

Report

Unlocking the Potential of Data Spaces

Critical Success Factors for Data Space Deployment



Authors

G. van Houwelingen, J. van Bommel,
R. Gilsing, G. Lamerichs, F. Berkens

Februari 2024



Contents

Executive summary	3
1. Introduction: Data spaces defined	5
2. Our recommendations in short	6
3. Findings through interviews	8
3.1. Trust	8
3.2. Organization of the data spaces	11
3.3. Financialization and investments for the data space	13
4. Conclusion	16
5. Acknowledgments	16
6. Referenties	17
Appendix 1 – Methods	18
Appendix 2 – Interview protocol	20
Appendix 3 – Splitting Investment	21
Contact details	22

Sponsor

TNO

Project name

Centre of Excellence for Data Sharing & Cloud

Project number

060.54823/01.05.17

All rights reserved.

No part of this publication may be reproduced and/or published by print, photoprint, microfilm or any other means without the previous written consent of TNO.

In case this report was drafted on instructions, the rights and obligations of contracting parties are subject to either the General Terms and Conditions for commissions to TNO, or the relevant agreement concluded between the contracting parties. Submitting the report for inspection to parties who have a direct interest is permitted.

© 2023 TNO

Executive summary

Parties interested in deploying a data space need not only to use the right type of technologies, they also need to make sure they get the business and governance of the data space right. That is, however, easier said than done: There is relatively little guidance out there how to deploy a data space successfully. Our aim in this whitepaper is to offer such guidance on the business and governance side of successful data space deployment. The whitepaper is meant for anyone or any organization that is involved with an initiative to deploy a data space or interested to become involved in one.

More specifically, we studied several cases of the development of data sharing initiatives in different sectors (i.e., agriculture, mobility, manufacturing) to distill generic business and organizational success and fail factors for data space deployment. Our goal is to inform interested parties about what they need to get right, and what pitfalls they need to avoid, to start a successful data space.

Each data space initiative is, obviously, unique, and, as such, the success (or failure) of the initiative is likely to depend as much on sector- or even initiative-specific factors as on the generic success and failure factors we identify here.

That said, our analysis has led us to three overarching recommendations. Put differently, whatever the sector, and whatever the purpose, data space initiatives need at least to get the following right:

- 1** Developing trust (both informal and formal) is crucial for the deployment and continuation of a dataspace. Stakeholders that distrust each other form a risk to the deployment and continuation of a data space.
- 2** The organization of the data space should be carefully aligned (and realigned) with the objectives of the data space over time. Create a clear (and limited) scope and process to better align with the needs of stakeholders in early development phases, while keeping in mind how the data space can be scaled in later phases once the user base expands.
- 3** Data space deployment calls for an understanding of its investment structure and value creation, as well as the selection of appropriate business models. Identify who are the main beneficiaries of the data space, and select business models which support this narrative to help finance data space development and operation.

Below, we will unpack these recommendations in more detail.

- 1** Developing trust (both informal and formal) is crucial for the deployment and continuation of a dataspace. Stakeholders that distrust each other form a risk to the deployment and continuation of a data space.

Few, if any, data space participants will be willing to exchange data if they distrust other parties or the data space itself. Trust is therefore a crucial ingredient to data space deployment. Trust can be established through informal means (mainly the management of objectives or interests of stakeholders). But it can also be fostered through formalized procedures (i.e. providing structure in terms of governance mechanisms, rules and policies to drive the desired behavior of data space stakeholders).

In the early phases of data space deployments, participants should rely on informal means to build trust mainly. Once the initiative matures and grows, we advocate for adding more formal means (while keeping in place informal practices). More specifically, we recommend to start by investigating the interests and strategic objectives for (future) data space

stakeholders in the exploratory phase of data space development. Subsequently, build informal trust between stakeholders through stakeholder engagement and representation for the data space. Once the data space is operational, formal mechanisms can be used, such as procedures and policies to drive beneficial behavior of stakeholders. The inclusion of a neutral third party driving decision making or acting as a liaison for stakeholders to support decision making can help in overcoming potential challenges faced.

- 2** The organizational aspects of data spaces, such as vision building, scoping of functionalities, and managing entry barriers, should be carefully aligned with the incentives and objectives of the data space and its intended users over time.

Successful data spaces are typically built by stakeholders who share a common vision early on in the development process. Although individual incentives can exist for stakeholders to participate, these should be aligned with the overall objectives for the data space. Misalignment of incentives forms a risk to the success of the data space.

Moreover, it is essential for the data space to achieve adoption with a substantial number of users early on in the operational phase. Without such critical mass, it will be difficult (if not impossible) to realize its intended objectives. This calls for a clear, simple, and shared vision on how the data space will support value creation. Too many functionalities or different use cases in early phases of data space development could lead to unclarity about the value, the investment into the data space will bring. Keeping a narrow focus initially regarding the functionalities and services offered, helps in determining the value to potential end-users.

Furthermore, financial and legal barriers for participation among end-user and other stakeholders should be minimized as much as possible, such that entry barriers are low for new adopters. As the data space is scaled, the vision as well as functionalities included should be evaluated against the (growing) group of data space stakeholders: what direction do we want to go, how do we want to grow in terms of stakeholders.

- 3 Data space deployment calls for an understanding of its investment structure and its value creation logic. This means investigating and

identifying the main beneficiaries for the data space and selecting suitable business models to help finance data space development and operation.

Successful data spaces are worth their participant's while. Simply put, participants need to get some sort of value out of their participation that offsets the time, effort and money that is required for participation. Value here may mean financial or economic value, but it may also mean a less tangible form of value (e.g., public value may be important for public partners). Partners in the consortium that found the data space, need compensation for the investments they have done to set up the data space. Later entrants will need to be asked a fee that fits with their business case.

In other words, appropriate business models to finance the deployment and operation of the data space need to be selected early on. These business models should heavily build upon understanding the business case for each stakeholder involved and on understanding what stakeholders are the main beneficiaries of the data space. For example, in supply chain or logistics, value creation typically resides at the (business) stakeholders involved through improved or automated

data sharing (facilitating improved decision making or operational efficiency). Therefore, in such settings, consortium players can make shared investments to support the deployment of the data space. Conversely, some data spaces are driven by public value, which is often more difficult to relate to individual stakeholders. In such cases, it can be worthwhile to have a public party making initial investments for the data space. Note that once the data space scales, new participants could be asked to pay a (one-time) fee to compensate for the initial investments necessary to set-up the data space.

