Exploring the value of decentralised personal data management solutions

6 June 2023

Special report produced by the ODI for ATxSG

Open Data Institute

Contents

About

This report has been researched and produced by the Open Data Institute (ODI), and published on 6 June 2023.

Its lead authors are Thomas Carey-Wilson and Lucas Stiglich. If you want to share feedback by email or would like to get in touch, please contact the ODI research team at research@theodi.org.

This is work in progress. It is likely to be updated as we continue our work. Keep an eye out for updates!

How can it be improved? We welcome suggestions from the community in the comments.

Executive summary

In recent decades, digital platforms have become an increasingly dominating medium through which consumers access goods and services. This comes with trade-offs as the exchange of personal data is often needed to sustain the viability of these platforms' commercial models, eg through personalised advertising strategies.

This, in addition to the dominance of the digital platform model, has led to a regime whereby many individual's personal data is held by an array of different platforms. These data can give platforms and companies a competitive edge in new markets over new entrants (eg Large Language Models, digital healthcare), as well as enabling invasive data practices like online behavioural advertising, or the sharing of these data with law enforcement.[1]

Over the last decade, in an attempt to answer the regulatory debates and challenges around data protection legislation, people and organisations have developed several technologies — providing digital spaces for users to store, transfer and withdraw their data from desired service providers, social media platforms and applications relating to banking, health and even retail. However, as the landscape of these products is at an early stage, there has not yet been a study looking at how these products might benefit their target audiences and the wider economy.

Through this study, we explore the main benefits of these decentralised personal data management solutions (PDMS) for enterprises, individuals, and the wider economy, and map the key challenges that seem to be limiting their widespread adoption. Through desk research, a content analysis of each product description, and supplementary interviews, we found that decentralised PDMSs have the potential to benefit enterprises in four areas:

1. **complying with data protection regulations,**
2. **increasing security and trust,**
3. **improving data-sharing, and**
4. **enabling business model innovation.**

[1] The Guardian (2022), 'Apple says it prioritizes privacy. Experts say gaps remain'

For individuals, the main benefits we identified are related to having more control over how their data is used, being able to move their data between different digital service providers (improving their consumer choices) and, finally, having a better, and more personalised, user experience across multiple platforms.

In addition to the benefits for enterprises and individuals, we also found that decentralised PDMSs can bring benefits to the wider economy. They can increase competition between different actors in the digital economy, by fostering innovation and creating new business models and markets, as well as by improving trust between the different actors in the data economy. These three benefits, in turn, can have the effect of increasing productivity and driving economic growth.

However, the domain is still in a very nascent stage and faces challenges that limit its potential growth in areas where it could provide value. We consider that governments can have a role to play in unlocking the long-term benefits of decentralised PDMSs through:

- **laying the foundations for development of the sector,**
- **building human capacities, and**
- **supporting research and development efforts.**

Lastly, the report covers some initial ideas of areas that could be further explored in future research to inform policy interventions aimed at supporting this sector. This guidance for policymakers will include insights about where these technologies can unlock the most benefits – and for whom – and where they might not.

This report begins with an introduction outlining the development of the web and how these technologies could potentially fit into recent overhauls through Web 3.0. We then move to an explanation of the methodology for this study, followed by an outline of our operating definition of a decentralised PDMS, and how their target audiences fit into this definition. Penultimately, we turn to a section describing the potential benefits of decentralised PDMSs according to practitioners. Lastly, we distil these findings into several high-level recommendations for policymakers to consider how these technologies may be shaped via regulation to provide further benefits.

Introduction

Proponents of Web 3.0 claim that it can offer a more decentralised, or distributed, paradigm for the World Wide Web (hereafter referred to as 'the web'). Purportedly, technologies covered under this technological development address some challenges of the current web, including improved transparency and privacy.[2]

There is great interest in this area, as suggested by, for instance, some sources estimating the market size for associated technologies could reach $33.53m by 2030.[3] But, there are still challenges, such as how it may be regulated, and how to break path dependence on older architectures to drive adoption.[4] Additionally, even fundamental questions about what technologies make up Web 3.0 are not entirely settled.[5]

This report delves into Web 3.0: what it is, and what the people looking to build a new version of the internet are striving to achieve. In particular, we examine decentralised personal data management solutions (PDMSs), a set of tools developed for the purpose of allowing individuals greater control over their data. In addition to gaining a clearer idea of their benefits for individuals and enterprises, we also offer context, caveats and clarifications for those claims. In this way, the report expands on an important area for business leaders and policymakers to understand if they are to unlock the potential of these technologies to enable greater trust on the web. There are a few interesting, speculative, institutional solutions to this problem, such as conceptualising platforms as public utilities,[6] but, for brevity, these are not covered here.

This research therefore examines the core value of this field of decentralised PDMSs, as expressions of a broader paradigm shift toward the Web 3.0 concept.

The report investigates a series of related questions:

- Who makes up the main audience for decentralised PDMSs?
- How do decentralised PDMSs benefit the identified audiences?
- In what way are decentralised PDMSs developing, both technically and commercially? Does this add value, and for whom?
- What role can policymakers play in shaping the future development of these products to unlock further value?

---

[2] TechTarget (2023), 'The biggest advantages and disadvantages of Web 3.0'
[3] Grand View Research (2022), 'Web 3.0 Market Size Worth $33.53 billion by 2030'
[4] Gan et al. (2023), 'Web 3.0: The Future of the Internet'
[5] ComputerWorld (2021), 'You can safely ignore Web 3.0'
[6] Bietti (2021), 'A Genealogy of Digital Platform Regulation'

Each of these explores a dimension of the 'value' of these products; first looking at the groups using them, examining their technical, commercial features and whether these constitute benefits for the identified audience. We will also consider a few challenges for these products to go beyond looking at the *current* features. This will help inform readers to understand where the decentralised PDMS landscape *could* develop to gain further adoption in appropriate sectors and use cases.

Background

Sir Tim Berners-Lee originally envisaged the web in 1989 as a 'distributed hypertext system', written using Hypertext Markup Language (HTML). This is an interconnected network of knowledge in the form of websites (containing text, images, tables, etc) that link to other websites, forming a distributed *web* of information.[7] [8] Decision-making and standardisation in this original form of the web was guided through multi-stakeholder consultation via the World Wide Web Consortium (W3C).

In the following decades, the web developed, through advancements in software, to enable users to move beyond interacting with static information,[9] instead using software to engage in more dynamic ways, such as personal blogs or early versions of social media. Looking back, this is generally characterised as the move from Web 1.0 to Web 2.0.

The emergence of Web 2.0 was underpinned by a wave of technical developments, enabling the dynamic delivery of websites. Some of the foremost innovations were:

- ECMAScript/JavaScript for dynamic client-side content generation and client-server communication,
- PHP for the generation of interactive websites,
- the MySQL database for lightweight server-side data storage, and
- the Apache Server making the configuration of web servers simple, 'cookies' for the local storage of information in a user's web browsers, and proliferation of fast and low-latency internet connections (ISDN/ADSL).

Perez (2002) emphasises how manifold these changes can be, when stating 'technological revolutions [are characterised by a] ... cluster of new and dynamic technologies, products and industries'.[10] While it is important to note

[7] Berners-Lee (1998), 'Information Management: A Proposal'
[8] Merriam-Webster (2023), 'hypertext'
[9] Microsoft (2023), 'Static files in ASP.NET Core'
[10] Perez (2002), 'Technological Revolutions and Financial Capital'

that other particular advancements contributed to this change in the web, for brevity these are not covered here.

As mentioned, during its first decade or so, the use of data on the web was less dynamic, amounting to a two-way process of users receiving, posting and viewing content, but with no way to share that content across platforms. With the aforementioned technical developments, this dramatically changed the nature of data exchange on the web, with users now able to contribute their data to online spaces and in return receive more interactive, personalised content. This development of the web is often referred to as the transition from Web 1.0 (characterised by the capability of users to 'read' digital assets), to Web 2.0 (characterised by user's being able to 'read and write' digital assets).

From this grew the proliferation of online platforms, many of which we now take for granted (such as MySpace, WordPress, Twitter, eBay, Google and many others).[11] As such, this transition towards Web 2.0 also entailed a shift away from the original multi-stakeholder rule-making towards a few select companies dominating values and standards in tech. Online digital platforms are now nearly ubiquitous in the activities of many communities and businesses, with consulting firm Gartner (2021) finding that 91% of businesses are engaged in some form of initiative on a digital platform.[12] As has been argued, the attraction of engaging on these platforms for firms is the personalisable advertising environment, fed by increasingly granular personal data from user activity and registration processes.[13]

Although this dynamic has led to significant growth and broader societal prosperity according to the World Economic Forum,[14] there is evidence it has also raised legitimate concerns among users, civil society organisations,[15] academia,[16] governments[17] and commercial organisations.[18] For one, there have been multiple ongoing debates about how the increasing concentration of data collection and processing in a few large platforms may have negatively impacted individuals and societies, affecting individuals' capacity to exert control over the data that is collected about them, and how it's used.

