

THE TECHNOLOGY LANDSCAPE OF DATA SPACES

Antti Poikola

Leading Specialist
Sitra

P J Laszkowicz

Advisor
Futurice

Ville Takanen

Head of Tech Advisory
Futurice

Teemu Toivonen

Principal Consultant Data & AI
Futurice

Data spaces are soft digital infrastructures that enable reliable and easy data sharing across organisational boundaries. In this emerging domain implementation technologies and architectures are evolving rapidly. This working paper provides a snapshot of the fast-moving technology landscape and recommendations to support data space builders. There are three main directions that data space builders can look to when making technology choices: existing tools and solutions (not specific to data spaces), data space technology initiatives, and commercial offerings focused on data spaces. It is recommended that data space builders test initial use cases with existing mature tools before committing to data space-specific solutions. However, working closely with one or more data space technology initiatives will allow the data space builder to better understand the development of the domain.

Sitra working paper

© Sitra 2023

The Technology Landscape of Data Spaces

Sitra: Antti Poikola

Futurice team: P J Laszkowicz, Ville Takanen and
Teemu Toivonen

Layout: PunaMusta Oy

ISBN 978-952-347-327-0 (PDF) www.sitra.fi

ISSN ISSN 2737-1042 (online publication)

Sitra working papers provide multidisciplinary information about developments affecting societal change. Working papers are part of Sitra's future-oriented work conducted by means of forecasting, research, projects, experiments and education.

Contents

Foreword	4
Summary	5
Tiivistelmä	6
Sammanfattning	7
1. Data spaces as an emerging technology domain	8
1.1. Core concepts for understanding data spaces	8
1.2. Studying the emerging technology landscape	10
2. Snapshot of the data space technologies	11
2.1 Existing tools and solutions (not specific to data spaces)	11
2.2 Data space technology initiatives	13
2.3 Commercial offerings focused on data spaces	14
2.4 User experience and trust	15
3. Recommendations for data space builders	16
3.1 Recommendation: Support actors with multiple roles in data spaces	16
3.2 Recommendation: Test business cases against existing mature solutions	16
3.3 Recommendation: Monitor the market and provide feedback	16
3.4 Recommendation: Pay attention to user experience	16
Glossary	17
Literature	18
Appendix 1. Organisations depicted in the landscape diagram	19
Appendix 2. Assessment of key data space technology initiatives	21

Foreword

Europe has an ambitious plan for its data economy. The goal is to create a pan-European single market for data, where data flows across company and national borders in the same way that people and goods do today. In its data strategy, the European Commission estimates that the value of the Europe's data economy will grow to EUR 829 billion by 2025. That was the estimate before the pandemic accelerated the digitalisation of our lives.

At Sitra, we believe that the current platform economy model is unfair. The benefits and the value of the data economy are only accumulated by a handful of large companies. The people and companies who use digital services do not control the data they upload or that the digital services gather from their actions and behaviour. The market does not readily welcome new players or innovations. Sitra is working towards a fair data economy where individuals, companies and data rights holders have more control. The aim is to create a level playing field and offer alternatives that benefit everyone. The goal is to provide better digital services that make everyday life easier without compromising privacy.

The European data strategy suggests that our data economy will soon be boosted by the exchange of data between all interested companies and organisations using a new data-sharing approach called data spaces. What are data spaces and for what business and data-sharing challenges do they offer solutions? How would the governance of data flows work in these new environments? If you want to build a data

space, how do you go about it? These are some of the questions we want to understand at Sitra. As the Finnish Innovation Fund, our goal is to ensure the competitiveness of Finnish companies in the European data economy. We are happy to share our experiences with the rest of Europe and beyond.

In this study, we delve into the technical aspects of building data spaces. We asked our technology partner to assess the current options for building the digital infrastructure needed for data space. The results were inspiring.

Data space-specific technologies are maturing at the level of common specifications, open source components and commercial offerings. However, a key finding is that there is no need to wait for these technologies to mature. For those who want to be the first movers in data spaces, existing enterprise solutions offer a path to secure and federated data sharing in line with the principles of data spaces. However, interoperability between different technology solutions – an essential requirement for European data spaces – is a challenge that remains to be overcome.

The first technological building blocks are in place, the fastest moving ecosystems are on the move, and the first implementations of the next chapter of the European data economy are taking place. For early adopters, now is the time to share data and build data spaces.

Anssi Komulainen

Project Director, Gaia-X Finland, Sitra

Summary

The amount of digital information, or data, is constantly growing, and managing it will be increasingly important for organisations. The biggest new opportunities in the data economy are related to co-operation between companies and organisations. No single company can meet all the needs of its customers, but in flexible business ecosystems, companies can work together to create seamless services for end users.

Sharing and combining data between organisations is a prerequisite for such seamless services. Reliable data sharing between different organisations requires a so-called soft infrastructure, meaning new types of technical, administrative and business solutions. The data space is a soft infrastructure that enables reliable and easy data sharing across organisational boundaries.

Implementation technologies and architectures for data spaces are being developed at a fast pace in Europe. This working paper provides an overview of the developing technology field and recommendations to support data space builders. While this working paper helps to make technology choices, the [Fair Data Economy Rulebook](#) published by Sitra is a tool for creating a governance model for data spaces.

While the individual standards and technologies related to data spaces are moving fast,

the overall structure of the landscape is a bit more static. There are three main directions that the data space builders can look to when making technology choices: 1. existing tools and solutions (not specific to data spaces), 2. data space technology initiatives, and 3. commercial offerings focused on data spaces. This working paper provides a starting point from which data space developers can learn more about the technology offerings in this area.