1. Introduction: Data spaces defined

Data sharing is crucial for both organizations and society as a whole in our digital age. For private organizations it facilitates collaboration, informed decision-making, and innovation by providing access to diverse insights and information. For society, data sharing promotes transparency, enables research advancements, and supports the development of solutions to complex challenges, ultimately fostering progress and collective well-being. However, organizations as well as individuals might often be reluctant to share their data for fear of losing data sovereignty (among other things). Traditionally, data sharing requires providing access to one's data to one, large, dominant player in the ecosystem. This typically means a loss of data sovereignty and privacy is paid as the price for receiving the benefits of digital transformation (benefits such as process automation, big data analytics).

Data spaces are a solution to avoid the loss of privacy and sovereignty: a data space allows for safe and secure sharing of data between parties within a data ecosystem¹, without losing data sovereignty (Otto, 2022). Data spaces are decentralized,

allowing for data sharing while retaining full sovereignty over the data, hence without the interference of dominant intermediaries. As such, data spaces have a key role to play in the European strategy for data, because they allow for data sharing, and innovative applications that require data sharing such as AI-driven applications, in accordance to European values (Jarke et al., 2019).

Building and deploying a data space requires work in five, interrelated, “lines” or disciplines. Two of these disciplines are of a more technical nature, relating to the infrastructure needed for data sharing as well as the semantics that allow for interoperability. The others, are about rules of engagement, agreements and contracts (legal and governance) as well as the value proposition of the use-cases supported by the data space (business). The whitepaper you are current reading describes success, and some fail, factors for data space deployment related to these last three disciplines.

Despite its benefits, data space deployment is still in an early stage, and lacks clear guidance on how

this can be supported or structured. For example, the data space *Smart Connected Supplier Network (SCSN)* is successfully and continuously used since its foundation (supporting data exchange between suppliers of high-tech manufacturers). Yet, SCSN followed a rather ad-hoc development approach and took several years of working with data space stakeholders to establish a viable data space.

In recent years, quite a lot of work and effort has gone into technical aspects of establishing data spaces and the technical definition and infrastructure of data spaces is relatively more mature than the definition of governance and business aspects. Probably, success and failure of data space deployment currently have at least as much to do with choices that were made in terms of business modelling and governance structures as with technical issues. Therefore, the following central question of this study is formulated:

What are the critical success and fail conditions for the deployment of data spaces in practice, and what recommendations can be based on these conditions to support its deployment?

An answer to this question could help stakeholders in different ways:

- 1 to understand and overcome implementation barriers;
- 2 to be better prepared for unforeseen risks;
- 3 to timely refrain from data space implementation if the conditions for doing so are not feasible or desirable.

Therefore, understanding the success and fail conditions would contribute towards the successful implementation of data spaces in practice.

¹ A data ecosystem represents is a platform-like infrastructure that enables interaction between different stakeholders to share data and supports the resolution of interoperability, security and privacy issues for shared data (Geisler et al., 2021)

2. Our recommendations in short

Through interviews with stakeholders that are or have been involved in a data space we identified a set of factors that play a key role in the development of a data space (see Appendix 1, for a discussion on the method we followed). Below, in chapter 3, we go deeper into how we derived these recommendations from our interview data. Here, we summarize our recommendations for ease of reference.

1 Developing trust (both informal and formal) is crucial for the deployment and continuation of a dataspace. Stakeholders that distrust each other form a risk to the deployment and continuation of a data space.

Trust can be established and maintained by considering the following actions and conditions:

- **Build informal trust** with data space stakeholders. Address the following conditions that influence the level of trust:
 - Look into power differences between stakeholders (e.g. due to differences in size, capabilities or dependencies in a supply chain).
 - In case such power dynamics exist, or may start to form a problem, agreements that protect the

interests of all stakeholders (big and small) should be made explicit early on.

- Competition between stakeholders or other conflicts of interest may be a cause of distrust. The rules of engagement between competing stakeholders should be clear to all stakeholders, that is: on what topics stakeholders collaborate and on what domains they compete.
- In case stakeholders have collaborated successfully before, the initiative can build on the trust established in these collaborations. But in case earlier collaborations were unsuccessful, considerable energy should be spent rebuilding trust in each other.
- Once operational, these informal measures should be amended with **formalized measures** for the data space. That is, while informal means to build and maintain trust remain important, formalized procedures become more important whilst the initiative grows:
 - Procedures, agreements and rules should be made explicit and should be formally agreed to by stakeholders.

- Assure representation of all stakeholders in the governance board.
- Install a neutral organization as governance board.
- Assure that the time, **effort and resources are available** for building trust in all phases of data space development. Trust takes time to develop, and requires a deep understanding of why (dis)trust exists and how this can be reduced, addressed, or supported.
- **Establish trust before deploying** the data space, even when from a technical perspective it could already be deployed.
- 2 **Organization of the data space: Create a clear (and concise) scope and process to deploy the dataspace, while keeping in mind how the dataspace can scale up in later phases.**

Several aspects are important to consider to successfully organize a data space. Considering the following conditions and actions help to reach this:

- **Develop a common vision** in the exploratory phase of the data space to ensure a proper understanding of the

purpose and the value of the data space. This helps to align different stakeholders.

- **Begin with the end(-user) in mind.** Besides considering the availability of data, think about **how the data is used to create value and for what end-users:** without end-user adoption, data space development will be futile and scaling impossible to achieve.
- **Keep things focused to support clarity on the purpose of the data space:**
 - Do not start too broad in terms of functionalities and applications, but focus on a set of core (cohesive) functionalities which support a clear use case for end-users that addresses their needs.
 - Assess the expected end-user group and whether this is sufficient in terms of critical mass to support data space operation.
- **Expand iteratively as the data space develops and scales over time:**
 - Additional functionalities can be added post pilot phase as the user group diversifies or increases in size, bearing in mind that the balance between suppliers and users of data should be maintained.
 - Continuously re-evaluate the barriers for new entrants, and adapt

accordingly to ensure that a critical mass of users can be retained.

3 Data space deployment, calls for an understanding of its investment structure and value creation, as well as the selection of appropriate business models.

The following factors should be considered or actions could be taken to account for this:

- **Identify and include capital-intensive stakeholders** (e.g. public bodies) that are willing and able to invest in the data space:
 - Obviously, also for these organizations value capture and creation should be balanced (i.e., data space deployment should be “worth their while”). However, especially for public bodies, but also in some cases for private organizations, value may be non-financial as well as financial.
- **Identify appropriate business models for data space operation** by identifying how value creation takes place and what stakeholders accordingly are willing to invest:
 - Understand the business or value case for data space participation for all stakeholders. Here, again, the motivating factors for investment may (also) be of a non-financial nature (e.g., in case data sharing creates significant public value).
- **Balance value creation and capture for all data space stakeholders,** by providing clarity on value and incentives:
 - This also includes understanding how data space participation may affect, and sometimes undermine, entrenched business models of participants. In such cases, participants may need to be convinced of the superiority of the new business model supported by the data space (vs., the existing model).