---

[11] Lee et al. (2010), 'Success factors of platform leadership in web 2.0 service business', Thorne et al. (2008), 'Cyberpunk-Web 1.0 "Egoism" Greets Group-Web 2.0 "Narcissism": Convergence, Consumption, and Surveillance in the Digital Divide', Fuchs (2010), 'Web 2.0, Prosumption, and Surveillance'
[12] Gartner (2021), 'Digitalization Strategy for Business Transformation'
[13] Zuboff (2018), 'The Age of Surveillance Capitalism'
[14] World Economic Forum (2022), 'Digital trust: How to unleash the trillion-dollar opportunity for our global economy'
[15] European Data Protection Supervisor (2018), 'Civil society organisations as natural allies of the data protection authorities'
[16] Imperial College Business School (2020), 'Opaque data practices keep Big Tech uncompetitive'
[17] OECD (2020), 'Government access to personal data held by the private sector: Statement by the OECD Committee on Digital Economy Policy'
[18] The New York Times, Warzel and Thompson (2018), 'Tech Companies Say They Care'

Enhanced data protection legislation, such as the European Union's (EU) General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), has tried to respond to many of the concerns raised in recent years regarding data autonomy and control by mandating measures such as the right for users to be able to access and use the data that organisations hold about them. These regimes have ambitious aims but are not without flaws. For example, they are still rarely enforced in practice, and thus struggle to meet their intended aims. This has led some to wonder whether a new iteration of the web could be conceived, returning to the original ethos of a *distributed* web of information and multi-stakeholder governance.

Therefore, in the last decade or so, a new suite of technologies has gained attention for their aim of returning control of digital assets, such as data, images or video content, to individuals.[19] Examples of this include the role of data standards that outline rules ensuring that data is interoperable between platforms, reducing friction in the cost of users switching between different options.[20] Another is the rise of blockchain, a type of database whereby activity is recorded on distributed ledger technology (DLT) shared throughout a network's nodes,[21] meaning validation of the identity of a user, or asset, can proceed without the need for a central authenticator. Blockchain solves a very specific type of problem related to validation. However, a key problem of DLTs is that they can embed many power structures that exist within current structures. As such, there is a risk with DLTs — and many of the nascent technologies that make up Web 3.0 — to shift power, but replicate the same inequalities between participants as in the current model of platform governance.[22]

But Web 3.0 is wider than just blockchain, as one of the crucial functions of Web 3.0 is architectural decentralisation. So we focus on any technology which returns control of digital assets to users in a decentralised fashion, including architectures such as edge computing, where computations are conducted at the 'edge' of networks nearer to users.[23] This collection of developments is often referred to as a shift from Web 2.0 (characterised by users being able to 'read and write' digital assets) and to Web 3.0 (or, users enabled to 'read, write and own' digital assets).[24]

[19] Investopedia (2022), 'Digital Assets'
[20] Diaz et al. (2012), 'Interoperable Search Mechanisms for Web 2.0 Resources', OECD (2021), 'Data Portability, Interoperability and Digital Platform Competition'
[21] Investopedia (2023), 'Blockchain Facts: What Is It, How It Works, and How It Can Be Used', Africa Renewal, UN Women (2023), 'Transforming internet governance to eliminate online inequalities'
[22] Brou et al. (2021), 'Corporate governance and wealth and income inequality'
[23] Accenture (2023), 'What is Edge Computing?'
[24] Competition and Markets Authority (2023), 'Insight Paper on Web3', Yazdinejad (2022), 'Web3 Challenges and Opportunities for the Market'

As mentioned above, Web 3.0 is a loose collection of lots of different technologies — and means a lot of different things to different people. We refer to Web 3.0, but with an understanding that it is amorphous. Given this, it has been said historically that there is a significant gap in the infrastructure necessary for users to effectively exercise control over their data,[25] especially given the array of platforms that store these data, including personal banking, social media, health facilities, and other types of digital platforms. A major development for proponents of Web 3.0 technology has been the emergence of distributed (or decentralised) tools for managing data in this way. This means systems that allow users (either individuals, groups or organisations) greater control over how data about them is stored and used, either by bringing data physically 'closer' to users in the network, spreading it across a network, or by providing rigorous and secure means of validating a user's identity so that valid control permissions are easily verifiable. These suites of tools, known as PDMSs, can allow users to manage data about them across the fragmented landscape of platforms in which it is stored.

Methodology

For this research, we examine the features of decentralised PDMSs using a mixed methods methodology — conducting a quantitative content analysis to examine the features of decentralised PDMSs as presented in web pages from the organisations deploying these products (eg 'product summary', 'features') and supplementary interview data. Following an **explanatory sequential design** approach, we used the results of the quantitative content analysis to identify themes that we then explored further through interviews.

This research methodology is appropriate for answering the above research questions as the content from the decentralised PDMS website is important, publicly-available information on each product's value proposition that explains the technical and, in some cases, commercial benefits as presented to clients and consumers. Using the qualitative interview data, we could then assess the private nuances of these value propositions in a more discrete environment, removed from any potential public, commercial sensitivities. In this way, the quantitative analysis of data collected from the web informs areas to explore for the qualitative data collection, which then expands on the former. A more detailed description of the methods used can be found in **Annex A**. For now, we turn to a clarification of what we mean by 'decentralised PDMSs'.

[25] Bal (2014), 'User Control Mechanisms for Privacy Protection Should Go Hand in Hand with Privacy-Consequence Information: The Case of Smartphone Apps'

Definitions

What are decentralised personal data management solutions (PDMSs)?

Broadly, as outlined in the previous section, we are characterising decentralised PDMSs as tools that give individuals greater autonomy over how and where their data is stored and used. Within this definition, there is a lot of room for nuance as, for instance, stakeholders of this ecosystem include both individuals and enterprises. Likewise, data 'autonomy' can be achieved in multiple different ways at the architectural level.

Observations from the available evidence (literature covered in this section, and the product summaries presented in **Annex B**) indicate a multitude of different tools providing the core service of allowing users various types and degrees of management capabilities (for example, storage, collection) over their personal data and the way it is used by third parties. Most of these are classified as PDMSs.[26] Although in the available evidence, these are referred to as *solutions*, for purposes of neutrality we will refer to them as *products* instead, to help avoid any foregone conclusion that decentralised PDMSs solve stated challenges.

These products are mainly those that manage data at the 'edge' of networks, where the data is stored close to the user's local environment, reflecting decentralisation.[27] However, we add the caveat that this study goes beyond just looking at this type of decentralised architecture, instead considering all types of products that purport to allow stakeholders greater 'control' of their data through decentralisation. This could also include types of decentralised PDMSs where data is 'split' into fragments across a network to prevent tampering by a central processor.[28] This distinction is explained in further detail below.

[26] Bouganim et al. (2022), 'Highly distributed and privacy-preserving queries on personal data management systems', Ladjel et al. (2019a), 'Trustworthy Distributed Computations on Personal Data Using Trusted Execution Environments', Anciaux et al. (2019b), 'Personal Database Security and Trusted Execution Environments: A Tutorial at the Crossroads'
[27] ibid., Ladjel et al. (2019b), 'A manifest-based framework for organizing the management of personal data at the edge of the network'
[28] Domingo-Ferrer et al. (2019), 'Privacy-preserving cloud computing on sensitive data: A survey of methods, products and challenges'

Decentralised versus distributed

A conceptual area to clarify is what we mean by ‘decentralised’. As suggested above, many PDMSs enable personal data to be moved and stored physically closer to the user. However, there are other products such as Arcana.Network, Sia.Tech and Storj that ‘split’ or ‘fragment’ the user’s data across a network, resulting in the same data being stored across multiple devices.

In this way, the latter type of network is ‘distributed’, whereas the former is ‘decentralised’. This is a widely recognised technical distinction with different benefits, as explained by Baran’s (1964) typology of communication networks, which noted that centralised, decentralised and distributed networks possess different levels of tolerance to system faults or attacks, with centralised architectures being the most vulnerable because all data is reliant upon a single, central point.

By contrast, decentralised networks place data in different points, deeming that ‘complete reliance upon a single point [is not] always required’. Distributed networks are where all data is spread across all nodes in a network, making it highly recoverable in case of a fault since each node can ‘flexibly delegate decisions about routing to alternative nodes’.