Practical business cases of inter-organisational data sharing can often be addressed using existing tools and technologies. This opens up the possibility of gradually embracing the concept of data spaces, using existing technology expertise and tools, before moving on to data space-specific solutions.

The experts we interviewed also emphasise the role of design and user experience in building trust in data spaces. Perceived trust in data spaces is key, especially for data rights holders of data rights to engage and dare to share data with other actors. Trustworthiness is a high priority for data space technologies, but even the best technology alone is not enough to achieve trust, if the user experience is not at a similar level. In addition to technology, legal design and service design are also needed for data space developers to gain the trust of users in data sharing.

Tiivistelmä

Datan määrä kasvaa jatkuvasti ja sen hallinnasta tulee organisaatioille entistä tärkeämpää. Datatalouden suurimmat mahdollisuudet liittyvät yritysten ja organisaatioiden väliseen yhteistyöhön. Yritysten on haastavaa vastata asiakkaidensa kaikkiin tarpeisiin yksinään, mutta joustavissa liiketoimintaekosysteemeissä yritykset voivat toimia yhdessä ja luoda saumattomia palvelukokonaisuuksia asiakkailleen.

Saumattomien palveluiden edellytyksenä on datan jakaminen ja yhdistäminen organisaatioiden välillä. Luotettava datan jakaminen vaatii toteutuakseen niin sanottua pehmeää infrastruktuuria, joka tarkoittaa uudenlaisia teknisiä, hallinnollisia ja liiketoiminnallisia ratkaisuja. Data-avaruudella tarkoitetaan sellaista digitaalista tietoa-alueita, joka mahdollistaa luotettavan ja helpon datan jakamisen yli organisaatorajojen.

Euroopassa kehitetään nopeasti data-avaruuksien toteutustekniikoita ja data-arkkitehtuuria. Tässä työpaperissa otamme katsauksen alan kehittyvään teknologiakenttään ja annamme suosituksia data-avaruuksien kehittäjien tueksi. Tämä raportti tukee kehittäjiä teknologiavalintojen tekemisessä, kun taas Sitran julkaisema Reilun datatalouden sääntökirja tarjoaa työkaluja data-avaruuksien hallintomallien luomiseen.

Data-avaruuksiin liittyvät yksittäiset standardit ja teknologiat kehittyvät nopeasti, mutta kehityksen kokonaiskuva ja pääsuunnat ovat pysyvämpiä. Tässä raportissa tunnistamme kolme pääsuuntausta, joita data-avaruuksien

rakentajat voivat hyödyntää teknologiavalintoja tehdessään: 1. olemassa olevat työkalut ja ratkaisut (jotka eivät koske erityisesti data-avaruuksia), 2. data-avaruuksia koskevat teknologialoitteet ja 3. kaupalliset data-avaruusratkaisut. Työpaperi tarjoaa data-avaruuksien kehittäjille lähtökohdan sekä lisätietoja alan teknologia-tarjontaan liittyen.

Organisaatioiden väliseen datan jakamiseen liittyvät käytännön liiketoimintatapaukset voidaan usein ratkaista jo olemassa olevien välineiden ja teknologioiden avulla. Tämä mahdollistaa sen, että data-avaruusteknologioiden käyttöön voidaan siirtyä vaiheittain, olemassa olevaa teknologiaosaamista ja välineitä hyödyntäen, eli jo ennen kuin siirrytään käyttämään varsinaisia data-avaruusratkaisuja.

Raporttia varten haastattelemamme asiantuntijat korostavat myös suunnittelun ja käyttäjäkokemuksen merkitystä, kun pyrkimyksenä on rakentaa luottamusta data-avaruusratkaisuja kohtaan. Koettu luottamus data-avaruuksia kohtaan on avainasemassa erityisesti siinä, että dataoikeuksien haltijat sitoutuvat ja uskaltavat jakaa dataa muiden toimijoiden kanssa. Luotettavuus onkin data-avaruusteknologioiden kannalta tärkeää, mutta parhainkaan teknologia ei yksinään riitä luottamuksen saavuttamiseen, jos käyttäjäkokemus ei ole luottamusta vastaavalla tasolla. Teknologian lisäksi tarvitaankin myös oikeudellista muotoilua sekä palvelumuotoilua, jotta data-avaruuksien kehittäjät saavuttaisivat käyttäjien luottamuksen datan jakamiseen.

Sammanfattning

Mängden data ökar ständigt och att hantera den blir allt viktigare för organisationer. De största möjligheterna inom dataekonomin ligger i samarbetet mellan företag och organisationer. Det är en utmaning för företag att på egen hand tillgodose alla sina kunders behov, men i flexibla affärsekosystem kan företag samarbeta för att skapa sömlösa tjänstepaket för sina kunder.

Sömlösa tjänster kräver delning och integrering av data mellan organisationer. Tillförlitlig datadelning mellan olika organisationer kräver en så kallad mjuk infrastruktur, vilket innebär nya tekniska, administrativa och affärs-mässiga lösningar. Ett datalager är ett digitalt informationsutrymme som möjliggör tillförlitlig och enkel delning av data över organisationsgränser.

Teknik för implementering av datalager och data arkitekturer utvecklas snabbt i Europa. I det här arbetsdokumentet går vi igenom det framväxande tekniklandskapet på det här området och ger rekommendationer för att stödja utvecklare av datalager. Rapporten hjälper utvecklare att göra teknikval, medan regelboken för rättvis dataekonomi som publicerats av Sitra tillhandahåller verktyg för att skapa styrmodeller för datautrymmen.

