



What are “personal data spaces”?

Viivi Lähteenoja

University of Helsinki, Finland and Aalto University, Finland
viivi.lahteenoja@helsinki.fi

ABSTRACT

While the concept of “data spaces” is no longer new, its specific application to individuals and personal data management is still undeveloped. This short paper presents a vision for “personal data spaces” in the shape of a work-in-progress description of them and some of the conceptual and implementation features envisioned. It is offered for discussion, debate, and improvement by professionals, policymakers, and researchers operating in the intersection of data spaces and personal data management.

CCS CONCEPTS

• **Social and professional topics** → Computing / technology policy; Government technology policy; Governmental regulations; Computing / technology policy; Commerce policy; Transborder data flow; Computing / technology policy; Commerce policy; Governmental regulations; • **Security and privacy** → Human and societal aspects of security and privacy.

KEYWORDS

Personal data, data spaces, digital policy, EU policy

ACM Reference Format:

Viivi Lähteenoja. 2023. What are “personal data spaces”?. In *Companion Proceedings of the ACM Web Conference 2023 (WWW ’23 Companion)*, April 30–May 04, 2023, Austin, TX, USA. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3543873.3587656>

1 INTRODUCTION

Personal data¹ is simultaneously essential for delivering on the promise of data spaces and at the heart of the most nightmarish scenarios to which increased accessibility and processing of data can lead. Any work in data spaces cannot afford to ignore the personal nature of at least some of the data that these spaces deal with or the normative implications thereof. In other words, personal data has a special role when discussing questions of (semantic and other kinds of) interoperability as well as the FAIRness of data spaces.

¹Personal data is here understood as defined in the EU GDPR article 4 (1), namely as “any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person”.



This work is licensed under a Creative Commons Attribution International 4.0 License.

WWW ’23 Companion, April 30–May 04, 2023, Austin, TX, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9419-2/23/04.
<https://doi.org/10.1145/3543873.3587656>

In view of addressing the needs for available personal data for the benefit of individuals, companies, and societies, and/or mitigating the potential harms stemming from personal data processing, a constellation of similar ideas has emerged over at least the past decade and a half. Various terms around giving individuals better control over personal data availability and processing, or solutions for “human-centric data management”, have been described, developed, or promoted by several people and organisations from at least the “infomediaries” of Hagel and Singer [1].²

In February of 2020, the European Commission published a communication entitled “A European strategy for data” [3], which highlighted the EU’s ambition for increased technological sovereignty both for the bloc as a whole and for its citizens. A centrepiece of this strategy was the adoption of the concept of “data space”, which had been promoted for some years by organisations like International Data Spaces Association (originally “Industrial Data Space Association”, see [4]).

In addition to announcing the creation of nine sectoral data spaces such as the “health data space”, the strategy document also introduced the concept of “personal data spaces”. They are included with the explicit motivation to support individuals “to be empowered to be in control of their data through tools and means to decide at a granular level about what is done with their data” [3 p. 20]. This motivation, as well as explicit references to the MyData movement and “consent management tools, personal information management apps, [...] as well as personal data cooperatives or trusts acting as novel neutral intermediaries in the personal data economy” [3 p. 10] place the strategy document firmly in the tradition of what was referred to as “human-centric data management” above.

While the basic idea of empowering individuals regarding personal data is not new, public-sector focus and efforts for realising these “human-centric data management” ideas have only seen limited uptake in the EU in the past.³ With the prominent promotion of these ideas and the introduction of the concept of “personal data spaces” in the EU data strategy, this is beginning to change.⁴

However, significant conceptual, and practical confusions persist about the specific term “personal data space”, exactly how it is related to the similar, preceding ideas of “human-centric data management”, and precisely how one is implemented.

This short paper presents a working description of what “personal data spaces” are. The concept is considered through several distinct but overlapping perspectives and the description here presented considers the following aspects: 1) definitions found in EU legal and policy documents, 2) conceptual features, 3) design and implementation features, and finally 4) what personal data spaces look like from the points of view of organisations and individuals.

²For references to some of these ideas since 1999, see [2] p. 5.

³See, e.g., the Finnish government programmes from that of Prime Minister Sipilä [5] onwards, following the 2014 publication of a Finnish-language white paper on MyData [6].

⁴See, e.g., the recent efforts by the regional government of Flanders, Belgium [7].

The ambition of this description is simultaneously to stimulate and to focus discourse on the topic among and between data space regulators, developers, businesses, and researchers.