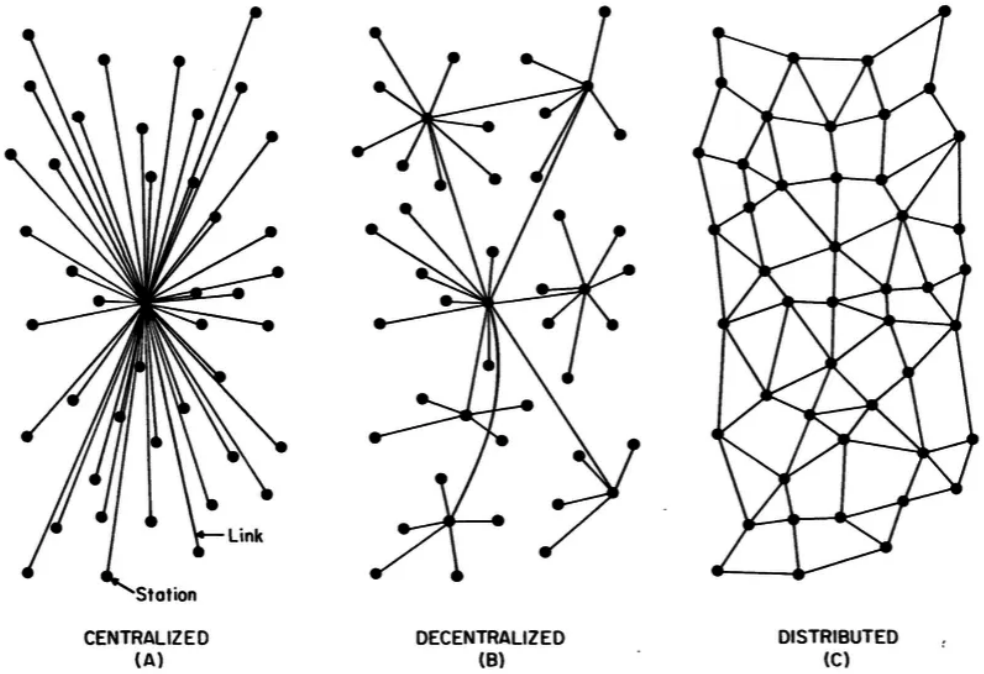

Figure 1. Decentralised versus distributed architectures[29]

[29] Baran (1964), ‘On Distributed Communications’

For personal data management, the difference between decentralisation and distribution is significant since (as shown in greater detail later) a core benefit is protection of a user's personal data against incursion by potentially malignant third parties.

As noted by Vergne (2021), 'the terms decentralised [...] and distributed [...] are often used interchangeably'.[30] We have also observed this in decentralised PDMS product descriptions online, versus more detailed descriptions given during interviews. Thus, where relevant during this study, we will highlight where certain architectures are decentralised and distributed, as relevant to any associated benefits (such as fault tolerance). Although, in keeping with the language used for the majority of the evidence, we will continue to refer to these products as *decentralised* PDMSs.

In addition to different types of architecture, another key part of the decentralised PDMS landscape is the differences in the type of audience. In this sense, we note below three types of decentralised PDMS and their target audience (types of users).

PDMS taxonomy

There are several distinctions between these products on grounds of functionality. The use cases for some are more designed for individuals, and others for businesses, with a few providing data management services on both sides. The evidence we analysed to inform this taxonomy can be found in the product summaries presented via **Annex B**.

Firstly, the 'individual' type of decentralised PDMSs are those products which give individual users the capability to manage their personal data. These include the products mentioned above, that allow users to manage personal data present on various digital platforms throughout the web. It is notable that these often provide unique functionalities for users, for instance dashboards that allow them to easily view what data is currently held by which platforms, and other usage trends. Examples of these individual-centred products include CozyCloud, Helixee and MyCloud. In this manner, the authors note that a majority of these individual decentralised PDMSs possess more decentralised technologies, creating a local 'home cloud' environment for users to store data in one place.

Secondly, we note 'enterprise' decentralised PDMSs are more shaped for data-sharing at an organisational level. They enable enterprise users, both inside and outside companies, to share data necessary for business operations, and to improve services, in a secure and consent-driven way.

[30] Vergne (2021), 'Decentralized vs. Distributed Organization: Blockchain, Machine Learning, and the Future of the Digital Platform'

This type of product lets companies collect data from their consumers and clients, and share data between internal users more easily (reducing friction in data-sharing) while still respecting user consent in adherence with data protection legislation. Examples of enterprise decentralised PDMSs include Datavillage, Inrupt's Solid PODs, Exonum and Storj.

Finally, the third category we have identified is what we have called 'marketplace' decentralised PDMSs. For this last category, the marketing narratives are often that of facilitating a kind of 'data marketplace' for consumers and enterprises, allowing the former to safely store and exchange their data with the latter in return for money. The key difference between these products and the above 'enterprise' types is that they intentionally facilitate the commercialisation of users' personal data, not just its sharing. Many of these types of decentralised PDMSs are blockchain-based, which may relate to some degree of interoperability between blockchain products in general with cryptocurrency and other decentralised finance products (DeFi).[31] Examples of this type of product include Filecoin, Ocean Protocol, Digi.Me,[32] Bitsabout.Me and Arcana.Network.

We summarise our observations about each of these types of PDMSs in the table below.

	Individual	Enterprise	Marketplace
Audience	Individual end users	Businesses, governments and other organisations	Both individual users and organisations
Architecture	Mainly decentralised	Mixture of decentralised and distributed	Mixture of decentralised and distributed
Feature(s)	• Web-wide personal data search (and sometimes retrieval, such as via web scraping) • Analytics for monitoring the use and location of an individual's sensitive data online • Data-sharing functionality with other users	• More scalable architecture than individual PDMSs • System integrations between vendors and external users (other enterprises, consumers) to share sensitive data • User interface to obtain consent from consumers	• More scalable architecture than individual PDMSs • Interface to allow companies to bid and/or buy users' personal data • Often accepting cryptocurrency

Table 1. Decentralised PDMS taxonomy

[31] PWC (2023), 'Making sense of bitcoin, cryptocurrency and blockchain'
[32] World Data Exchange acquired digi.me in October 2022

Although many of these products are referred to in existing studies as *personal data*[33] management solutions, we note that their focus is sometimes more diverse than this. Despite their name, multiple PDMSs, especially enterprise-facing products, also facilitate the exchange of other types of data. For instance, Datavillage outlines that its product helps enterprises to 'keep control of your data *assets*' and 'collaborate on *sensitive* data'.[34] So, to capture the benefits of these more multifaceted products, we will refer to these data as 'sensitive data', which encapsulates personal data and PII, as well as other non-identifiable data that is still sensitive in nature, such as commercial data (relating to transactions, but not necessarily individuals). This helps us include the interests of enterprises as well as individuals through our analysis, but we will refer to personal data specifically where relevant.

We will now explore the significance of this decentralised PDMS taxonomy in the full context of our findings from the content analysis and interviews.

[33] ICO (2023), 'What is personal data?'
[34] Datavillage (2023), 'Datavillage for Media'

Findings

In this section, we describe the benefits and challenges of the decentralised PDMSs that we have identified through research and interviews, alongside some challenges that the sector currently faces.

Adopting new data management products can benefit or challenge not only an organisation, but also other stakeholders in its data ecosystem, such as consumers, providers and commercial partners. Within the organisation, the adoption of these products may also affect constituents differently, as all data users – not just the technical staff (programmers, engineers) but those across different operational functions (eg HR, communications) – have to adapt to a different way of working.

Thus, when discussing the benefits and challenges that decentralised data management products bring, it's important to distinguish the different types of users and stakeholders that may be affected positively or negatively by the adoption of these products.

With this in mind, throughout the section we note, where necessary, how different kinds of products have the potential to provide different benefits and present varied challenges to organisations, to consumers and, in some cases, to other stakeholders. But first, we outline the high-level value propositions of the products overall, as they are presented by the vendors.

Value propositions

The solutions vendors we interviewed tended to have different value propositions depending on the decentralised PDMS type previously identified in this study, such as ‘individual’, ‘enterprise’ or ‘marketplace’ solutions. The first two placed more emphasis on the benefits for end users (individuals) and enterprise consumers respectively, while ‘marketplace’ solutions typically balance benefits from both the other two perspectives.