Enskilda standarder och tekniker relaterade till datalager utvecklas snabbt, medan den

övergripande bilden är mer statisk. I rapporten identifierar vi tre huvudtrender som de som bygger datalager kan använda när de väljer teknik: 1. befintliga verktyg och lösningar (inte specifika för datalager), 2. tekniska datalagerinitiativ och 3. kommersiella datalagerlösningar. Arbetsdokumentet ger en utgångspunkt för utvecklare av dataområden och ytterligare information om teknikutbudet inom området.

Praktiska affärsfall för datadelning mellan organisationer kan ofta lösas med hjälp av befintliga verktyg och tekniker. Detta möjliggör en gradvis övergång till användning av dataområdesteknik, som bygger på befintlig teknisk kunskap och verktyg, dvs. till och med innan de faktiska dataområdeslösningarna tas i bruk.

De experter som vi intervjuade för den här rapporten betonar också vikten av design och användarupplevelse för att bygga upp förtroende och tillit till dataområdeslösningar. Det upplevda förtroendet för dataområden är avgörande, i synnerhet för rättighetsinnehavarnas engagemang och förtroende för att de skall våga dela data med andra aktörer. Förtroendet är därför ytterst viktigt för teknik för dataområden, men även den bästa tekniken kommer inte att vara tillräcklig för att uppnå förtroende om användarupplevelsen inte står i proportion till förtroendenivån.

1. Data spaces as an emerging technology domain

Seamless data flow between organisations enables better products and services and creates tremendous potential for increasing work productivity. Multiple stakeholders collaborating and sharing data are often referred to as a data ecosystem. The essential starting points for the development of such data ecosystems are shared rules and visions for cross-organisational data use and the first concrete use cases. Reliable data sharing between parties requires a soft infrastructure such as common standards and practices, architectures, and a governance framework. Data spaces are implemented by this sort of soft infrastructures.

Data spaces support the ongoing business transformation in which many organisations are beginning to see data more as a product and to produce it with reuse in mind. Data products are gaining widespread traction in enterprise data architectures within organisations. Data spaces offer the next step whereby data products can be distributed and used in cross-organisational data ecosystems.

This chapter provides a small set of key concepts to help readers understand the relationship between the data space as an infrastructure, value-generating use cases, data products deployed in use cases, and technology that enables the implementation of data spaces in practice.

1.1. Core concepts for understanding data spaces

The concept of data space is evolving and the term has slightly different definitions in different contexts (see Literature). While there are different definitions of data spaces, they all share the same basic objective – to facilitate trusted data flows fairly and transparently for the parties involved in data sharing. In data spaces, individuals and organisations, as data rights holders, are in the driver's seat, deciding

who can use their data and on what terms. By comparison, in more centralised and traditional data platforms, the decision-making power is in the hands of a few. The benefits also often accumulate more for the platform owner.

This working paper uses the terms and definitions from the [Data Spaces Support Centre](#) glossary. Sitra is a partner of the DSSC, a project funded by the European Union.

The DSSC glossary defines data spaces as follows:

A distributed system defined by a governance framework that enables secure and trustworthy data transactions between participants while supporting trust and data sovereignty. A data space is implemented by one or more infrastructures and enables one or more use cases.

A data space use case is a potential value-creating usage scenario for the data provided through the data space. The use cases generate demand for the data space infrastructure that supports the cost-efficient implementation of these use cases. A data space provides a shared infrastructure for one or more use cases. They can speed up the development of

use cases in a given domain (such as logistics, tourism, skills) because the use cases often require the same data sources. If the data needed by one use case is productised (instead of tailored to a single use case), the data products can be used directly in subsequent use cases. Ready-made data products also facilitate the identification of new use cases and the development of innovations.

Data spaces can facilitate the trusted data flows required by the use cases because all participants of the data space are committed to the common rules encoded in the data space governance framework. The participants associated with a use case can fulfil the role of data rights holders, data providers, data recipients or data users. The same actors typically participate in data spaces in multiple roles. Also, having a role in one data space may create opportunities to have another aligned role in another data space. For example, one company we interviewed had identified a business opportunity to work in the energy data space because it was already a data provider in another data space

for the types of data needed by data users in the energy data space.

Some data space participants (such as data intermediaries, identity providers) provide services that enable data transactions for others while not directly participating in those transactions. The governance authority is the party that develops, maintains, and enforces a data space's governance framework.

The data space initiative is a collaborative effort of committed partners to implement and maintain a data space. The data space initiative provides common parts of the infrastructure that are shared by all the participants. Common components can be, for example, a clearing house, an identity broker, a data catalogue, and suchlike. Data spaces are distributed in the sense that all data space participants are autonomous and individually implement or acquire the technologies that are necessary for them to operate within the data space. However, the requirements as to which technologies and standards the data space participants must follow are defined as part of the data space governance.

Case: Maritime Data Space Finland

Port congestion and maritime transport emissions challenge the global shipping industry. The movements of cargo ships generate a constant stream of data that could be used to overcome many logistical problems. Ships often rush at full speed to get near the mainland to wait for the port to confirm a berth. Unnecessary emissions and fuel costs resulting from the moving full speed at sea and waiting times near ports could be reduced if cargo ships had systems to share traffic data with other operators, such as the ports, shipping companies and other vessels.

The maritime data space in Finland is a data space initiative co-funded by Sitra, where data space participants are looking for ways to reduce maritime transport congestion through the use of data. The coordinator and governance authority for the maritime data space is Fintraffic, a Finnish state-owned company that provides traffic control and management services for all forms of transport.