2 DESCRIPTION

This description of personal data spaces is normatively informed by three existing sets of principles that the author is aware of as being especially relevant for the topic. Others no doubt exist and should be considered in any future work. The three sets of principles are: the basic ideas of human-centricity as described by the World Economic Forum [8], European Union digital rights and principles [9], and the MyData principles [10], which are listed below.

Human-centric approaches generally share the following features: HC-1. Human as the logical point of integration, HC-2. Empowerment with data, HC-3. Ecosystems approach, HC-4. Pluralism, HC-5. Equity and proportionality [8].

The European digital rights and principles are defined as the following: EU-1. Putting people and their rights at the centre of the digital transformation, EU-2. Supporting solidarity and inclusion, EU-3. Ensuring freedom of choice online, EU-4. Fostering participation in the digital public space, EU-5. Increasing safety, security and empowerment of individuals, EU-6. Promoting the sustainability of the digital future [9].

The MyData principles are the following: MD-1. Human-centric control of personal data, MD-2. Individual as the point of integration, MD-3. Individual empowerment, MD-4. Data portability: access and re-use, MD-5. Transparency and accountability, MD-6. Interoperability [10].

The topics covered in section 2 have been identified as particularly relevant for personal data spaces but do not represent a conclusive, comprehensive, or exhaustive list of types of perspectives on personal data spaces. The choice of these specific perspectives is the result of the author's academic, professional, and volunteer activity in the field since 2016 and, as a result, it is necessarily subjective. Further work to describe personal data spaces will benefit from a more systematic review of the relevant discourse, academic and other.

2.1 Legal and policy descriptions of personal data spaces

The purpose of personal data spaces is for individuals to “be empowered to be in control of their data”⁵. They are “tools and means to decide at a granular level about what is done with their data”, [3 p. 20]. (HC-2, EU-5, MD-1, MD-3.)

Personal data spaces enable people to reuse, share, and enable the joint use of data. Through, for example, data intermediation service providers, personal data spaces facilitate individuals “exercising the rights of data subjects in relation to personal data”, [11 Article 2 (11)]. (MD-4.)

Personal data spaces can be described in the language of sectoral data spaces as purpose-specific (but only in that their purpose is to

empower individuals), cross-sectoral or horizontal, and “interoperable frameworks of common standards and practices”, [11 Article 30 (h)]. (MD-6.)

2.2 Conceptual features of personal data spaces

The phrase “personal data space” should be read as “data space for a person” and not as a “data space for personal data”. In other words, personal data spaces are data spaces that are particular to an individual (as opposed to being specific to a domain or vertical industry like health, mobility, public administration, etc.). This conceptualisation is in line with the basic tenets of human-centricity, namely that the focus is on the person (and not the data). (HC-1, EU-1, MD-2.)

Because personal data spaces are always specific to individuals, one individual may have more than one personal data space, but one personal data space can only be controlled (directly or through an authorised representative) by one individual.

Personal data spaces, like sectoral data spaces, can be described as collaboration environments with models of interoperating between organisations. Organisations collaborating in the context of data spaces will each fulfil one or more roles in the architecture of the data spaces. In addition to other roles, all data spaces will always include roles, and actors fulfilling these roles, that provide infrastructure services and those that provide end-user services. (HC-3, MD-6.)

In this vein, personal data spaces can also be described as that end-user service layer which caters specifically to individuals (as opposed to companies and other organisations), and which operates on top of the infrastructure layer provided (at least in part) by sectoral data spaces like health, mobility, public administration, and so on.

Multiple natural and/or legal persons can and most often do hold different kinds of rights over the same personal data,⁶ and this is one of the challenges to be addressed by personal data spaces and their implementers and regulators.

The same data can be simultaneously relevant for two or more individuals for example in the case of the energy consumption of a household of two or more people. In this case, it's easy to see how this data is relevant to all the people in the household and how they each have rights over its use.

The same data can also be simultaneously relevant for individuals and organisations for example in the case of the location data generated by ride-sharing service provision. In this case it's easy to see how this data is relevant to the individual driving the car, the individual passenger(s), as well as to the company whose platform is used to make the connection between the driver and the rider(s).

Further complication in this type of cases is that the rights held by individuals and organisations (the latter in the plural because we can also add the telecommunications company whose network is being used for connectivity etc.) are of different kinds from each other. The rights of the passengers are data subject rights. The rights of the driver are both data subject rights and workers' rights (and different depending on whether they are legally defined as

⁵Cf. sectoral data spaces, where “businesses [...] have easy access to an almost infinite amount of high-quality industrial data, boosting growth and creating value,” [3] pp. 4–5.