The enterprise value propositions typically revolve around three factors:

1. promising improved access to data about users that can be processed for valuable insights,
2. reducing the costs and risks associated with legal compliance when handling data about consumers, and
3. being able to offer their end users a better experience while simultaneously offering them more control over how their data is used and shared.

For end users, the value propositions of PDMSs typically promise enhanced privacy and security, as well as additional features that can improve user experience when interacting with different digital services. In that sense, the value proposition for users, beyond just offering them enhanced control of how the data about them is used and shared, typically also includes the promise of a more personalised service offering, increased convenience, or a more seamless experience when dealing with digital services that require data use. These different propositions, in turn, translate into different driving factors for adoption and challenges for each of these decentralised PDMS types, which will be explored in more detail in the next subsections.

Adoption drivers and challenges

Adopting decentralised PDMSs entails a paradigm shift not just for how data is stored, but also for how it is shared across different stakeholders. Consequently, this also determines how these stakeholders may be able to access data and/or generate insights from that data.

For example, if an organisation (enterprise consumer) adopts a decentralised PDMS to manage the data generated from end users using their services, this decision may have consequences for multiple parties or constituents. It not only affects the company and its operations, but also impacts the end users, who may now be able to access new functionalities or interact with different interfaces that the decentralised PDMS offers. Beyond that, it can also have

consequences for other stakeholders which either the organisation or the end user normally share data with.

In that sense, the adoption of a decentralised PDMS affects not just one player but rather an entire ecosystem, and the decision to adopt may well be influenced by the preferences or priorities of multiple stakeholders within such an ecosystem. While some stakeholders may have incentives to adopt PDMSs because of the benefits it brings to them, other stakeholders within the same ecosystem may resist its adoption because of the challenges or risks that it can present to their own interests.

Based on interviews conducted with nine employees of decentralised PDMS vendors and two enterprise consumers (see **Annex C**), in this section, we analyse the array of factors that drive adoption of PDMSs among certain stakeholders, and the challenges that they face that may slow down or even hinder the rate of adoption.

Adoption drivers

In the interviews with vendors and companies deploying decentralised PDMSs, we were able to identify certain factors that have driven or continue to drive their adoption.

- **Data protection legislation:** Introduction of data protection laws, such as the approval of the GDPR in the EU and the CCPA have created an appetite for companies to develop new ways of storing and managing personal data in order to comply with these regulations. This has opened a window of opportunity for decentralised data management products, which are well suited to address some of the challenges that companies face when complying with these data protection regulations. For example, by giving end users the capacity to store their own data, and control directly how and when it is accessed, used and deleted, organisations solve many of the challenges associated with handling personal data themselves, as long as the product providers are able to provide the assurance that they comply with the necessary regulations.

- **Technological changes and platform changes:**
  The abovementioned regulatory changes also influenced trends in the development and use of technologies for generating insights from data. For example, third-party cookies have become less common and Google recently announced that they will be phased out of all Chrome browsers by 2024.[35] Decentralised PDMSs, particularly those that store data close to the end user or in personal data stores and allow users to control how it can be used and with whom it may be

[35] Forbes (2022), 'The Slow Death Of Third-Party Cookies'

shared, offer a new paradigm for enterprises that combine data about specific users (generated across different digital services) in order to personalise these services or to create useful insights. This is a vital shift, given that it can potentially enable end users to control how the data generated from interactions with different companies may be shared and with whom, and for what purposes it may be used.

- **Convenience and improved user experience:** Decentralised PDMSs can also present opportunities for organisations to improve the user experience. With decentralised PDMSs, organisations may be able to access end users' preferences across multiple services and to potentially integrate and automate processes across more than one service. Across most interviewees, there seemed to be a general consensus that, although privacy and control are valued by users, convenience and an improved user experience are key to decentralised PDMSs catching user attention, and therefore drive more service adoption than privacy.

- **Privacy and control:** Finally, although not the main driver, increased awareness about the importance of privacy, and of being able to control how data about oneself is used, may also have an influence on end users' willingness to adopt decentralised PDMSs. This is also indicated to a degree through the narrativised and mission-driven approach to the marketing of decentralised PDMSs, many of which are communicated as delivering control and privacy for users. One example being LifeHash, which promises: 'Data Freedom: Share your information and data without losing ownership. Your data controlled by you: Your data controlled entirely by you!'.[36]

> *The key challenge for the adoption of decentralised data management systems is to build user-friendly and secure applications that allow people to control their data without compromising on convenience and usability.*
>
> — Executive, tech-sector representative from a vendor of decentralised PDMSs

[36] LifeHash (2023), 'Your Digital History'

Challenges for adoption

Although there are multiple factors that have driven adoption of decentralised PDMSs in recent years, the sector also faces certain challenges that have limited its growth. Here we outline some of the main issues that surfaced in the interviews with vendors and companies deploying decentralised PDMSs.

- **Network effects and dependencies:** Like other data-enabled products and services, PDMSs become more valuable as more actors become connected to them. PDMSs that allow organisations or end users to combine data from multiple sources only become valuable when such sources can be connected to them. Moreover, for enterprise clients, adopting decentralised PDMSs typically requires changes in their tech stack that affect not just their own operations, but also their interactions with other actors within their data ecosystem, such as data providers, IT providers, partner companies, among others. In that sense, changing from conventional, centralised data management products to decentralised PDMSs creates a challenge to continue getting data in and out of the organisation, and requires getting buy-in from, and being able to onboard, multiple actors in their ecosystem.

> *The challenge that I think we face is that it's a little bit of a chicken and egg. As in, it becomes far more valuable the more organisations that participate in that data [ecosystem] because the data gets richer through sort of, you know, multiple actors as opposed to one organisation.*
>
> — Senior, public health sector user from an organisation which deployed a decentralised PDMS

- **Lack of interoperability with other technologies or systems:** The challenge of network effects and dependencies can potentially be attenuated if decentralised PDMSs are able to interoperate and integrate with other data systems, technologies and applications. However, many existing applications and systems are not yet designed to work with them. This can limit the decentralised PDMS's usefulness and makes it difficult for potential users to adopt them. For instance, a few of the 'individual' type products examined allow only manual uploading of data, lacking integrations with other applications and platforms entirely.

- **Culture:** Interviewees also mentioned cultural factors as a challenge for increased adoption, as people in some industries and sectors with well-established ways of working are reluctant to adapt to new technologies and ways of work. Additionally, in some cases and among certain organisations there is a reluctance to share data and a fear of 'losing control' over the data they currently hold if it is decentralised. This can also be related to fears of losing competitive advantage or of risking compliance with data protection laws when it has already been established within their current infrastructure.

> *I think one of the biggest challenges with decentralisation is that it requires people to change the way that they think about how data is managed and who controls it.*
>
> — Senior, media-sector user from a company that has deployed decentralised PDMSs

- **Pricing and commercial models:** Commercial models seem to remain a challenge for wider adoption of decentralised PDMSs. Users interviewed mentioned that some vendors offer pricing models that only large-scale operations can afford, thus making it difficult or prohibitively expensive for small and medium enterprises that may be interested in adopting them. Interviewees also mentioned that prices are in any case generally higher for decentralised products than for centralised products. This challenge is further exacerbated because centralised storage vendors not only tend to have lower costs than decentralised vendors, but in some cases — such as Amazon Web Services or Google Cloud Storage — they are also able to cross-subsidise their operations by offering complimentary services, which makes it even harder for new players to compete with them.

> *If I want to go and implement this solution and use the service, I've got to bring like a baseline number of PODs at an annual cost, which is sort of, you know, getting close on to a seven-figure sum. So as an SME, if you want to go build an app in the App Store and put them on pods for like, you know, 30 to 40 grand, you can't do that.*
>
> — Senior, public health sector user from an organisation which deployed a decentralised PDMS

- **Technical complexity:** Decentralised PDMSs are a relatively new technology that requires a certain level of technical expertise to implement and use. Among potential enterprise consumers, many may find the learning curve too steep or may not have the necessary technical skills within their teams to implement and use decentralised PDMSs.

Overall, these challenges may be slowing down the rate at which decentralised PDMSs are adopted, and thus can limit the potential benefits that advocates of decentralised PDMSs claim they could bring. Further down in this report, we propose three areas for possible government intervention to tackle some of these challenges.

Benefits of decentralised PDMSs

As mentioned, adopting decentralised PDMSs can potentially bring a suite of benefits to multiple constituencies or stakeholders. This section outlines the main benefits for enterprises, end users, governments and the wider economy. These benefits have been identified by analysing the literature for the project, by conducting a quantitative content analysis of providers' websites, and through interviews with a sample of vendors and users of decentralised PDMSs.