1.2. Studying the emerging technology landscape

Data spaces are an emerging domain where implementation technologies and architectures are evolving rapidly. The emergent nature of the domain makes it hard to make solid predictions or recommendations – as ‘winner takes it all’ is often a truism in the battle between technologies and standards. The business cases are still evolving, and the technology implementations will vary for some time. From a technology perspective, data spaces’ distributed nature supports individual and rapid innovation and market-driven consensus of practices and protocols.

We used a two-fold strategy in order to gain an understanding and evaluate the maturity of available and used technologies in this rapidly evolving field. First, we conducted a desktop study based on the openly available material and developed an initial idea of the data spaces technology landscape. We then interviewed experts from pioneering organisations to gather

their real-world experience in building data spaces and to test the initial ideas from the desktop study. During the desktop study, we identified many technology projects, support organisations, and commercial providers related to the area of data space technology. Based on the interviews, we selected some of these for more detailed analysis. The expert interviews also helped us develop recommendations for organisations willing to join the pioneers in providing data space services or solutions.

We interviewed experts from three organisations involved in the development of data spaces: Mtech Digital Solutions Oy, which is a solution provider for the Finnish food supply chain, Agdatahub, which is a French Agricultural Data Space, and Fit Our Future, a Dutch energy sustainability consultancy. We also interviewed three companies having commercial offerings focused on data spaces: Dataspace Europe, Nexyo Io, and Sovity. EU-funded data space projects were not included in the interviews.

2. Snapshot of the data space technologies

It is a challenge for data space builders to keep track of all the potentially useful technologies. Each data space has different technology requirements, depending on the business case, governance choices and applicable regulations. Best practices for technology choices are also subject to change as the field matures. This chapter provides a snapshot of the rapidly changing technology landscape as a starting point and initial reference for data space builders.

While the individual standards and technologies associated with data spaces are moving quickly, the overall structure of the landscape is somewhat more static. The landscape image below shows three main directions that people making technology choices for data space initiatives can consider:

1. Existing tools and solutions (not specific to data spaces),
2. Data space technology initiatives, and
3. Commercial offerings focused on data spaces. We follow these three directions to give data space developers a perspective on which technologies to watch for – or adopt for a business case. This study does not cover the general-purpose enterprise data architecture components in detail. The organisations depicted in the landscape diagram are listed in Appendix 1.

2.1 Existing tools and solutions (not specific to data spaces)

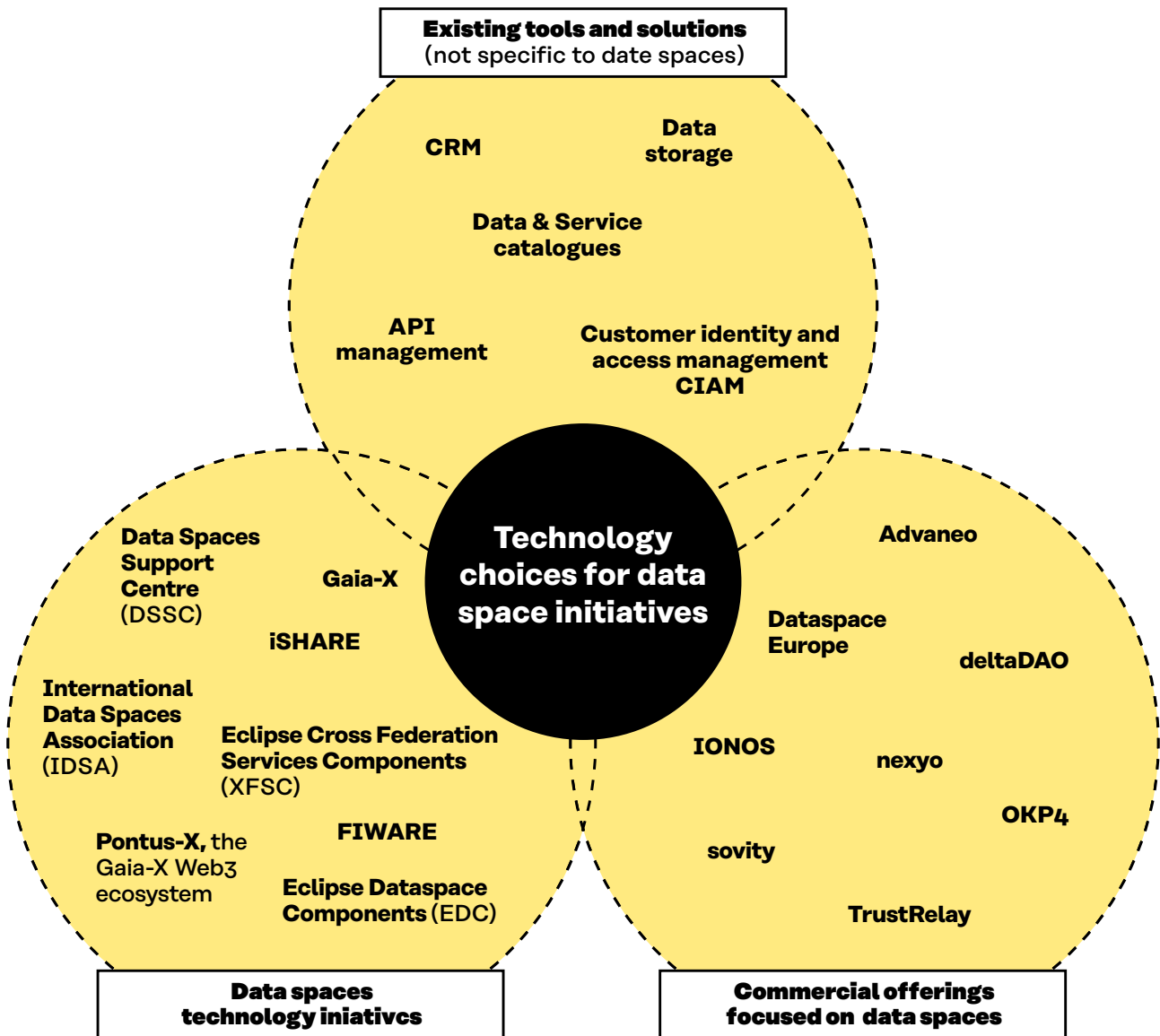
Testing business cases iteratively in an emerging technology domain, such as data spaces, is easier than finding a working end-to-end busi-