3. Findings through interviews

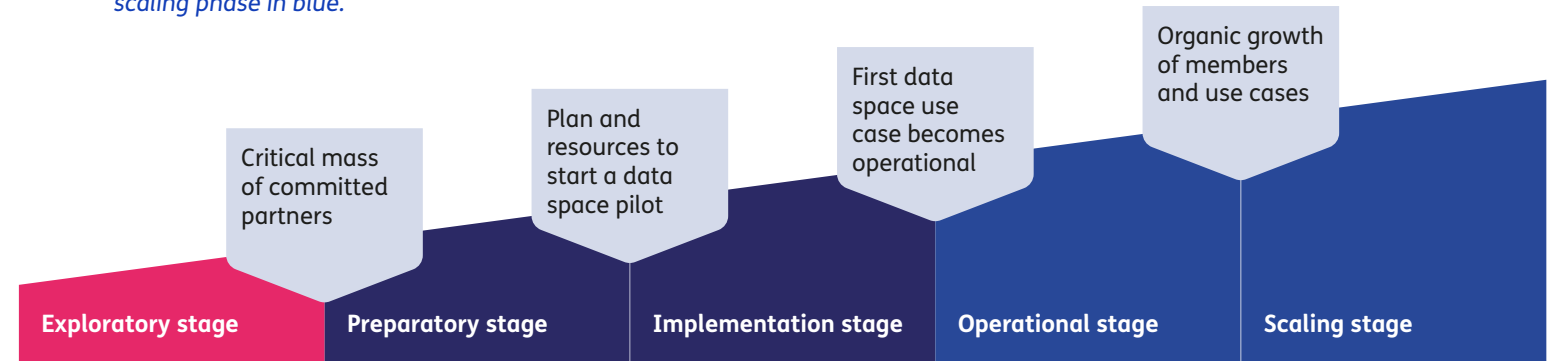
According to the Data Space Support Centre (DSSC), there are five phases to data space deployment (see figure 1). In this chapter, we map our recommendations on a simplified version of this model, which considers here phases: the *exploratory phase*, *piloting phase* (encompassing the DSSC's preparatory stage and implementation stage) and *scaling phase* (operational and scaling stage). We do this, to provide insight into when what recommendation should be implemented. In addition, we provide information on how we derived these recommendations from our interview data.

3.1. Trust

Effectively all interviews indicate that *trust* is key in supporting data space deployment, as it influences the degree to which stakeholders are willing to collaborate and exchange and share data as part of a data space infrastructure. Different perspectives on trust were mentioned during the interviews (see Table 1)². Without trust, the willingness of stakeholders to exchange and share data

² N.B. If there are no crosses in a particular cell, it only means it has not been discussed in the interview. This says nothing about whether the interviewee thinks this particular point is important or not.

Figure 1 Process followed to support the deployment of data spaces (source: DSSC.eu). Piloting phase in dark blue, scaling phase in blue.



as part of data spaces (or even participate at all) decreases. This reduces the potential value a data space can create. Additionally, this may also lead to limited adoption of the data space, resulting in financing problems to sustain its operation.

For the interviews, we see that trust can be established through informal or formal means – although in most cases: through both type of measures. Informal trust measures refer to non-formalized trust building activities, such as promoting interpersonal contact. Formal trust measures include formal agreements and rules of engagement. Hence, trust may be the product of how stakeholders view

Table 1 Overview of forms of trust mentioned in the different interviews.

Interviews	Developing trust is crucial			
	Build informal trust	Formalized measures	Resource availability	Establish trust before deployment
SCSN 1	X	X	X	X
SCSN 2		X		X
Talking Traffic		X	X	
DMI	X		X	
Join Data 1				
Join Data 2				X
D-Just Connect	X			X
DITM	X			X

each other, their past experiences, their positioning for the market or their previous relationships and interactions with other data space stakeholders (informal trust). Formal trust on the other hand is fostered through governance mechanisms which provide security or 'trust' for the collaboration. Both aspects are considered as important for data space deployment but should be considered at different phases (see Figure 2).

3.1.1. Building informal trust

Informal trust largely results from the dynamics between and composition of stakeholders involved for a data space. This calls for stakeholder engagement and management to ensure that these dynamics lead to (informal) trust for partners. We can see in Figure 2 that even though informal means to trust building do play a role throughout the data space deployment process, informal trust is particularly important to consider in early phases of data space deployment.

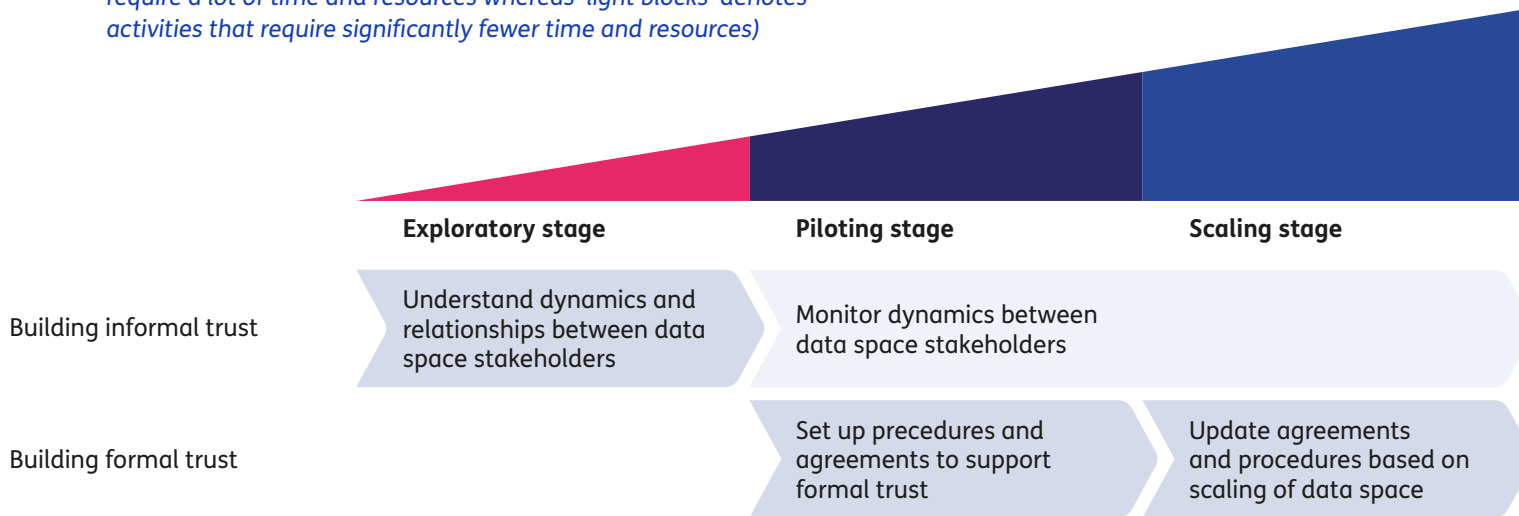
The composition and dynamics between stakeholders influences the degree to which formal mechanisms are needed (for example as stakeholders operate in similar competitive environments or in case clear power dynamics are at play). Interviews indicate that informal trust can be the result of the following factors which should be considered:

- (Dis)trust as a result of previous collaborations.

It is indicated that trust exists because of (past or ongoing) positive or successful collaborations between data space stakeholders, and that this can aid the deployment of data spaces in practice. Note that these collaborations can even take place outside of the context of the data space (as part of supply chain operations, alliance-like structures or partnerships). As previous collaborations proved positive (and therefore understand what value other partners may bring to the table), stakeholders may be more inclined to work together as part of data space deployment. Note that unsuccessful (or negative) collaborations can result in a lack of trust, or even distrust. In such cases, considerable time and effort should be spent addressing the lack of trust, or distrust, and to rebuild trust.

- Power dynamics between data space stakeholders

Figure 2 Trust in data space deployment ('dark blocks' denotes activities that require a lot of time and resources whereas 'light blocks' denotes activities that require significantly fewer time and resources)



differentials that cause smaller players to feel forced by a larger player to adopt practices or even business models that do not fit their objectives. As a result, the incentives of stakeholders within this supply chain may not be properly aligned, reducing trust for data space participation.

Power dynamics do not merely have to result from supply chain collaborations. It can also be the case that power dynamics exist due to differences in size and capabilities between data space stakeholders. For example, in case an organization has access to significant data sources to share as part of the data space, it may also feel the need to 'exert control' regarding whom can access and use this data to ensure that sufficient value can be captured by this organization. In some cases, such control may limit the ability of other players to capture value for themselves.

- *Competition between data space stakeholders, conflicting business models or incentives*

Interviewees indicate that trust is dependent on whether the incentives of stakeholders align or can be aligned. A data space may bring together stakeholders which operate in similar

markets or can even be considered as direct competitors (as they offer similar services or products). As a result, distrust may exist as stakeholders pursue conflicting incentives and objectives (they compete for the same market share or customers). If such stakeholders start sharing data as part of data space, they 'expose' part of their operations and may (inadvertently) put themselves at risk to weaken their competitive position. Depending on the likelihood and severity of this risk, it may make stakeholders reluctant to enter in a new data space or to start data exchange.

It should be noted that the presence of informal trust is not necessarily a prerequisite for data space success: data spaces exist and operate successfully which feature largely competing stakeholders (which do not have to trust each other outside of the boundaries of the data space) – this is where formal mechanisms come into play (see next section). Nonetheless, informal trust can help to strengthen the collaboration over time. Logically, as the composition of the data space changes over time (new players enter whereas incumbent players may leave), these informal trust may have to be continuously reshaped.

3.1.2. Building formal trust

As explained, formal trust is established through governance mechanisms, procedures and policies which serve as explicit means of driving the behavior of stakeholders (and support the establishment of trust between stakeholders). We see in Figure 2 that building formal trust takes a more prominent role as part of the piloting and scaling phase: in these phases, governance mechanisms can be realized and enacted. This requires setting up structures to support formal trust building (either through agreements, rules of engagement or other mechanisms) and enacting upon these structures. Such governance mechanisms take particular importance in the scaling phase which is characterized by potentially new applications offered as part of the data space or new potential entrants. Hence, in the operational phase governance mechanisms should be set up for scaling in terms of onboarding new participants, supporting new functionalities or even federation of the data space with other data spaces. Through our interviews, we identified the following mechanisms which help in building formal trust:

- *Use of formalized procedures, policies and rules on data sharing and participation*

Formal procedures, policies and rules on the type of data to be shared, how data sharing takes place, and who has access to data shared for the data space are ways to influence data space stakeholder's behavior and, therefore, trust in each other and the solution. Even if competitors may be sharing sensitive information as part of the data space, these procedures and policies should safeguard the competitiveness of the respective stakeholders involved for the exchange: It should be clear for all stakeholders in which areas they can collaborate and where they can compete. 'Breaking the rules' results in some form of penalty or action. This should reduce the barrier for stakeholders to start sharing data.

- *Representation of all stakeholders in data space decision making*

It is indicated that the constitution of a governance board in which all stakeholders are represented (and are able to influence decision making on the purpose and functioning of the data space) helps in onboarding stakeholders for data space deployment. If stakeholders can influence decision making (being allowed to 'vote' on strategic, tactical or operational concerns regarding the data space),

it helps in reducing perceived risks as well as fosters the building of trust between data space stakeholders.

- Presence of a neutral organization as a liaison for stakeholder collaboration

We conclude that the role of an intermediary or ‘neutral’ organization in shaping the collaboration or as the governance board can help to overcome a lack of trust in the solution or the consortium. This organization therefore takes the lead in building relationships

between data space stakeholders as well as can act as a liaison in accelerating decision making in case decisions need to be escalated to a neutral party, meaning that potential conflicts between data space stakeholders (as they may pursue conflicting business models) can be resolved in a ‘fair’ and ‘transparent’ way. Note that the neutrality of this organization should be clearly demonstrated. In many cases, we observed that the third party organization takes on a not-for-profit model, which also support trust of participants.

Table 2 Overview of organizational aspects mentioned in the different interviews.

Interviews	Alignment organization and objectives			
	Develop a common vision	Begin with the end(-user) in mind	Keep things focused	Expand iteratively
SCSN 1	X		X	X
SCSN 2	X		X	X
Talking Traffic	X			
DMI				
Join Data 1	X	X		
Join Data 2			X	
D-Just Connect	X			
DITM	X			

3.2. Organization of the data spaces

In addition to supporting the collaboration between data space stakeholders (for which formal and informal trust are key to consider), we have also identified conditions that should be in place for a data space to be deployed and sustained over time. These insights are based on several topics that were mentioned during the interviews (see Table 2). As illustrated in Figure 3, we see that *establishing a common vision, start with a clear scope, ensure critical mass and provide low entry barriers* are key to either 1) define and drive the strategy for data space deployment or 2) ensure that the onboarding of stakeholders takes place for value creation and capture. In the following, we detail each of these factors.