Before conducting the interviews, we conducted a quantitative content analysis of the benefits that vendors of decentralised PDMSs claim to offer on their websites, vis-à-vis the benefits claimed by centralised data management products. We accumulated 16 centralised data management products (with personal data storage functionality) and coded the benefits outlined in their websites. We then did the same for 28 decentralised PDMSs (see **Annex B**) and compared the results. The outline of the main benefits coded for each is below in Figure 2 and Figure 3, respectively.

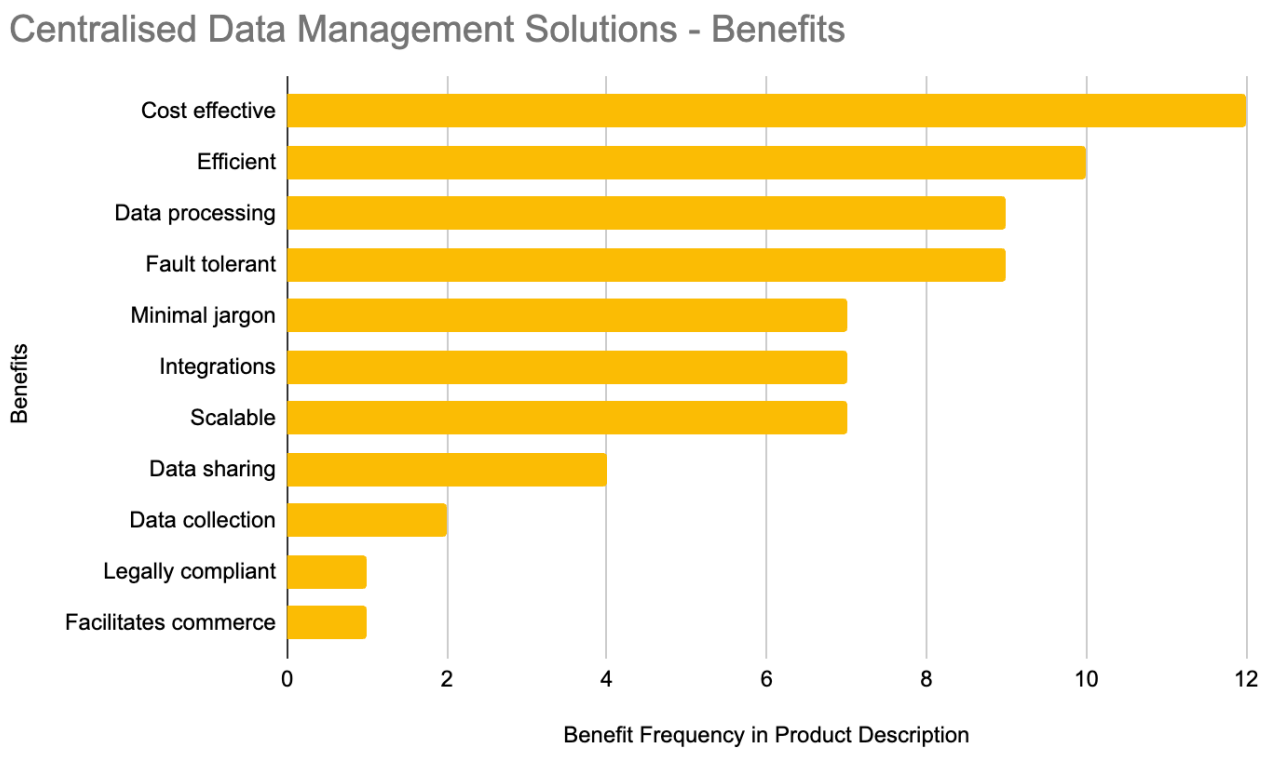

Figure 2. Benefits from centralised data management products

Figure 3. Benefits from decentralised data management products

It is important to note that **'data sharing'** functionality as a list benefit (both between and outside users of the same products) is much more common with decentralised PDMSs than with conventional centralised products.

There is also a clear difference in the frequency of **'legal compliance'** as a benefit (five for decentralised products, and just one for centralised), which relates to the previous points about these products striving to deliver on the promises of data protection legislation.

The number of companies communicating their product with **'minimal jargon'** was comparatively more prevalent in decentralised PDMSs (12 for decentralised solutions, versus seven for centralised). This supports the aforementioned point that having a non-technical marketing narrative is a feature that is valued by end users.

Lastly, **'data collection'** capability is an important benefit highlighted with decentralised products. More specifically, this indicates an ability for the products to identify and retrieve user data on the web. These benefits serve to highlight the user-centric approach emphasised with these products. This includes the wider assurance of a consent-driven service, clarity of language in marketing materials (benefitting users operating across multiple levels of technical proficiency) and data collection functionality, helping users to understand the placement and use of their sensitive data.

We used this initial quantitative approach to inform the next phase of the research design, which consisted of interviews with vendors and end users. In the interviews, we discussed in more depth what benefits both vendors and users perceive these products can bring to enterprises, end users and the wider economy.

Benefits for enterprises

As previously indicated, according to the interviewees and to the claims made by vendors on their websites, the benefits of decentralised PDMSs for organisations that adopt them are manifold. Firstly, these solutions provide users with control over their data and this helps enterprises comply with data privacy regulations and manage data protection risks more easily.

Additionally, interviewees mentioned that decentralised PDMSs can enable more responsible data-sharing across organisations and between organisations and end users, promoting greater trust as end user data is used and shared in a transparent and ethical manner.

Finally, decentralised PDMSs can potentially lead to innovation opportunities for organisations, allowing them to build more personalised and engaging experiences for their customers, through gathering more user data for insights across the web and multiple digital platforms. This can lead to the creation of new products and services, or even entirely new markets potentially, all underpinned by this technology.

> *My application reaches to wherever the data is, so multiple different applications can use the same data. I can get explicit permission from users to use that. It's much easier for me to do things like comply with GDPR, for example. And when you separate the two, the user may decide to store their data somewhere else, I don't necessarily have to store data. So if I'm selling books or selling shoes, storing data is just a side effect of the digital world. I don't necessarily go out saying: I want to store data.*
>
> — Executive, tech sector representative from a vendor of decentralised PDMSs

- **Reducing friction in compliance**

  The continuous presence of ‘legal compliance’ in the content analysis above shows that the capacity to alleviate some of the burden that enterprises face in complying with data protection legislation remains desirable for enterprises. For example, reducing friction for companies obtaining consent for a user’s personal data while safeguarding their ability to deny or grant access. In this way, friction is reduced for end users as they are able to allocate their sensitive data from a single user interface, potentially reducing the risk of fatigue related to an abundance of platforms[37].

  Decentralised PDMSs can relieve organisations from having to create procedures and protocols to comply with some of the requirements introduced in data protection laws, such as the right to data portability and the right to be forgotten. By storing data in personal stores linked to individuals, and allowing individuals to choose how and with whom to share their data, and when to withdraw and/or delete it, such regulatory requirements can be fulfilled without the need for enterprises to develop new procedures.

  Moreover, in comparison with centralised data management solutions, decentralised solutions allow for greater flexibility in complying with different data protection requirements across multiple jurisdictions (supported by ‘legal compliance’ being of comparative value in our content analysis). According to one interviewee, platforms built on decentralised data architectures can leverage metadata to create rules that model compliance locally, and then associate those rules with the location where the data is collected or stored.

> *If you're building a general platform, what I've learned is that you can build it such that compliance can actually be modelled in your metadata and live with your data. [...] That metadata can be things like access control policies. It can be sets of rules. It can be sets of notifications.*
>
> — Executive, tech sector representative from a vendor of decentralised PDMSs

[37] Wang et al. (2022), ‘User characteristics, social media use, and fatigue during the coronavirus pandemic: A stressor–strain–outcome framework’,
Forbes (2022), ‘Digital Agility Depends On A Platform Approach: Three Things To Consider’

- **Increasing security, transparency and trust**

  Decentralised architectures can reduce reliance on centralised platforms and therefore may potentially reduce security risks such as data breaches or leaks at scale. Interviewees mentioned that, given that data is distributed across multiple nodes and devices, the risk of massive data leaks is significantly reduced in comparison to the risk in centralised products.

  Additionally, decentralised PDMSs may be able to offer greater transparency to end users and to other stakeholders, since end users are able to control and monitor how their data is being used, and how it is being transferred and to whom.