ness case. Most innovations in the data spaces relate to legal-, service-, and business design. These can often, if not always, be tested using standard and existing technologies such as customer identity and access management (CIAM), CRM, data storage, API management, data- and service catalogues, etc.

Data space builders should be pragmatic about standards and try to use existing solutions outside the data space domain to test the use cases before committing to development efforts. As with a start-up company: explore the business cases while simultaneously developing the technology at a rapid pace.

The development of data spaces and inter-organisational data sharing should be linked to the organisation's internal data architecture solutions. In this light, the data mesh is a promising modern approach to building a distributed data architecture for enterprises with many similarities to data space thinking. The data mesh can be seen as a miniature example of how a data space might work within an organisation. The primary value proposition of the data mesh is to reduce the cost of data engineering and analytics within the organisation.

Figure 1. The Data Space Technology Landscape consists of three directions.



Data products, data space, and data mesh

Data meshes are gaining widespread ground in enterprise data architectures. Data meshes and data spaces have a lot in common – the data meshes focus on data management within organisations and data spaces focus on data management across the organisational borders. One of the fundamental principles of data meshes is the perception of data as a product.

Viewing data as a product is a mindset change in organisations – the notion of starting to generate data with reuse in mind. This addresses the root cause of many problems concerning traditional approaches where business processes generate data without designing it to be shared. Productised data can be easily consumed, even by users who had no initial connection to the source of the data.

Data spaces are a distributed and standards-based approach to enabling data sharing and use across organisations while overcoming some of the problems encountered in centrally owned data platforms.

Data mesh is a modern approach to building a distributed data architecture for enterprises. It has four principles: domain ownership, data as a product, self-serve data platform, and federated (and computational) governance. Zhamak Dehghani coined the term data mesh in 2019. It borrows ideas from domain-driven design and builds on software paradigms that encourage agile, functional teams with autonomy and responsibility. The data mesh is an in-company miniature example of how a data space ecosystem could work.

The data mesh and data space can converge to create a more holistic data management paradigm that builds on top of data products and encompasses internal and external data sharing and use. Data governance, authorisation and connection capabilities of data spaces complement the data mesh capabilities. The idea of a data marketplace thinking that covers internal use while transcending organisational boundaries can be a significant driver for data sharing and the adoption of data spaces.

It is also possible to compare data spaces and meshes to forecast the market development of these areas. Although the data mesh is new and its practical implementation is still growing, there is a lot of excitement about it in the industry. Many IT firms have jumped on board to promote their data mesh capabilities, and consultancies are promoting how they can help companies on their data mesh transformation journeys. In data spaces, the commercial offering is just beginning to emerge. Internal data sharing in businesses drives the demand for data mesh solutions. As companies mature with their internal data capabilities, the next step is to focus on sharing data across organisations. That would create demand for the data spaces in combination with data meshes.

2.2 Data space technology initiatives

The data space builders can implement the first business cases with existing mature technologies outside the data space domain. At the same time, the data space builders should pay close attention to data space-specific technol-

ogy frameworks that are maturing and can be used alongside more generic tools. According to the experts interviewed, it is essential for the long-term success of the whole domain of data spaces that the early data space builders work closely together and provide constant feedback to the projects striving for standardisation.

Several players in the data space technology domain are working towards data space-specific standards and common reference technology frameworks. When one starts learning about data spaces, these names soon pop up: International Data Spaces Association (IDSA) Gaia-X, iSHARE, Eclipse Cross Federation Services Components (XFSC), Gaia-X Web3 Ecosystem (Pontus-X), Eclipse Dataspace Components (EDC) and FIWARE. Knowledge of these data space technology initiatives gives a good basis for evaluating others. To give the data space builders a starting point for their explorations, we have made an initial effort to assess the maturity, adoption, and potential of some key data space technology initiatives (Appendix 2).

These initiatives are not directly comparable alternatives to each other. They contribute to the data space technology landscape at different levels, from reference architectures to trust frameworks and open source components. IDSA specifies a reference architecture (IDS Reference Architecture Model), parts of which are implemented by EDC, FIWARE and other IDS connector providers (IDSA connector report). Gaia-X also specifies a reference architecture (Gaia-X Architecture model) that to which GXFS-DE, Pontus-X, and to some extent EDC and iSHARE are aligned. Together they form a mesh of intertwined and co-evolving initiatives that are advancing data space technologies.