3.2.1. Establish a common vision

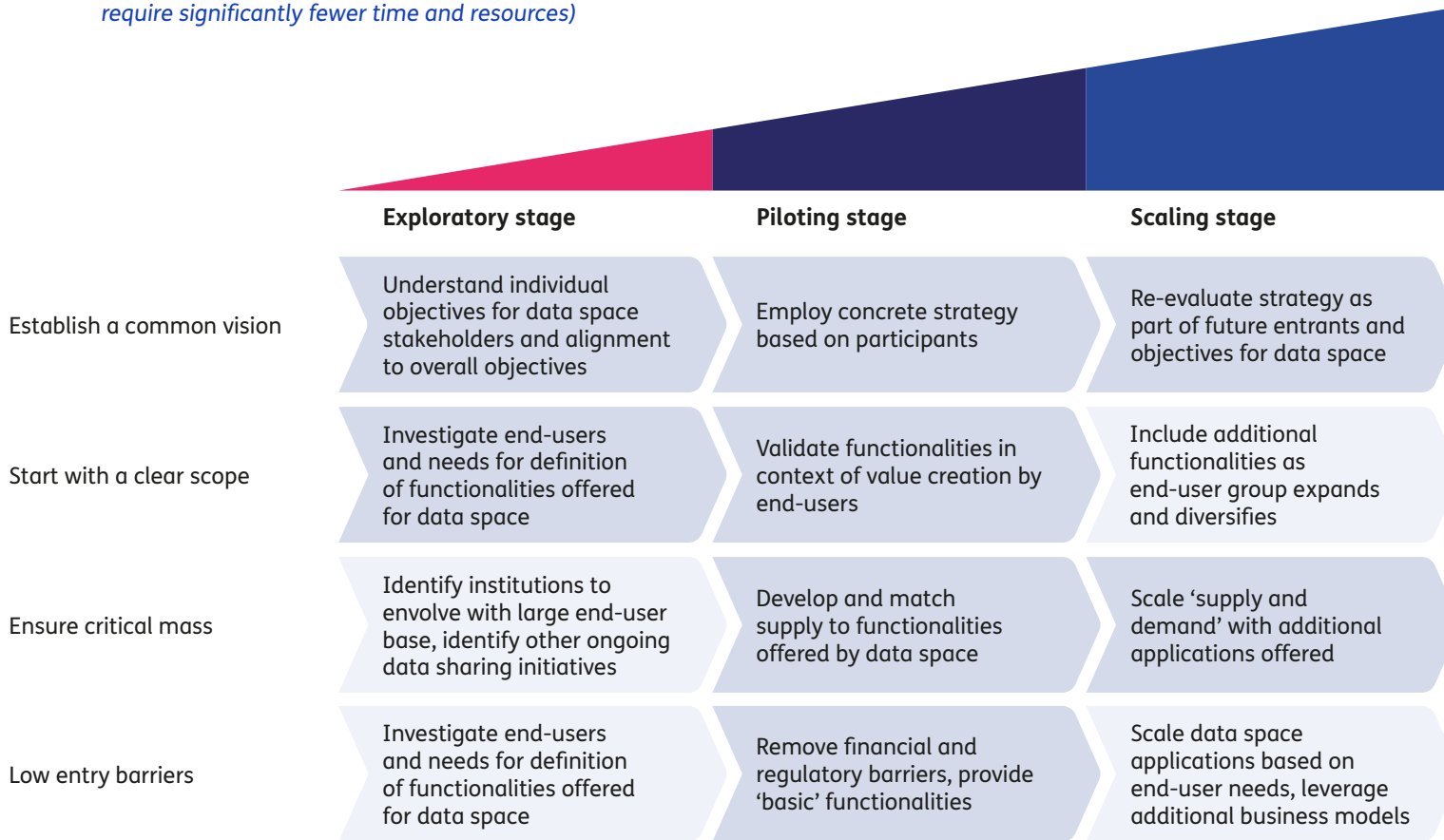
Interviewees emphasize developing a *shared vision* for the data space, as without a shared vision the objective for the data space (as well as its long-term goals) becomes ill-defined. It focuses on explicating why collaboration between stakeholders is needed and to engage stakeholders as part of building a shared story, identity or vision towards operating the data space (why do we need a data space, and what does this mean for us?). A concrete example of this can be observed

for SCSN (see appendix for a description of this data space): service providers initially considered the data space as ‘counter-productive’, as it meant shifting away from their current business model (i.e., selling individual, customized solutions to clients) towards a different (and therefore uncertain) business model (i.e., offering a standardized solution which can be scaled to all potential clients). Building a shared vision between service providers as well as highlighting the added value of the data space (i.e., to increase the scale at which service providers can operate) helped in driving the deployment of the data space in practice.

Building a common vision is not only important for the early phases (piloting) of the data space, but also on how stakeholders intend to support the data space long-term. A bottom-up strategy of defining a vision is advocated by several interviewees, which will help in stimulating its subsequent adoption. There should be a clear driver or need for organizations to consider the deployment of a data space, rather than a ‘mandated’ push to develop data spaces in practice (which often tends to be the case for European projects, which aim for the deployment of new technologies but lack clear coordination or monitoring mechanisms to sustain their use post-project).



Figure 3 Organizational factors for data space deployment ('dark blocks' denotes activities that require a lot of time and resources whereas 'light blocks' denotes activities that require significantly fewer time and resources)



3.2.2. Start with a clear scope and focus on value creation for end-users

In line with establishing a common vision, we also conclude from our interviews that it is important to start the development of a data space with a clear scope and focus, including coherent functionalities which support value creation for end-users. It is mentioned that data spaces are often pursued as a result of their technological novelty or the interest of combining and integrating different sources of data, without a clear understanding of why this is actually relevant for end-users to consider. As a result, functionalities and applications are included for the data space which do not match the needs of end-users, or are highly diverse and as a result dilute how value creation takes place for end-users. This is troublesome as it increases the entry barrier (3.2.4.) for end-users to participate for the data space. Therefore, interviewees suggest that it is better to start with a small scope (featuring a limited but coherent set of functionalities offered) which support a clearly defined value creation logic. From this initial set of functionalities, the data space can subsequently be scaled as the amount of users increases or diversifies.