  In turn, increased security and transparency can build trust among consumers and other stakeholders in the enterprise's data ecosystem. According to an ODI-commissioned study into the economic impact of trust in data ecosystems, conducted by Frontier Economics, trust can increase the different actors' willingness to share data with each other, and thus the potential benefits that can be obtained from data within a given ecosystem.[38]

  However, it is noted in the literature that cloud-based decentralised solutions have certain security limitations, the main one perhaps being that the data security is dependent, to an extent, on the user's local network environment being trustworthy. If the user's network is compromised, this could lead to security breaches, which are harder for providers to oversee or prevent.[39]

[38] Frontier Economics for the ODI (2021), 'The economic impact of trust in data ecosystems'
[39] Bouganim et al. (2023), 'Highly distributed and privacy-preserving queries on personal data management systems', Anciaux et al. (2019a), 'Personal Data Management Systems: The security and functionality standpoint', Anciaux et al. (2019b), 'Personal Database Security and Trusted Execution Environments: A Tutorial at the Crossroads'

- **Making data more easily findable and enabling data-sharing**

  Decentralised PDMSs can play an important role in enabling findable, accessible, interoperable and reusable (FAIR) data. FAIR data is a key challenge for organisations looking to share data with other stakeholders. Typically, making data FAIR requires preparing adequate metadata to ensure that data is easily processable by computational systems with little human intervention, and thus that it can be more easily reusable by different organisations without requiring much extra processing.[40]

  Decentralised data products can potentially help make data more findable. Just as the web allows users to find content by typing in an address, some decentralised PDMSs are built over an infrastructure that allows data resources or personal data stores to be associated with a unique ID that can be findable through the web. We found that these products could also improve data interoperability. This is explored more in the next section on benefits for end users. But we did not find that these products would improve, or present issues to, accessibility or reusability.

> *It works the same way as web pages work. You know, the whole idea that we can have a unique ID, a Web ID for the personal data store itself so it's easily discoverable.*
>
> — Senior, public health sector user from an organisation that deployed a decentralised PDMS

[40] The ODI (2023), 'Understanding the social and economic value of sharing data', GO FAIR (n.d.), 'FAIR Principles'.

- **Enabling new business models and product and service development**

  With greater control and access to their data, and with a clear channel to obtain consent for their customers' data, decentralised PDMSs can allow organisations to develop new business models that rely on data-sharing, data monetisation, and data collaborations with other organisations.

  For example, there have been initiatives to leverage decentralised PDMSs to offer customers compensation in exchange for sharing certain data with specific stakeholders (our identified 'marketplace' type product). This opens the door to new business models that allow businesses to benefit from data-sharing while respecting end users' rights and adequately incentivising them for the data they provide.

  However, it is important to note that so far, initiatives to create data marketplaces tend to struggle because of the many challenges associated with the particular economic characteristics of data, which make it difficult to price and trade datasets. According to the Organisation for Economic Co-operation and Development (OECD), in 2019, only 2.1% of firms with more than 250 employees were selling data and only 4.6% purchased data. These percentages were lower for smaller firms.[41]

  Additionally, decentralised PDMSs not only help organisations engage with their customers in new and innovative ways, but also to develop more personalised experiences and offerings. This can be a competitive advantage for companies building their services on decentralised data architectures when compared with those who use centralised data management solutions.

[41] OECD (2022), 'Measuring the value of data and data flows', see also The ODI (2023) 'Understanding the social and economic value of sharing data'

Benefits for end users

Decentralised PDMSs also benefit end users who interact with them in multiple ways. The main benefit is that they provide individuals with greater control over the data about them that is generated or collected through their interaction with digital services, allowing them to choose how and with whom they share such information. By storing the data close to the user instead of on centralised platforms, decentralised PDMSs also allow individuals to interact with multiple applications and services, and, more importantly, to transfer data from one application to the other easily. This has the double benefit of, on the one hand, allowing users to customise and enjoy a more personalised experience across multiple services, and on the other, reducing user lock-in and giving them more control and choice over the services.

- ## Control, privacy and transparency

  Decentralised PDMSs can put users in control of their personal data and allow them to decide who can access it and for what purpose. Additionally, these products may provide enhanced privacy and security for users, as their data is stored at the edge of the network, in addition to users being able to control who can access and use it. Several authors have noted how the deployment of privacy enhancing technologies (PETs), in particular trusted execution environments (TEEs), can be easily deployed with decentralised PDMSs.[42]

  Both decentralised and distributed solutions are less likely to be affected by serious long-term damage in the event of breaches, leaks and attacks, compared to centralised products. There is evidence that integrating blockchain-based technology ensures that users are personally unidentifiable and so, in the event of a security breach, the risk of user identification is drastically reduced to practically zero.[43] Also, fault tolerance is a significant benefit for decentralised PDMS solutions, which distribute personal data across a network (distributed architecture). This, coupled with blockchain technology to provide a tamper-proof data structure, is noted as particularly effective in preventing corruption of PDMS network nodes.[44]

[42] Anciaux et al. (2019b), 'Personal Database Security and Trusted Execution Environments: A Tutorial at the Crossroads', Loudet et al. (2018), 'SEP2P: Secure and Efficient P2P Personal Data Processing'
[43] Zainal et al. (2022), 'A decentralized autonomous personal data management system in banking sector'
[44] Bouganim et al. (2023), 'Highly distributed and privacy-preserving queries on personal data management systems', Ladjel et al. (2019b), 'A manifest-based framework for organizing the management of personal data at the edge of the network'

Another notable blockchain use case to highlight the benefit of data security in PDMSs is the use of smart contracts to set and enforce the 'rules' for user access to the data without the need for a central authenticator, which is also a valuable security mechanism identified by those interviewees who were decentralised PDMS vendors.

Finally, a repeated finding evident from the interviews is that by allowing users to see who is accessing and using their data, personalised PDMSs generally offer more transparency to end users than other data management products. This again has the potential to support the propagation of 'trust' in the exchange of personal data from users.

- **Enabling data portability and reducing lock-in**

Decentralised PDMSs can play a role in upholding the right to data portability, which is sanctioned in data protection laws such as the GDPR. The ODI has previously defined data portability as *'the ability to share data between people, groups and organisations. A company, for example, might 'port' data – which could involve the transfer of data, or the provision of access to it – to a third party in order to deliver a particular service.'*

While data portability has been included in multiple data protection laws and regulations around the world, its implementation hasn't necessarily fulfilled its goals. This is at least partially because traditional data management infrastructures and the business models designed around them are not well suited to share usable data with users. Data portability is technically possible within current centralised architectures, but business models of centralised platforms that are built around data collection tend to restrict users' access and control over their data.

For instance, to exercise their right to portability, users typically need to file a one-off request to a service provider, and then the service provider has to process this request internally and share the data about the user that they hold at that given moment.[45] In a context in which decentralised PDMSs adoption was widespread, users could theoretically have continuous access to the data generated through their interactions with digital services, and control in real-time with whom to share it, and for what purpose. However, this comes with certain caveats: for example, with decentralisation of data storage, gaining a 'master view' of the data and coordinating any activity across a network can become challenging.[46]

[45] The ODI (2022), 'Bottom-up data institutions: mechanisms for government support'
[46] Forbes (2022), 'A Decentralized Approach to Database Management'

Because of this, decentralised PDMSs also have the potential to reduce switching costs for users and thus improve market competition and reduce users' lock-in. Decentralised PDMSs give users more control over their data by allowing them to store and manage their data independently of any particular platform or service, and they may also enable interoperability (for example, via integrations) between different platforms and services. This can allow users to switch easily between different services without losing their data or having to start from scratch.

If decentralised PDMSs do improve data portability and interoperability, this can potentially improve competition in digital markets, as noted by the OECD, potentially leading to customers being able to access better digital services offerings.[47]

> *Combining portability with control lets me say [...] 'I don't like what you're doing. I'm bringing my stuff over here.' And if you want to change your mind every two days and to make it seamless, just press a button. [...] it's one of the important things that prevents the consolidation over time of data into a few huge places controlled by a few people or organisations, right? You have to give the individual or the organisation or the entity the power to move where they want.*
>
> — Executive, tech sector representative from a vendor of decentralised PDMSs

- **Improved user experience**

  In addition to giving users more power over how their data is used, decentralised PDMSs can improve end users' experience through integrating and potentially automating processes across multiple services. By allowing users to selectively share certain aspects of their data with specific companies and applications, decentralised PDMSs can allow companies to access the data gathered through their interaction with services from different organisations. Companies may use the data to better understand their end users' preferences and use this to provide more tailored experiences.

  For example, decentralised PDMSs can potentially allow individuals to integrate data about them gathered from sources such as social media accounts, online shopping history, fitness tracker data or bank transfers, as suggested by the 'data collection' functionality benefit

[47] OECD (2021), 'Data portability, interoperability and digital platform competition'

from the content analysis. This can enable companies to have a more holistic view of their end users' preferences and activities, which they can use to provide more personalised recommendations and services, such as workout routines based on fitness data, products based on online shopping history or facilitating more seamless, automated loan requests. At the same time, this can allow customers to simultaneously update their information across multiple services, and to keep their preferences on data-sharing and consent when switching between digital platforms.