Technical convergence refers to the integration of previously separate technologies, functionalities, or standards, resulting into a cohesive framework. Technical convergence is underway across the data space technology initiatives. A collaborative forum, the Data Spaces Business Alliance (DSBA), is working towards a common reference technology framework based on the technical convergence of existing architectures and models from Gaia-X, IDSA and FIWARE. This collaboration aims to achieve the interoperability and portability of solutions across data spaces by harmonising technology components and other elements. In a broader context, an EU-funded project, the Data Spaces Support Centre (DSSC), will also contribute to the technical convergence by

analysing and recommending existing technologies and providing guidance to data space builders through a common blueprint.

A key finding from the interviews is the positive correlation between developer experience and the adoption of a framework or a technology. A tool or technology with a good developer website, relevant open source components and active feedback channels will gain more traction in the market than solutions that lack any of these. In terms of developer experience, some data space technology initiatives are further along than others, but none are yet outstanding. For example, documentation may be technically available, but the usability is not at the level that would drive developer adoption. The documentation needs to be more digestible for developers and accompanied by specific examples of how to use the services and concepts. A prominent way to support the adoption is the network of national hubs that Gaia-X, IDSA and FIWARE have.

2.3 Commercial offerings focused on data spaces

An important part of the technology landscape is the available market offering of commercial data space -tools and -services. The market development of commercial technology services and solutions indicates the level of maturity of the data space domain.

The reference architectures mentioned earlier (IDS and Gaia-X) continue to evolve, challenging implementers. Implementations may be better aligned with the older versions of the reference architectures. On the other hand, implementations also push the architectures forward. The data space builders should have proper versioned planning in place. Commercial offerings focused on data spaces can provide valuable support in dealing with the evolving versions of the architectures and software.

Currently, a small but steadily growing number of companies are focusing primarily on data spaces or are launching data space-specific products and services as part of a larger port-

folio. Several players are offering a form of “data space as a service” solution, which enables a full data space to be set up with fewer technical hurdles. As part of this study, we contacted the following commercial data space providers: Advaneo, Dataspace Europe, deltaDAO, IONOS, nexyo, OKP4, soivity and TrustRelay (Appendix 1).

2.4 User experience and trust

During the interviews, we identified an important challenge for the development of data space technologies that is not directly part of the technology landscape: how to create a sufficiently good user experience that it conveys the trustworthiness of the data space.

Trust in digital services refers to the confidence and reliance that users place in the reliability, security, privacy and ethical practices of technologies. It includes the belief that digital service providers will act in the best interests of users, protect their data and privacy, and deliver on their promises. However, trust and trustworthiness do not always go hand in hand when it comes to digital services. While trust is the belief that a service provider will act in the best interests of the user, trustworthiness is the demonstrated ability of the provider to consistently fulfil those expectations. In some cases, users may initially trust a digital service based on brand image and good user experience, only to find that the provider fails to deliver on its promises. This disconnect between trust and trustworthiness can undermine user confidence and lead to scepticism about digital services as a whole.

Many data space initiatives face a common challenge: attracting data rights holders and gaining their trust. Typically, the business case and motivation for sharing data across organisa-

tions come from those who would use the data. The viability of these business cases depends on the willingness of data rights holders to share data. In order to release their data, rights holders should be confident that it will not be misused or exploited. As many companies have tried to misuse and monetise crowdsourced or scraped data without explicit consent, the utility and safety of data spaces can be hard to communicate to potential data rights holders.

This creates a two-fold problem for users:

1. Potential data rights holders and data providers want to avoid sharing or releasing their data for fear of exploitation.
2. Existing large-scale data users have solved the first problem, without data spaces, through excellent onboarding, desirable features, and other means to cater to people and companies regardless of their fears and doubts.

The primary vision of data spaces is to address the first problem by providing trustworthy mechanisms for data sovereignty and the fair use of data. In data spaces, a fair share of the value generated by the data user should be distributed to the data rights holders. In data spaces, there should be means to track and monitor data transactions and enforce policies to prevent exploitation. Based on the basic design principles, the data space technologies would be technically immune to the first problem arising. However, as the trust is often already lost, the players in the fair data economy will need to adopt the user experience and tools that their competition in the legacy markets are already using. To mitigate the above issues, we recommend focusing on legal experience design and user experience design for data space participants in all roles.

3. Recommendations for data space builders

From the interviews, the desktop study and the assessment of key technologies, we were able to derive a number of recommendations to provide a starting point for data space builders to explore the domain further.

3.1 Recommendation: Support actors with multiple roles in data spaces

The same actors typically participate in data spaces in multiple roles. Therefore, the requirements for the different roles (data rights holders, providers, recipients or users) should be harmonised as far as possible. The primary focus of development should be to ensure that the actors can take on different roles and participate in other data spaces using the same tools and set of technologies.

3.2 Recommendation: Test business cases against existing mature solutions

A data space business case should be tested using existing tools and solutions wherever possible and new data space-specific tools should only be adopted if existing options are insufficient. Most innovations in the data spaces relate to legal, service and business design. These can often be tested using existing technologies such as customer identity and access management (CIAM), CRM, data storage, API management, data- and service catalogues, etc.

3.3 Recommendation: Monitor the market and provide feedback

After the initial use cases, it may become relevant to choose a technology framework specific to data spaces. There are currently no clear winners among the data space technology initiatives, so the most rational option for data space builders is to use the framework which delivers the fastest business value with the least investment. This will depend on the business case. By working closely with one or more data space technology initiative, the data space builder can better understand the domain and contribute to its development with feedback.

3.4 Recommendation: Pay attention to user experience

To gain traction and participants (especially the data rights holders), the user experience of the data space needs to match or exceed the level offered by existing data platforms. With this in mind, a good product owner who can funnel stakeholder needs into a meaningful backlog for developers will be essential. The role of the product owner is crucial in channelling the needs of potential data rights holders and other data space participants into the design of the data space.