3.2.3. Ensure that a critical mass exists

For some of the interviews, the importance of having a critical mass (sufficient early adopters as well as continuous growth in future participants) to support data space development is stressed. A data space in essence works as a *platform* or *multi-sided business model*: it requires the participation of data providers and data users in order to support value creation. Without a network of stakeholders to support the data space, it will be difficult to demonstrate its effects and hence to sustain, but also grow the data space over time (as adoption is dependent on the amount of stakeholders onboarded, so called *network effects*). Additionally, stakeholders will be reluctant to participate and commit resources as the intended effects of the data space cannot be shown.

Establishing this critical mass can be achieved through ensuring that entry barriers are low (as discussed in 3.2.4), but can also be realized through leveraging previous or on-going data sharing initiatives: in such cases, the data space can adopt existing infrastructure and onboarded members to work towards the critical mass needed. Alternatively, institutions such as cooperatives or trade associations (which typically represent a

large set of users) can serve as valuable partners in working towards a critical mass, as such institutions can quickly mobilize potential users and offer centralized support for the adoption of the data space.

3.2.4. Provide low entry barriers

We also conclude that in order to support data space deployment, it is necessary to ensure low entry barriers for stakeholders to participate. That is, end users should not have to invest greatly in things like upskilling or technology. For example, if the data space is geared towards farmers (such as the case for *JoinData* and *DJust-Connect*) which do not generally have the infrastructure needed (nor the knowledge of data spaces) to support data sharing applications, it may be troublesome for such end-users to join a data space. However, participation and onboarding of such stakeholders is essential to ensure that a critical mass can be achieved. Accordingly, the barrier towards entry for the data space should be low, meaning that policies and procedures for the data spaces should be flexible and understandable and the applications supported by the data space should be user-friendly.

More so, in the early phases of data space deployment, it can help to offer ‘basic’ functionalities in terms of data sharing: this focus should improve the clarity on what the data space intends to achieve and what value it can yield for end-users. As the data space is scaled over time, new functionalities can be introduced. Additionally, financial barriers should be addressed for such stakeholders to partake for the data space (further explained in Section 3.3): this calls for

a deep understanding of where value creation and capture takes place (and at what stakeholders).

3.3. Financialization and investments for the data space

Our interviews also address the financialization and investments needed to support data space deployment (see Table 3). Data space often pose significant investments but involve many different

Table 3 Overview of topics related to financialization and investment mentioned during the interviews.

Interviews	Financialization and investments		
	Identify and include capital-intensive stakeholders	Identify appropriate business models	Balance value creation and capture for all data space stakeholders
SCSN 1		X	X
SCSN 2		X	
Talking Traffic			X
DMI	X		X
Join Data 1	X	X	
Join Data 2	X		
D-Just Connect	X	X	
DITM	X	X	

stakeholders. As a result, identifying the stakeholders at which value creation and capture takes place can be difficult, which influences the degree to which stakeholders are willing to invest for a new data space. Through our analysis,

we identified three factors to consider for data space deployment: *the need for support in terms of initial investments for data space deployment, the selection of appropriate business models to support data space operation, and the challenge of*

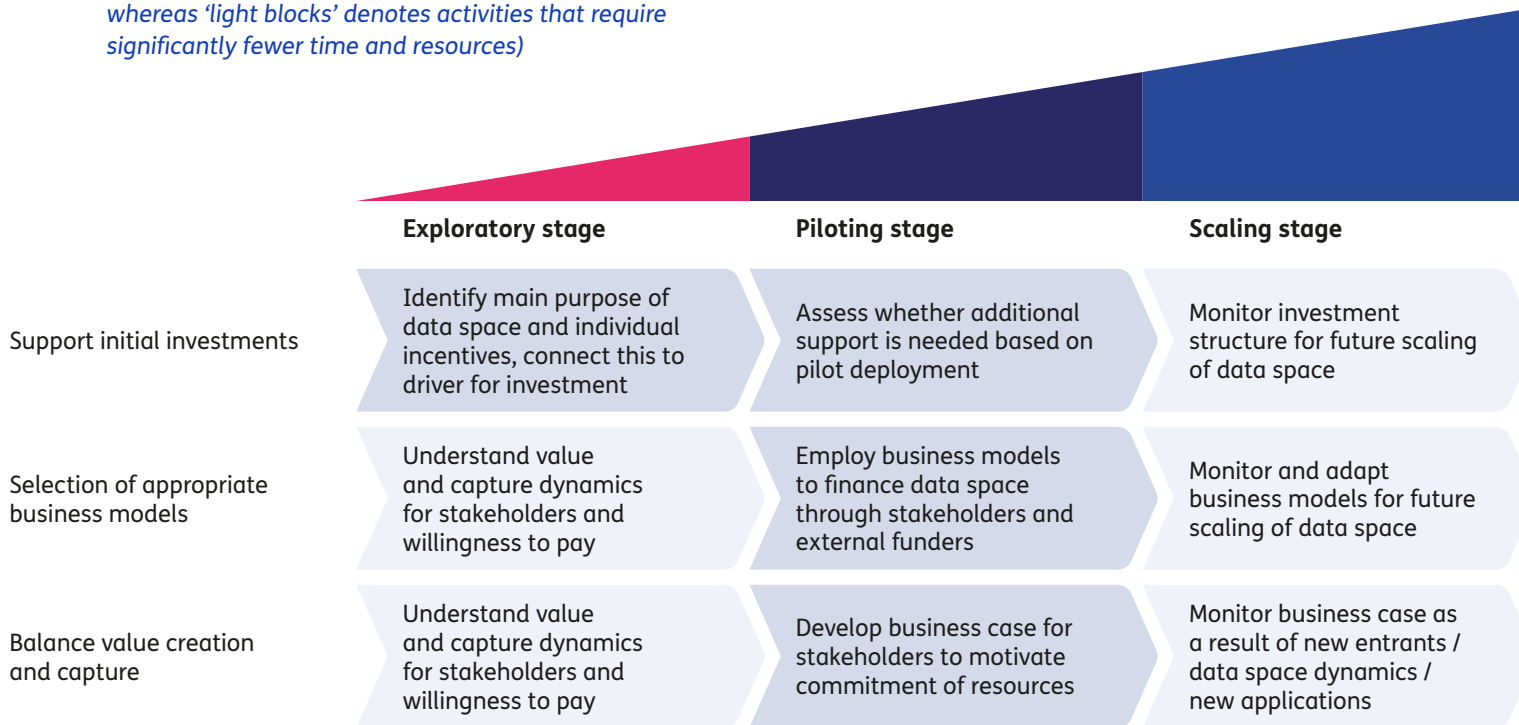
achieving a balance between value creation and capture for data space stakeholders. As illustrated in Figure, these can again be plotted to the typical process followed for data space deployment and describing how these can be taken into account.