Many of these functionalities are also achievable by using centralised data storage vendors, such as login providers that allow a user to have the same identifier across different services, or by allowing companies to aggregate data from different sources to develop more tailored offerings. However, decentralised PDMS vendors claim to be able to offer them without compromising users' privacy and while at the same time allowing them to keep control over their data, thus helping to address the concerns that have created distrust about big technology vendors in recent years (as noted in the introduction of the report).

> *We see this in the simplest of things even. When I change my address, how many places do I need to actually tell in an ecosystem? None, because once you change it, the notifications layer within this solution can actually inform all interested parties that you've changed it. So you may have given permission to your utility company, your bank, the post office, to the weird little subscription to a tech magazine that you had 10 years ago and all the other plethora of things in your life, just by changing it in one place.*
>
> — Executive, tech sector representative from a vendor of decentralised PDMSs

Benefits for the wider economy

Beyond the benefits to enterprises and end users, the adoption of decentralised PDMSs can help societies and governments foster digital innovation and economic growth, while protecting the privacy and security of their citizens' personal data. By reducing end user lock-in and allowing them greater choice between services, decentralised PDMSs can support governments in improving market competition, and driving innovation.

- ## Competition

  Data can be a source of market power, especially in the digital economy. This is not only because of the important role it plays in enabling digital business models, but also its economic characteristics, such as economies of scale, economies of scope and network effects, which create market failures. Additionally, data can reinforce the position of incumbents against that of new entrants – data collection and processing can allow first-movers in a market to gain a position that is difficult for new entrants to overcome. They may also be in a position to leverage that advantage into other markets.[48]

  As mentioned in previous sections, by giving users more control over their data and enabling them to share it or to request for it to be shared with other organisations on their own terms – that is, by improving data portability and interoperability – decentralised PDMSs can play a role in improving market competition.

  With improved data portability and interoperability, entrants and smaller businesses could potentially gain access to similar, or the same, data as their larger competitors. This would enable them to better understand customer needs and preferences and develop products and services more tailored to their customers' needs. Although in theory this benefit is possible, the landscape needs to develop more large-scale actors for tangible examples to become clear.

  Additionally, with the increased transparency and accountability that comes with decentralised PDMSs, businesses could have a stronger need to compete on the quality and price of their products and services, driving efficiencies and productivity. Estimates for the UK suggest that greater personal data portability and interoperability could increase GDP by £27.8bn per annum, just by improving competition, creating efficiencies, and driving productivity.[49]

[48] OECD (2021), 'Data portability, interoperability and digital platform competition'.
[49] CtrlShift for the Department for Digital, Culture, Media & Sport (2018), 'Data mobility: The personal data portability growth opportunity for the UK economy', The ODI (2022), 'Introducing the Smart Data Innovation Guidebook'

> *I think the decoupling and decentralisation of data also creates a whole market in data solutions, which gives the consumer choice. So the portability means I can actually decide to move from party A to B. There's no lock-in.*
>
> — Executive, tech sector representative from a vendor of decentralised PDMSs

- **Innovation**

  In addition to improving competition, increased data portability and interoperability can drive innovation in many areas, including the creation of new products and services, business models, and markets.

  For example, several interviewees suggested that enhancing access to data may enable the creation of new business models that rely on data management, sharing and analysis. This, in turn, can support the creation of new products and services built on top of them, while at the same time increasing data use and sharing, and thus the potential value that can be created from it. For example, studies on 'data mesh' architectures for enterprise data sharing have found that the decentralised and consent-driven exchange of data via data marketplaces inside enterprises can 'establish an innovation ecosystem and drive data-driven business decisions'.[50]

  Finally, the possibilities that decentralised solutions offer in terms of collaboration between different stakeholders, such as individuals, organisations and governments, can also lead to the development of new ideas and solutions that would not otherwise be developed were the data centrally held by separate organisations.

[50] Tata Consulting (2022), 'Enabling business-centric data platforms in financial services - Part 1: Data mesh shows the way'

- **Trust**

With decentralised PDMSs holding the promise of giving end user control, improving transparency on the way data is being used, and enhancing privacy for the end user, widespread adoption of decentralised PDMSs can help strengthen trust between individuals, organisations and the different actors within the digital economy.

A study commissioned by the ODI from Frontier Economics found robust quantified evidence that greater trust data ecosystems are associated with increased data sharing, which in turn can drive value creation.[51] [52] This effect is not specific to decentralised PDMSs, but refers instead to the impact of increasing trust in data ecosystems more generally. If the adoption of decentralised PDMSs can foster trust among users, as their vendors claim, their adoption could potentially have an impact on value creation and economic growth in the medium term.

> *This solution is really about giving back trust to users. If we are able to give back that trust, then we can really change the way that data is being used and shared on the web.*
>
> — Executive, tech sector representative from a vendor of decentralised PDMSs

Not only are competition, innovation and trust positive outcomes in themselves, decentralised PDMSs can also have the effect of driving productivity and boosting economic growth. Therefore, the widespread adoption of decentralised PDMSs could theoretically yield benefits, not just to those who adopt them, but also to the wider economy. Although the empirical quantitative evidence for the magnitude of these impacts is still lacking, we consider these findings should spark greater debate and research into how governments can support the development of this sector.

[51] Frontier Economics for The ODI (2021), 'The economic impact of trust in data ecosystems'
[52] See Bennett Institute (2020), 'The Value of Data – Accompanying Literature Review', The ODI (2021) 'Policy to unlock the economic value of data'

Potential areas for government involvement

As seen in the previous subsections, decentralised PDMSs have the potential to bring benefits not just to organisations and users but also to the wider society. However, the adoption of decentralised PDMSs may happen at a slower rate than what would be socially optimal, in the view of relevant actors. This is because of the existing challenges for wider adoption of these solutions by enterprises, and because of the economic characteristics of data that incentivise certain players to amass it to maintain their market position.

Given the early-stage development of these solutions, and the potential they may hold for creating societal benefits in the future, we believe that governments can play an important role in supporting the development of this sector. We also believe this development can translate to greater international discourse, wider interest and even adoption of these solutions in the medium term. In this section, we propose three areas in which governments could potentially have a role to play.

Different political, cultural and technological contexts around the world may call for different kinds of government interventions. However, we believe the recommendations given in this section can inform areas for governments to explore to achieve further societal benefits from decentralised PDMSs. Rather than strict policy recommendations, what we propose here are paths that governments can examine further, to learn more about the potential benefits of decentralised PDMSs and the most effective ways of realising them.

Several of these recommendations might also provide some useful information for enterprises and other non-government organisations looking to support or even adopt decentralised PDMSs. In particular, ‘Building capacities’ and ‘Create an evidence base to steer action for future developments’.

Laying the foundations for PDMSs

Governments can have the capacity to establish foundations to shape and promote the development of the decentralised PDMS sector. Some areas for this are:

1. Establishing regulatory frameworks to consider how decentralised PDMSs can uphold data subjects’ rights.

2. Developing guidance to support organisations interested in adopting decentralised PDMSs, to assess how these organisations can use decentralised PDMSs effectively and still comply with existing data protection regulation.

3. Promoting the development of standards, enabling interoperability of decentralised PDMSs and other data management systems.

Building capacities

While some of the challenges limiting the growth of the sector may be related to legal, regulatory and technical factors, the human factor is also key. Awareness of decentralised PDMSs and their possible benefits, as well as the literacy and skills to deploy them, are still challenges to be solved. Additionally, the challenges related to network effects and dependencies between different actors within an ecosystem could also be addressed with government support. Some possible efforts are:

1. Supporting campaigns to create awareness about the benefits of decentralised PDMSs. This can help reduce cultural barriers to adoption and drive sector growth.
2. Providing or promoting educational and training programmes to develop the literacy and technical skills needed to develop and implement decentralised PDMSs.
3. Fostering collaboration among multiple stakeholders to uptake adoption at the ecosystem level. For instance, by hosting fora and running events (for example roundtable discussions) to spark dialogue about these technologies.