Glossary

Data space is a distributed system defined by a governance framework that enables secure and trustworthy data transactions between participants while supporting trust and data sovereignty. A data space is implemented by one or more infrastructures and enables one or more use cases.

Data space initiative is a collaborative project of a consortium or network of committed partners to initiate, develop and maintain a data space.

Data space governance authority is the data space participant that is accountable for creating, developing, operating, maintaining and enforcing the governance framework for a particular data space, without replacing the role of public enforcement authorities.

Data space participant is a party that has committed to the governance framework of a particular data space and may have one or more roles in it.

Data space use case is a specific setting in which two or more participants use a data space to create value (business, societal or environmental) from data sharing.

Data transaction is the result of an interaction between two participants with the purpose of sharing, accessing, exchanging or processing data.

Data rights holder (role) is party that has (legal) rights and/or obligations to use, grant access to or share certain personal or non-personal data. Data rights holders may transfer such rights to others.

Data provider (role) is a transaction participant that, in the context of a specific data transaction, technically provides data to the data recipients that have a right or duty to access and/or receive that data.

Data recipient (role) is a transaction participant to whom data is, or is to be technically supplied by a data provider in the context of a specific data transaction.

Data user (role) is a natural or legal person who has lawful access to certain personal or non-personal data and has the right, including under Regulation (EU) 2016/679 in the case of personal data, to use that data for commercial or noncommercial purposes (DGA Art.2)

Data space enabling service is a compulsory or optional core function of a data space that enables data transactions for the transaction participants and/or data space operations for the governance authority. Examples of enabling services include identity, observability, catalogue, membership management, and connector services.

Data space intermediary (role) is a data space participant that provides one or more enabling services while not directly participating in the data transactions itself.

Data product is a standardised data unit packaging relevant data resources and services into a consumable form complying with data product specifications.

Source: The Data Spaces Support Centre (DSSC) Glossary 2.0

Literature

BDVA. 2019. Towards a European Data Sharing Space. (accessed 26 June 2023).

Curry E., Scerri S., Tuikka T. 2022. Data Spaces: Design, Deployment, and Future Directions.

DSSC. 2023. Starter Kit for Data Space Designers. Data Spaces Support Centre (DSSC). (accessed 2 July 2023).

DSSC. 2023. Data Spaces Blueprint Version 0.5. Data Spaces Support Centre (DSSC). (accessed 26 October 2023).

DSSC. 2023. Glossary 2.0. Data Spaces Support Centre (DSSC). (accessed 26 October 2023).

EC. 2022. Staff working document on data spaces. European Commission. (accessed 26 June 2023).

EHDS. 2022. European Health Data Space (website). European Commission. (accessed 26 June 2023).

Nagel L., Lycklama D. 2021. Design Principles for Data Spaces. Position Paper. Version 1.0. (accessed 26 June 2023).

Otto B., Hompel M., Wrobel S. 2022. Designing Data Spaces: The Ecosystem Approach to Competitive Advantage.

Pitkänen O, Luoma-Kyyny J. 2022. Rulebook for a fair data economy. Sitra.

Steinbuss, S. et al. 2023. Data Spaces Landscape – Overview and relations of data spaces initiatives, standards, and tools (1.0). International Data Spaces Association. (accessed 26 June 2023).

Appendix 1. Organisations depicted in the landscape diagram

Data space technology initiatives

Data Spaces Support Centre (DSSC) is a project funded by the European Commission as part of the Digital Europe Program. The DSSC explores the needs of data space initiatives, defines common requirements, and establishes best practices to accelerate the formation of sovereign data spaces as a crucial element of digital transformation in all areas.

International Data Spaces Association (IDSA) provides a reference architecture that enables an ecosystem for the sovereign data exchange with clearly defined usage rights.

Gaia-X aims to establish an ecosystem of data spaces where data is shared in a trustworthy environment so that users retain control and data sovereignty. It is developing the Gaia-X technical framework, compliance scheme, and open source software reference implementations. See the [documentation](#) and [repositories](#).

iSHARE is a European trust network for international and sovereign business data sharing managed by the iSHARE Foundation. The iSHARE trust framework enables the federated trust governance of data spaces. It provides components of data spaces in line with the [design principles for data spaces](#) from the [Open DEI -project](#), International Data Spaces Association and Gaia-X. See the [documentation](#) and [repositories](#).

Eclipse Cross Federation Services Components (XFSC) is an open source software project developing basic components necessary to set up federated data sharing systems. Before transferring to the Eclipse foundation the project was known as Gaia-X

Federation Services (GXFS). See the [documentation](#) and [repositories](#).

Pontus-X, the Gaia-X Web3 ecosystem, under the governance of the Gaia-X member institutions, aims to provide decentralised and federated approach to data management, allowing data, software, infrastructure and federation services to be securely built, collated, shared, and monetised. See the [Gen-X](#), [Ocean protocol](#) and [Polygon](#) documentation and the [deltaDAO](#), [Ocean protocol](#) and [Polygon](#) repositories.

Eclipse Dataspace Components (EDC) is an open source project that aims to implement the International Data Spaces standard (IDS) and relevant protocols and requirements associated with the Gaia-X and thereby provide implementation and feedback to these initiatives. See the [documentation](#) and [repositories](#).

FIWARE open source technology is being used to develop Smart Solutions, Digital Twins and Data Spaces in several domains of digital transformation. See the [documentation](#), [repositories](#) and the [marketplace](#).