3.3.1. Support initial investments for data space deployment

Through our interviews, we see that initial investments to set up a data space are usually between 1 and 1.5 million euros to kickstart a data space. These ‘instigation’ costs (i.e., to set up the data space and to support its minimal viable operation) are often difficult to appropriate or subdivide amongst initial adopters, as in its initial phases the data space has not yet proven its value to stakeholders involved. Here, a clear and strong business case is required for the stakeholders involved: what value are you able to capture at what cost? What does this imply for how investments can be subdivided (see also 3.3.2. and 3.3.3.) and what return on investment can be defined for stakeholders making these investments? This can help in shaping the investment structure for the data space, and also highlights the role that stakeholders play in realizing or operationalizing the data space.

Our interviewees indicated that there is much variation around the numbers mentioned in the previous paragraph: the investment required for a data space depends on the services the data space will have to support, the sector and kind of product market the stakeholders are operative in, as well as other choices.

Figure 4 Financial factors for data space deployment (*‘dark blocks’ denotes activities that require a lot of time and resources whereas ‘light blocks’ denotes activities that require significantly fewer time and resources*)



Hence, the level of investment required is very much context specific, and generic recommendations are hard to give. However, in appendix 3 we have created a table which sets apart different topics which influence the level of investment based on our desk research and interviews. For each topic several elements (situations or choices) have been described based from least costly to most costly. For example, a data space that needs to support several use cases will require higher levels of upfront investment than one that only supports one use case. Each of the initiatives we have spoken has a mixed profile, that is, for all of the cases studied, some choices were made that increased required investment, whereas others reduced it. We hope the table in appendix 3 provides some guidance to decision makers on how to structure and shape their investment choices.

Although these investments can largely be covered by the consortium of data space stakeholders, we see that alternative means of financing can be considered, particularly if the data space can create public or societal value (for example in sharing mobility data or sharing data to support sustainability). For example, in some of our interviews, governmental institutions took a role to cover initial

investments made for the data space was present. Here, the government institution did not necessarily take an active role in driving or even leading the collaboration between data space stakeholders (as trust should ultimately be intrinsically established by stakeholders involved), but rather served to kickstart the data space to overcome initial finance barriers faced in early phases of data space deployment. Through involvement of an external funding party (government institution), financial barriers for data space stakeholders to participate can be reduced or removed.

3.3.2. Selection of appropriate business model to support data space operation

In regards to financial requirements, we conclude from our interviews that the selection of appropriate subscription or business models is important to sustain operations as part of a data space. Sustained operation of a data space can range from 200.000 to 800.000 euros on a yearly basis, depending on the amount of users and the amount and type of data that is exchanged. These yearly expenses should be compensated to sustain data space operation, but it raises the question what data space

stakeholders should ultimately be billed (and why this is the case). Our interviews indicate that billing of the data space is often placed at the service providers: such service providers pay a fixed fee to receive access to the data space as well as variable fees depending on the amount of data exchange and interactions that take place. Data providers and data consumers (such as farmers in the case of *JoinData* or mobility providers in the case of *Talking Traffic*) are not charged, as the primary benefit for the data space lies at the service providers in creating a standardized solution for these providers to support data sharing. The service providers subsequently set up contracts with the data providers and data consumers for their services in supporting data sharing through the data space.

3.3.3. Balancing value creation and capture for data space stakeholders

Many of the interviews also stress that the expectations of stakeholders in terms of 'contributing and capturing' value should be managed, as it creates the incentive for stakeholders to participate. Here, it should be evident what the business case is for data space stakeholders to participate, which logically may differ per data space

stakeholder and as well may depend on the individual objectives each stakeholder has. For example, in the case of *DITM*, public (road) authorities indicate the need to support mobility data sharing, as it can help to reduce traffic accidents and contribute traffic road safety and efficiency. Accordingly, there is significant value capture to be obtained for these authorities to start investing in a data space, and therefore can take a more active role in terms of making investments for the deployment and operation of the data space. However, this data sharing may also have benefits for (private) users, although these benefits may only become apparent once the data space is scaled or once additional functionalities are introduced. Hence, understanding this value and aligning individual motivations as part of concretizing the investment structure can help in supporting the subsequent deployment of the data space, which can be refined as the data space is scaled over time. Without clarity on value and incentives, it will be difficult to motivate stakeholders to commit resources to do so.

4. Conclusion

In sum, whatever data space you want to set up and deploy, in whatever sector, you would do well to carefully consider at the least three questions:

- 1 How are we going to build and maintain trust among stakeholders, even when the group of participants starts growing?
- 2 How will we organize our data space in order to achieve our vision and objectives?
- 3 Who will pay for initial deployment and later participation, and how do we make it worth their while?

In the paper above, we have identified several generic recommendations that answer these questions. But we recognize, of course, that the way our recommendations work out and get implemented depend highly on the specifics of the case at hand. That said, we hope to have given anyone interested to become involved with a data space some guidance as to how to maximize their chances to success, and minimize the chances to failure.

5. Acknowledgments

This research has received financial support through the *Centre of Excellence for Data Sharing and Cloud* (CoE – DSC) in the Netherlands. In addition, we would like to thank the interviewees and the respective organizations – Dat.Mobility (a Goudapple company), ILVO, Monotch, SupplyDrive, and TNO – for our selected case studies for their willingness and openness to participate in this study.

6. Referenties

Braun, V., & Clarke, V. (2012). *Thematic analysis*.

Geisler, S., Vidal, M., Cappiello, C., Loscio, B., Gal, A., Jarke, M., Lenzerini, M., Missier, P., Otto, B., Paja, E., Pernici, B., & Rehof, J. (2021). Knowledge-Driven Data Ecosystems Toward Data Transparency. *Journal of Data and Information Quality*, 14(1).

Jarke, M., Otto, B., & Ram, S. (2019). Data Sovereignty and Data Space Ecosystems. *Business & Information Systems Engineering*, 61, 549–550.

Otto, B. (2022). The Evolution of Data Spaces. In *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage*.

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Publications.

Appendix 1 – Methods

We investigated data sharing initiatives in three different sectors: Mobility, agriculture and manufacturing. Since several different data spaces have become operational in these three sectors across Europe, our selection helped us to ensure that our findings would be applicable across sectors. In this chapter, we shortly describe how we collected and analyzed the data for our research question.