⁶This, among other factors, is the reason this paper consciously avoids all language of “ownership” about personal data.

employee or as independent entrepreneur). The rights of the ride-sharing platform company can be related to their right to business secrets or intellectual property. The rights of the telecommunications company are likewise, with potential additional rights as providers of critical infrastructure.⁷

As a result, the data (such as energy consumption in a home or location details from trips around town) that is controllable by an individual via a personal data space can be, and usually will be, also in some way controllable or processable via one or more personal and/or sectoral data spaces. This control will be relative to the rights (and restrictions) relevant for the natural and legal persons involved.⁸ (HC-4.)

Collective data spaces are also conceptually possible for formal and informal groups and communities. An example of the former might be a collective data space for a housing association, through which collective utility consumption data might be controllable. An example of the latter might be a collective data spaces for the people who all use a public space like a park, through which data of park usage and occupancy might be controllable. Collective data governance such as would be required for collective personal data spaces have been explored especially by and in the context of indigenous peoples (see, for example, [13]) and in connection with variously defined collectivist concepts like data trusts [14], commons [15], cooperatives [16]. However, individualist models continue to dominate in the EU and collective personal data governance remains at present underexplored especially outside research institutions. (HC-4.)

2.3 Features of personal data space design and implementations

The description of the necessary and optional functional elements of personal data spaces in this section have been informed by the following ideas about personal data spaces:

- Because they are specific to people, they must be human-centric. (HC-1, EU-1, MD-1.)
- Because people have a diversity of needs and preferences, they must be many. (HC-4.)
- Because they must be many, they must be interoperable. (MD-6.)
- Because they must be interoperable, their design should be modular.
- Because the different modules have a diversity of technological and architectural requirements and options, they should be by default agnostic.

A description, not the only one or an entirely uncontroversial one, of the necessary and optional elements that make up the functional whole of personal data spaces is included at the end of this section. The starting point for this list was the definition of MyData

⁷There exists a legal and philosophical debate worth noting, though not discussing at length, here. That debate is over whether rights (fundamental or human rights, specific legal rights applicable in a given jurisdiction, etc.) are *hierarchical* or not and, if they are, *what exactly* is that hierarchy. Does the right to privacy trump the right to pursue economic prosperity? Does the right to be forgotten in the EU trump the rights to processing in other jurisdictions? For a recent attempt to provide a kind of hierarchy, see [12]

⁸Note that this does not imply that personal data spaces are controllable by multiple people, only that the data controllable through them can be.

operators “functional elements” [17], which are: Identity management, Permission management, Service management, Value exchange, Data model management, Personal data transfer, Personal data storage, Governance support, and Logging and accountability.

The list was further developed and refined during engagements with implementers of technical solutions and services for personal data spaces. Further work to understand the interrelations and dependencies, as well as the set of items and their categorisation in general, would be beneficial.

Necessary functionalities:

- Identity management (including authentication)
- Permissions management (including authorisation and access control),
- Activity logging & support for auditing,
- User interfaces for the individuals who are the data rights holders and end users,
- Technical interface for data providers with existing technical access to the data,
- Technical interface for data recipients who are granted technical access to the data,
- Cybersecurity.

Optional functionalities:

- Data storage or hosting,
- Data transfer,
- Data source or using service discovery,
- Value exchange,
- Data model management,
- Identity wallet (storage for attributes, claims, etc.)
- User interfaces for the data rights users who receive rights to use the data,
- Data visualisation.

Implementations of these functionalities are interoperable within their categories and these functionalities may also be provided as add-on services independently of a core personal data space service. (MD-6.)

2.4 Organisation and individual perspectives on personal data spaces

An organisation providing personal data space services to individuals will need to provide (independently or in conjunction with partners) at least: Customer interface and UX for individuals; Legal compliance and means for lawful operations for the regulator; Technical capacity for the core functionalities and possible additional services.

For the individual, a personal data space can appear as an app they use on their device, through a browser log-in, or through other services where they see who is using what data about themselves and for what purpose. They can add data sources like online services they use. They can also accept incoming requests by services or products to access and use data from the data sources added.

3 CONCLUSION

Before concluding, short remarks on potential adoption of personal data spaces. For individuals actually and en masse to adopt services like personal data spaces, these services need to address three major

considerations. First, they need to solve real problems real people have in the real world. Second, the usability standards that big tech platforms provide are the reference for individuals and personal data spaces must match or exceed these standards. Third, as a unique selling point, personal data spaces must operate and be perceived as operating according to higher ethical standards than the incumbent suite of tech companies and online service providers.