Commercial offerings focused on data spaces

Advaneo offers a one-stop-shop solution for participating in data spaces. Companies can easily use their components to build innovative business models and products. Their Data Marketplace contains around 2.5 million datasets, an AI workbench and a hackathon solution for open innovation. Their Trusted Data Hub allows sensitive data to be used without exposing the raw data. Data Catalog, Data Marketplace and Trusted Data Hub provide the infra-

structure for sovereign data sharing through data spaces solution building blocks.

Dataspace Europe provides the Triton data intermediation service to enable the collaborative use of data and improve the operational capabilities of industry players.

deltaDAO built the Gaia-X Web3 Ecosystem “Pontus-X” based on the Ocean Protocol and Distributed Ledger Technology in 2021. deltaDAO enabled the first instant settlement layer to consume data, software, and infrastructure services in Gaia-X with Euro. deltaDAO was the first to integrate the Gaia-X Trust Framework.

IONOS is a European cloud provider that offers its customers the automated provisioning of data space connectors and components on its cloud, ensuring seamless integration and sovereign management of their data.

nexyo DataHub connects decentralised data sources via EDC connectors and offers additional services that allow non-technical users to be part of the emerging data ecosystems or to build ecosystems themselves. The goal is to enable cross-company and cross-industry innovation for data-driven business models while preserving data autonomy and sovereignty.

OKP4 is a public layer-1 blockchain designed to coordinate digital assets such as datasets, algorithms, software, storage or computation. Anyone can create and join custom data spaces where rules are shared and value flows between participants.

sovity gives companies access to data sovereignty, allowing them to create new data-driven business models and build innovative products based on data space technologies. With its holistic and easy-to-use software, Connector-as-a-Service customers can easily participate in data ecosystems, sharing data while maintaining full control.

TrustRelay is a SaaS solution that enables corporates to collaborate on data in a privacy-preserving manner using Confidential Computing and following a Data Mesh approach. With TrustRelay, corporates can share and apply analytics on data without actually centralising the data – easily, securely and in a legally compliant way. The solution facilitates the drafting and signing of so-called “Data Sharing Agreements”, which provide the legal foundation for cross-company data collaboration via Data Spaces.

Appendix 2. Assessment of key data space technology initiatives

As the concept of data space is still emerging, assessing the state of data space technologies is challenging in many ways. Despite the known challenges, we have made a first attempt to assess the maturity, adoption and potential of some key data space technology initiatives: **International Data Spaces Association (IDSA)**, **iSHARE**, **Eclipse Cross Federation Services Components (XFSC)**, **Gaia-X Web3 Ecosystem (Pontus-X)**, **Eclipse Dataspace Components (EDC)** and **FIWARE**. According to the interviews, these are the most relevant data space technology initiatives today that every data space builder should be aware of.

The present assessment is not a comparison of the selected initiatives. As described earlier (Chapter 2), these initiatives are not substitutable alternatives to each other, as they provide assets that are useful for data space builders, but at very different levels: reference architectures (IDSA), trust frameworks (iSHARE), open source frameworks (XFSC, Pontus-X, EDC, FIWARE). Each initiative is different and is assessed on its own merits. This rudimentary assessment gives the data space builders a starting point for their own explorations.

By reviewing the public materials, developer portals and code repositories of these initiatives we estimated their adoption and assessed the quality and amount of documentation they provide. This process mimics the work done to assess the maturity and potential of a technology product for a commercial transaction. We validated the assessment with the results of the interviews. Finally, we sent the results to be fact-checked by the representatives of the assessed initiatives and received a substantial number of links and new information that had not been captured in the original data collection. Brief descriptions of the initiatives assessed and links to the repositories and documentation pages used in the assessment are in Appendix 1.

The overall picture is that the maturity of data space technologies is still evolving. Many initiatives have stable releases in use by multiple actors, active community, and support from some commercial actors, and they will likely be a key part of the offering in the spaces domain also in the future. However, there is still a way to go before these data space technologies are delivered in multiple products and become part of the core internet technologies.

Table 1. Ratings for the data space technology initiatives assessed.

Note that the initiatives are not comparable to each other. The initiatives are rated on their own merits focusing on the benefits they provide to data space builders.

Technology	Maturity	Adoption	(current) Potential
IDSA	3	3	3
iSHARE	3	3	3
XFSC	1	2	2
Pontus-X	2	2	2
EDC	2	3	3
FIWARE	3	3	4

The scale and ratings for maturity, adoption, and current potential.

	Maturity	Adoption	(current) Potential
1	Emerging or beta-phase technology	Used experimentally or in an early phase, startups	Will have limited use within the domain
2	Early release	Used by multiple actors and has an active open-source community	Will likely be used widely for a subdomain, or a type of data
3	Stable release, or releases	Used by multiple actors and has continuous support from a commercial actor and an active community	It will likely be a key part of the domain offering
4	Multiple stable releases over time	Widely supported by commercial actors, part of multiple open-source projects	Will have significant usage outside the domain and be part of the domain offering
5	Multiple stable releases over time by multiple forks	Shipped in multiple products, a key part of current technology	Likely to become part of the core internet technologies

SITRA

SITRA WORKING PAPER 31 October 2023

Sitra working papers provide multidisciplinary information about developments affecting societal change. Working papers are part of Sitra's future-oriented work conducted by means of forecasting, research, projects, experiments and education.

ISBN 978-952-347-327-0 (PDF) www.sitra.fi

SITRA.FI

Itämerenkatu 11–13

PO Box 160

FI-00181 Helsinki, Finland

Tel: +358 294 618 991