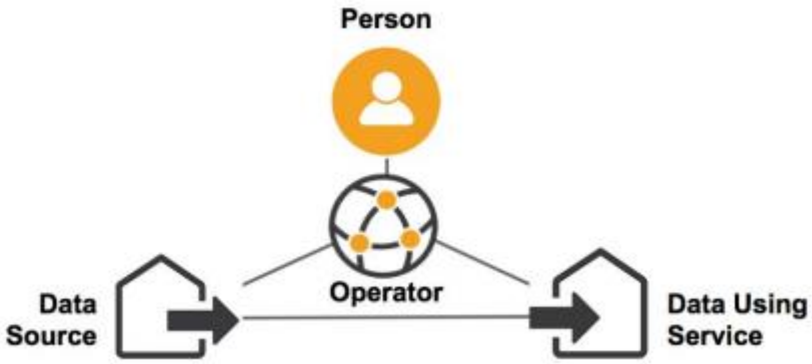
Data protection regulation and corporate ethics codes are designed to protect people from abuse and misuse of their personal data by organisations. While these will remain necessary, we intend to change common practices towards a situation where individuals are both protected and empowered to use the data that organisations hold about them. Examples of such uses include simplifying administrative paperwork, processing data from multiple sources to improve one's self-knowledge, personalised AI assistants, decision-making, and data sharing under the individual's own terms.

### 2.1.1.3. From closed to open ecosystems

Today's data economy creates network effects favouring a few platforms able to collect and process the largest masses of personal data. These platforms are locking up markets, not just for their competitors, but also for most businesses who risk losing direct access to their customers. By letting individuals' control what happens to their data, we intend to create a truly free flow of data - freely decided by individuals, free from global choke points - and to create balance, fairness, diversity and competition in the digital economy.

## MyData Roles: who does what

Please note: "Roles" are not "Actors" – an individual or organisation may fulfill one or more roles at once.



#### 2.1.1.4. Person

An individual that manages the use of their own personal data, for their own purposes, and maintains relationships with other individuals, services or organisations.

#### 2.1.1.5. Data source

A data source collects and processes personal data which the other roles (including Persons) may wish to access and use.

#### 2.1.1.6. Data using service

A data using service can be authorised to fetch and use personal data from one or more data sources.

#### 2.1.1.7. Personal data operator

A Personal Data Operator enables individuals to securely access, manage and use their personal data, as well as to control the flow of personal data with, and between, data sources and data-using services. Individuals can be their own operator. In other cases, operators are not using the information itself, but enabling connectivity and secure sharing of data between the other roles in the ecosystem.

The role of personal data operator is defined by the MyData Operator Reference Model. This model is based on nine functional elements (listed below). Please note that not all these elements need to be covered, only the ones you are focussing on.

1. **Identity management**: how do your operations handle authentication and authorisation of individuals and organisations.
2. **Permission management**: how do your operations enable people to manage, have an overview of data transactions and connections and to execute their legal rights.
3. **Service management**: how do your operations use connection and relationship management tools to link operators, data sources, and data-using services.
4. **Value exchange**: how do your operations facilitate accounting and capturing value, created in the exchange of data.
5. **Data model management**: how do your operations manage the semantics of data.
6. **Personal data transfer**: how do your operations implement the interfaces to enable data exchange.
7. **Personal data storage**: how do your operations allow data to be integrated from multiple sources.
8. **Governance support**: enables compliance with the underlying governance frameworks.
9. **Logging and accountability**: entails keeping track of all information exchanges taking place.

For more information about the MyData Operator reference model, please refer to the following factsheet.

### MyData principles: what we want to achieve

In order to produce the shifts that are needed for a human-centric approach to personal data, we commit to working towards and advocating the following principles:

#### 2.1.1.8. Human-centric control of personal data

Individuals should be empowered actors in the management of their personal lives both online and offline. They should be provided with the practical means to understand and effectively control who has access to data about them and how it is used and shared.

We want privacy, data security and data minimisation to become standard practice in the design of applications. We want organisations to enable individuals to **understand privacy policies** and how to activate them. We want individuals to be **empowered to give, deny or revoke their consent** to share data based on a clear understanding of why, how and for how long their data will be used. Ultimately, we want the terms and conditions for using personal data to become negotiable in a fair way between individuals and organisations.

### 2.1.1.9. Individuals as the point of integration

The value of personal data grows exponentially with their diversity; however, so does the threat to privacy. This contradiction can be solved if individuals become the "hubs" where, or through which **cross-referencing of personal data** happens.

By making it possible for individuals to have a 360-degree view of their data and act as their "point of integration", we want to enable a new generation of tools and services that provide deep personalisation and create new data-based knowledge, without compromising privacy nor adding to the amount of personal data in circulation.

### 2.1.1.10. Individual empowerment

In a data-driven society, as in any society, individuals should not just be seen as customers or users of pre-defined services and applications. They should be considered **free and autonomous agents**, capable of setting and pursuing their own goals. They should have **agency and initiative**.

We want individuals to be able to **securely manage their personal data** in their own preferred way. We intend to help individuals have the tools, skills and assistance to transform their personal data into useful information, knowledge and autonomous decision-making. We believe that these are the preconditions for fair and beneficial data-based relationships.

### 2.1.1.11. Portability: access and re-use

The portability of personal data, that allows individuals to obtain and reuse their personal data for their own purposes and across different services, is the key to make the shift from data in closed silos to data which become reusable resources. Data portability should not be merely a legal right but combined with practical means.