In sum, the author hopes the description above will provoke diverse reactions from professionals, policymakers, and researchers at the intersection of data spaces and personal data management. None of the above is conclusive and all of it is meant to be up for debate, refinement, and improvement. This paper will have achieved its objective when it succeeds in provoking thought and reaction in the shape of better ideas than those contained here.

ACKNOWLEDGMENTS

The author wishes to thank the following individuals and groups for their engagement that informed ideas in this paper: Seth van Hooland, Joss Langford, Paul Theyskens, Max Leonard, Emmet Townsend, and the participants of The First and Second Workshops on Personal Data Spaces by the European Commission [18]; Esther De Loof, Ruben Verborgh, Pieter Colpaert, and colleagues at SolidLab Flanders; Kai Kuikkaniemi and Sanna Toropainen; Tuukka Lehtiniemi. All errors and confusions remain strictly the author's.

REFERENCES

- [1] Hagel III, John, and Marc Singer. 1999. 'Unbundling the Corporation'. *Harvard Business Review*, 1 March 1999. <https://hbr.org/1999/03/unbundling-the-corporation>.
- [2] 'Understanding MyData Operators'. 2022. White paper. MyData Global. <https://www.mydata.org/publication/understanding-mydata-operators/>.
- [3] European Commission. 2020. 'European Data Strategy'. Text. European Commission - European Commission. 2020. [https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=\\$1593073685620&uri=\\$CELEX%3A52020DC0066](https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=$1593073685620&uri=$CELEX%3A52020DC0066).
- [4] 'Industrial Data Space: A New Idea of Sharing Data'. 2017. Industrial Data Space Association. https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/About-the-International-Data-Spaces-Association_1.pdf.
- [5] 'Government programme of Prime Minister Sipilä'. 2015. <https://valtioneuvosto.fi/-/74-sipila-hallitusohjelma>.
- [6] Poikola, Antti, Kai Kuikkaniemi, and Ossi Kuittinen. 2014. 'My data - johdatus ihmiskeskeiseen henkilötiedon hyödyntämiseen'. Ministry of Transport and Communications, Finland. <https://julkaisut.valtioneuvosto.fi/handle/10024/77875>.
- [7] Service, Flanders News. 2023. 'Government of Flanders to Partner with Microsoft to Develop Data Vaults'. *belgianewsagency.eu*. 17 January 2023. <https://www.belgianewsagency.eu/government-of-flanders-to-partner-with-microsoft-to-develop-data-vaults>.
- [8] World Economic Forum. 2021. 'On the Importance of Human-Centricity and Data'. World Economic Forum. https://www3.weforum.org/docs/WEF_On_the_Importance_of_Human_Centricity_2021.pdf.
- [9] European Commission. 2022. 'Declaration on European Digital Rights and Principles'. <https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles>.
- [10] MyData Global. 2017. 'MyData Declaration'. MyData Global. <https://mydata.org/declaration/>.
- [11] European Union. 2022. Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European Data Governance and Amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA Relevance). OJ L. Vol. 152. <http://data.europa.eu/eli/reg/2022/868/oj/eng>.
- [12] Müller, Fernando Suárez. 2019. 'The Hierarchy of Human Rights and the Transcendental System of Right'. *Human Rights Review* 20 (1). <https://doi.org/10.1007/s12142-018-0537-z>.
- [13] Carroll, Stephanie Russo, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, *et al.* 2020. 'The CARE Principles for Indigenous Data Governance'. *Data Science Journal* 19 (1): 43. <https://doi.org/10.5334/dsj-2020-043>.
- [14] The Open Data Institute. 2018. 'Defining a "Data Trust"'. Oct 2018. <https://www.theodi.org/article/defining-a-data-trust/>.
- [15] Rosnay, Mélanie Dulong de, and Felix Stalder. 2020. 'Digital Commons'. *Internet Policy Review* 9 (4). <https://policyreview.info/concepts/digital-commons>.
- [16] Hafen, Ernst. 2019. 'Personal Data Cooperatives – A New Data Governance Framework for Data Donations and Precision Health'. In *The Ethics of Medical Data Donation*, edited by Jenny Krutzinna and Luciano Floridi, 141–49. Philosophical Studies Series. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-04363-6_9.
- [17] 'Understanding MyData Operators'. 2022. White paper. MyData Global. <https://www.mydata.org/publication/understanding-mydata-operators/>.
- [18] 'Personal Data Spaces | Joinup'. n.d. Accessed 15 March 2023. <https://joinup.ec.europa.eu/collection/semic-support-centre/personal-data-spaces>.