Create an evidence base to steer action for future developments

Finally, we consider the need to develop stronger evidence to continue supporting the development of the sector and inform future policy interventions:

1. Encouraging and supporting research and development efforts in decentralised PDMSs, for example through funding academic institutions, organisations interested in researching and experimenting with decentralised PDMSs.
2. Creating multi-stakeholder working groups from different sectors, including industry, academia, civil society, and government, to collaborate, and share knowledge and expertise.
3. Running small-scale experiments to demonstrate the impacts of decentralised PDMSs, such as technical trials, focus group discussions and sandbox programmes. This will also support policies and regulations directed at this sector.

Research limitations

Before elaborating on our findings, there are some methodological and topical limitations of this research to mention. Firstly, the scope of solutions examined for this research is not exhaustive, as there are other solutions that could be classified as decentralised PDMSs. Instead, this research is meant to provide a useful snapshot of the solutions that appear prominently from a conventional web search.

However, given there is evidence that search engine optimisation is positively correlated with market share,[53] we can infer that the solutions identified are at least *mainly* strong market actors in this vertical. It should also be mentioned that there are initiatives promoting greater data control, such as MyData[54] that do not produce their own technical solutions but are influential in this space. However, these are not covered in this research for brevity.

Through the course of this research, we have discovered that these solutions are largely early-stage technologies. So the findings are potentially time-sensitive, as there is plenty of leeway for the decentralised PDMS market to develop in new ways that depart from this study's findings. We also note that given the early stage of development of this sector, empirical evidence of its impacts is still scarce, and therefore it is not yet possible to conduct a comprehensive, purely quantitative analysis of its costs and benefits to society.

Additionally, there is a lack of examples of past policy initiatives aimed at supporting this sector, which made developing evidence-based policy recommendations to drive positive impact challenging. Even so, based on the qualitative findings of this study, we proposed some recommendations for policymakers to support the development of this sector, alongside evidence that can inform future policy interventions, in the last part of this report.

Throughout the research, a major undertaking has been coming to a definition of decentralised PDMSs, especially given their technical diversity and immaturity. Therefore, it is essential to discuss our operating definition of decentralised PDMSs before we explore the analysis and interview findings.

[53] Zhang et al. (2017), 'Search Engine Optimization: Comparison of Link Building and Social Sharing', Bhandari et al. (2019), 'An Analysis Between Search Engine Optimization Versus Social Media Marketing Affecting Individual Marketer's Decision-Making Behavior'
[54] MyData (2023), 'MyData'

Conclusion

The web has been through a series of changes over the last couple of decades regarding how its users interface with digital assets. Specifically, from the static, linear milieu of Web 1.0, to manifold recent developments categorised under Web 3.0 that are shaping digital assets to be more consonant with notions of 'control', engendering greater consumer trust in markets.

It is increasingly evident that, as Luciano Floridi (2014) argues in 'The Fourth Revolution', personal data-driven platforms and services are shaping our realities as an environmental force.[55] Therefore, outside of technical, moral, political and regulatory debates, there is little argument that if high-potential technologies like decentralised PDMSs continue to develop, those holding the levers of power should pay attention.

Through this research, we have investigated the value of decentralised PDMSs and discovered that there is significant potential for these technologies to unlock a mutually-beneficial economic regime between consumers and enterprises, undergirded simultaneously by trust and growth.

We have observed that these products have an edge on centralised data management products in their ability to deliver on the principle of data portability, strengthening choice and autonomy for data subjects, while developing a more competitive environment for businesses; buttressing new market entrants by reducing barriers to consumer choice.

Conversely, there are still challenges to be addressed when making these products widespread. We believe governments can play a role in addressing some of these challenges by working together with the private sector and civil society to develop adequate legislation and guidance, nurturing the right skills and creating awareness among relevant audiences, and to develop research that can yield the evidence to inform future investments and policy decisions in this area.

We will continue to investigate these issues, and are interested in hearing from others working on this topic across industry and civil society. If you'd like to stay informed about our work in this area, provide feedback on this report, we would love to hear from you! You can get in touch with us by emailing the ODI's Research and Development team at research@theodi.org

[55] Floridi (2014), 'The Fourth Revolution — How the infosphere is reshaping human reality'

Annex A

Methods

As mentioned in the Methodology chapter, our methods for this project include a quantitative **content analysis** of the features of decentralised PDMSs as presented in 28 web pages reviewed, drawing out the different types of benefits of the products highlighted on publicly-available data.[56] These web pages were discovered using several search engines including Google and Bing, deploying search terms including the following:

- “Decentralised” data management solution
- “Decentralized” data management solution
- “Decentralised” data management product
- “Decentralized” data management product

The bias potentially introduced by search engine optimisation (SEO) was disregarded as problematic given we were searching for major market actors.

Then, we conducted **interviews** with 11 representatives from organisations both producing and deploying these products (seven of the former, three of the latter) across government organisations and industry, covering three continents (Europe, North America and Asia). These representatives were discovered and approached for an interview via LinkedIn. We chose the platform as we were mainly looking for business community representatives, and this is LinkedIn’s main audience.[57]

Although they were anonymised to ensure discretion, the typical mix of these interviewees included representatives from technical, product and business development backgrounds, with a range of seniorities from junior to executive staff.

Lastly, the transcripts from these interviews were analysed according to a **thematic analysis** method, with common themes (as relevant to the project objectives) identified and coded from the interview transcripts.[58]

[56] Columbia University (2023), ‘Content Analysis’
[57] LinkedIn (2023), ‘LinkedIn Audience Network’
[58] Braun et al. (2006), ‘Using Thematic Analysis in Psychology’

Annex B

List of centralised storage products

Name	Country
AWS Cloud Storage	US
Microsoft Azure Storage	US
Oracle Cloud storage	US
HPE Multi Cloud Storage	US
Alibaba Cloud Drive	China
VMWare Cloud Flex Storage	US
IBM Cloud Storage	US
Radar Healthcare	UK
ISCSI	US
Data Dynamics	US
Filecloud	US
N-Able	US
DropBox	US
Google Cloud Storage	US
Google One	US
Box	US

List of decentralised storage products

Name	Country	Active since	Sector	Classification
Cozy Cloud	France	2012	Banking, Healthcare, Multi-purpose	Individual
Digi.me	UK	2009	Healthcare, Financial, Media, Multi-purpose	Marketplace
Bitsabout.me	Switzerland	2017	Multi-purpose	Marketplace
Inrupt Solid PODs and PODBrowser	US	2017	Multi-purpose	Enterprise
Helixee (formerly by NovaThings)	France	2014	Multi-purpose	Individual
MyCloud (but also Western Digital)	US	2012	Multi-purpose	Individual
OpenPDS/SafeAnswers	US	2014	Multi-purpose	Individual
Exonum (part of Bitfury Group)	Netherlands	2019	Commerce, Multi-purpose	Enterprise
MedRec (part of Sirma Solutions)	US	2020	Healthcare	Individual
Datavillage	Belgium	2019	Multi-purpose	Enterprise
DataSwift	UK	2015	Multi-purpose	Enterprise
VITO	Belgium	2018	Healthcare	Individual
LifeHash	Australia	2021	Legal, Insurance, Supply Chain, Copyright & IP	Marketplace
Filebase	US	2019	Multi-purpose	Individual
Arweave	UK	2017	Multi-purpose	Enterprise
Filecoin	US	2014	Commerce	Marketplace
Sia.Tech	US	2013	Multi-purpose	Marketplace
Storj	US	2014	Multi-purpose	Enterprise
Ionian (part of Ionian Network)	US	2022	Multi-purpose	Enterprise

OrbitDB (part of Protocol Labs)	US	2017	Multi-purpose	Individual
Aleph.im	France	2020	Media, Multi-purpose	Individual
Swarm	Switzerland	2019	Multi-purpose	Marketplace
DB3Network	N/C	2022	Multi-purpose	Individual
Zata.Network	N/C	2023	Multi-purpose	Marketplace
Dolpin (part of Dosier)	India	2022	Multi-purpose	Individual
Space and Time (part of SpaceandTime Labs)	US	2022	Finance, Gaming, Commerce	Enterprise
Arcana Network (part of Arcana Technologies)	India	2021	Multi-purpose	Marketplace
Ocean Protocol	Singapore	2018	Multi-purpose	Marketplace

Annex C

List of interviewees

ID	Role	Profile	Type of solution
1	CEO	Vendor	Marketplace
2	Co-Founder	Vendor	Enterprise
3	Co-Founder & Growth	Vendor	Enterprise
4	Technology Executive	Vendor	Enterprise
5	CBO	Vendor	Individual
6	Software Engineer	Vendor	Individual
7	CEO	Vendor	Marketplace
8	CDO	Vendor	Marketplace
9	Technical Advisor	Vendor	Marketplace
10	Head of Architecture	User	N/A
11	Digital Health and Care Strategic Advisor	User	N/A