We want to empower individuals to **effectively port their personal data**, both by downloading it to their personal devices, and by transmitting it to other services. We intend to help Data Sources make 3 these data available securely and easily, in a structured, commonly used and machine-readable format. This applies to all personal data regardless of the legal basis (contract, consent, legitimate interest, etc.) of data collection, with possible exceptions for enriched data.

### 2.1.1.12. Transparency and accountability

Organisations that use a person's data should say what they do with them and why and should do what they say. They should take responsibility for intended, as well as unintended, consequences of holding and using personal data, including, but not limited to, security incidents, and allow individuals to call them out on this responsibility.

We want to make sure that **privacy terms and policies reflect reality**, in ways that allow people to make informed choices beforehand and can be verified during and after operations. We want to allow individuals to **understand how and why decisions** based on their data are made. We want to create easy to use and safe channels for individuals to **see and control what happens to their data**, to alert them of possible issues, and to challenge algorithm-based decisions.

### 2.1.1.13. Interoperability

The purpose of interoperability is to decrease friction in the data flow from data sources to data using services, while eliminating the possibilities of data lock-in. It should be achieved by continuously driving towards **common business practices and technical standards**.

In order to maximise the positive effects of open ecosystems, we will continuously work towards interoperability of data, open APIs, protocols, applications and infrastructure, so that all personal data are **portable and reusable**, without losing user control. We will build upon commonly accepted standards, ontologies, libraries and schemas, or help develop new ones if necessary.

# 3. Solid

Solid is an initiative led by Sir Tim Berners-Lee, an editorial team and an active community and larger ecosystem of implementers including public and private organisations. The description below is coming from the Solid website. For more information, an open training was developed by the SEMIC team for introducing Solid.

## 1. What is Solid?

Solid is a specification that lets people store their data securely in decentralized data stores called Pods. Pods are like secure personal web servers for your data.

- Any kind of information can be stored in a Solid Pod.
- You control access to the data in your Pod. You decide what data to share and with whom (be it individuals, organizations, and/or applications). Furthermore, you can revoke access at any time.
- To store and access data in your Pod, applications use standard, open, and interoperable data formats and protocols.

### 1.1.1. What are Solid Servers and Pods?

A Solid Server hosts one or more Solid Pods. Pods are where you store your data:

- Each Pod is fully controlled by the Pod owner (i.e., you).
- Each Pod’s data and access rules are fully distinct from those of other Pods.

You can get a Pod from a Pod Provider, or you may choose to self-host your Pod.

You can even have multiple Pods. They can be hosted by the same Pod Provider or by different Providers or be self-hosted or any combination thereof. The number of Pods you have as well as which Solid Server or Servers you use is effectively transparent to the applications and services that you use. This is because, in the Solid ecosystem, data is linked through your Identity and not through the specifics of your Pod. This is true for your own data as well as for data that others have shared with you.

### 1.1.2. What Data can a Solid Pod contain?

You can store any kind of data in a Solid Pod. This makes Solid special because it offers the ability to store data in a way that promotes interoperability. Specifically, Solid supports storing Linked Data. Structuring data as Linked Data means that different applications can work with the same data.

### 1.1.3. Who has access to a Solid Pod?

With Solid’s Authentication and Authorization systems, you determine which people and applications can access your data. You grant or revoke access to any slice of your data as needed. Consequently, you can do more with your data, because the applications you decide to use can be granted access to a wider and more diverse set of information.

And just as you can share your data with others, they can also share their data with you. This creates rich and collaborative experiences across a combination of both personal and shared data.

### 1.1.4. Solid Applications

Solid applications store and access data in Pods using the Solid Protocol.

Within the interoperable Solid ecosystem, different applications can access the same data instead of requiring separate data silos specifically for the applications. For example, instead of inputting your email with your bank statement notification service, with your phone’s billing service, etc., you can instead store this information in your Pod and grant access to read your email information to these disparate services/applications.

For a listing of some Solid applications, see Solid Applications.

## 2. What does SOLID offer?

Solid (derived from "social linked data") is a proposed set of conventions and tools for building decentralized social applications based on Linked Data principles. Solid is modular and extensible, and it relies as much as possible on existing W3C standards and protocols. Briefly, here is what Solid offers:

### 2.1.1. True data ownership

Users should have the freedom to choose where their data resides and who is allowed to access it. By decoupling content from the application itself, users are now able to do so.

### 2.1.2. Modular design

Because applications are decoupled from the data they produce, users will be able to avoid vendor lock-in, seamlessly switching between apps and personal data storage servers, without losing any data or social connections.

### 2.1.3. Reusing existing data

Developers will be able to easily innovate by creating new apps or improving current apps, all while reusing existing data that was created by other apps.

# 4. Source materials

**Legal Framework**

- https://digital-strategy.ec.europa.eu/en/policies/data-governance-act last visited on 13/12/2022.
- https://joinup.ec.europa.eu/collection/semic-support-centre/welcome last visited on 13/12/2022.

**MyData**

- https://www.mydata.org/participate/declaration/ last visited on 13/12/2022.
- https://mydata.org/wp-content/uploads/2021/09/MYD-Operator-Awards-Factsheet.pdf last visited on 20/12/2022.

**Solid**

- https://solidproject.org/about last visited on 13/12/2022.
- https://academy.europa.eu/courses/introduction-to-solid/view/?fromPath=dashboard last visited on 13/12/2022.
- https://solidproject.org/users/get-a-pod last visited on 13/12/2022.
- https://solidproject.org/self-hosting/ last visited on 13/12/2022.
- https://solidproject.org/TR/protocol#identity last visited on 13/12/2022.
- https://solidproject.org/TR/protocol last visited on 13/12/2022.
- https://solidproject.org/apps last visited on 13/12/2022.
- https://w3.org/ last visited on 13/12/2022.