For each of the sectors, we identified and approached interviewees with deep and extensive knowledge of data space deployment in the sector. Specifically, we invited interviewees who either:

- 1 Have been or are currently responsible for the coordination of the deployment of a data space.

Or

- 2 Have been or are currently heavily involved in the deployment process.

Or

- 3 Are currently active in supporting the continuation of the data space (post pilot) (and thus can reflect on success conditions).

Using these selection criteria allowed us to safeguard the quality and relevance of insights generated. Each interview was conducted by two researchers from TNO, each with a background in business and innovation management. We used an interview protocol was devised (see Appendix) to structure the interviews and guide the elicitation of results. Each interview was recorded to support analysis of the results post-interview.

We used a thematic analysis (Braun & Clarke, 2012) identifying ‘themes’ regarding data space deployment which can either be linked to *business* or *governance aspects* of data space deployment which may influence their success (or failure). We first closely examined our interview results to identify initial categories and themes called codes, without a pre-defined list. This process is called *open coding*. Next, we organized our initial codes into broader themes and categories through a process that is known as *axial* (sense making) coding (Strauss & Corbin, 1990) to structure this analysis.

The result is a set of drivers, barriers and challenges towards data space deployment which can be considered as success or failure conditions for data spaces (i.e., drivers should be in place to motivate

and incentive data space deployment, whereas barriers and challenges should be mitigated). These drivers, barriers and challenges are discussed above in terms of success and failure factors.

Case study selection

We intentionally focused on initiatives from different sectors, and with different levels of maturity. This allowed us to capture insights that are applicable across sectors as well as levels of maturity. Not all of the initiatives studied are actually a data space. More in particular, while JoinData – an agricultural datasharing initiative we studied – has certain features of a data space (in particular safeguards for data sovereignty), it uses a more centralized solution. We did include JoinData to help us understand why a collaboration would not choose to organize datasharing according to data space principles. Each of the initiatives studied has mainly Dutch participants, apart from D-JustConnect, which is a Belgian initiative.

The list of case studies included for our analysis is presented in Table 1 – per case study, we offer a brief description on its purpose and objectives.

Smart Connected Supplier Network (SCSN) is a data standard that makes information sharing in the supply chain more efficient, allowing companies to share data more easily, quickly and reliably.

Talking Traffic focuses on facilitating the exchange of data between road users and intelligent infrastructure within a public-private data chain in order to reduce the adverse effects of mobility on the basis of real-time data.

Dutch Metropolitan Innovations (DMI) is an initiative of a consortium of companies, knowledge institutions, (pioneering) municipalities, provinces and several Dutch Ministries that wants to give concrete, scalable and data-driven substance to smart, sustainable urbanization and mobility renewal and remove (technological) barriers on the side of supply and demand.

Join Data is an independent data platform aimed at improving processes and data sharing in the Dutch agricultural sector. It is a cooperation owned cooperatively by both farmers as well as other cooperatives in the sector such as FrieslandCampina and Rabobank.

D-Just Connect provides a digital infrastructure for safe data transport, primarily in the Flemish agricultural sector. DJustConnect is a neutral data exchange platform, hosted by ILVO, available for all data users in the Agrifood sector.

Digital Infrastructure for Future-proof Mobility (DITM) is a National Growth Fund project subsidized by the Dutch government. Within this project TNO is working with partner companies to implement the digital infrastructure for automated transport, with the aim of making mobility safer and more efficient.

Table 4 Descriptives for selected case studies and associated interviewees

Data space case study	Domain	Interviewees	Role for data space deployment	Date of interview
SCSN	Manufacturing	1 Community manager	Both interviewees involved in exploratory phase and responsible for current business	1 15-11-2023
		2 Owner data intermediary		2 30-11-2023
Talking Traffic	Mobility	Partner	Responsible for current deployment	29-11-2023
DMI	Mobility	Software architect	Involved in exploratory phase	27-11-2023
Join Data	Agriculture	1 Lead architect	Both interviewees involved in exploratory and early operational phase	1 27-11-2023
		2 Business modelling expert		2 5-12-2023
D-Just Connect	Agriculture	Senior engineer	Involved in exploratory phase and responsible for current business	30-11-2023
DITM	Mobility	Senior project manager	Involved in exploratory phase and responsible for current business	13-12-2023

Appendix 2 – Interview protocol

Introduction

Getting know each other and brief introduction to the research at hand.

Context and Background

- Ask the interviewee to provide a brief overview of their role and responsibilities within their organization, and the data space.
- How is the data space used within their specific organization. What is the impact on decision-making and operations?
- What is the role of the data space in the context of the project?
- What kind of data space is this?
 - Private and / or public parties?
 - Cooperative entity?
 - Data sharing pool: Organizations pool data to generate additional value.
 - Data cooperative: Data subjects pool data at cooperative, which manages data in their interest.
 - Public data trust: Public bodies collect data to create public value.
 - Personal data sovereignty: Subjects pool data for collective understanding, but maintain control.
 - Unilateral data sharing: Organizations collect data at

subjects to provide services, generate control over data.

- How mature is the data space currently?

Process

- What is the history of the project (how it started – how is it going).
 - Which phases did you see yourself in the project?
 - Pre-phase
 - Did you already know the other partners upfront? How mature was your collaboration before you got started?
 - Preparatory stage
 - Who took the initiative? And why them?
 - How was trust established between partners?
 - Did you understand the needs and interests of other users going in? If not, how did you come to understand them and when? And what were your own incentives/interests? Where they already clear from the beginning?
 - Implementation stage
 - How was the project scoped? Did it always keep that scope?
 - What agreements did you make about (data) governance early

on? Did you need to update these agreements later?

- Operational stage
 - Did you consider other data sharing solutions? What were your considerations for the solution chosen?
- Scaling stage
 - How was the data space scaled (if at all)?
 - ◆ Did new partners join and if so, what kind of partners?
 - ◆ How were contacts established?
 - ◆ Did the collaboration change due to new membership?
 - ◆ Was new functionality added to the data space? If so, which and why?

Challenges and roadblocks

- What were the main challenges and roadblocks?
 - Organizational
 - Regulations and contractual arrangements
 - Trust and stakeholder engagement
 - Did you experience a breach of trust? If so, how was it addressed?
 - Human resources (skills & capacity)
 - Data governance and security.

- Discuss data governance practices related to data spaces, including data access, data sharing, and data security.
- Explore any compliance or regulatory issues they have encountered.
- How is data sovereignty secured?
- Did you use materials developed by DSSC or IDSA in the development process?

Business value & finances

- What is the major benefit for your organization to be part of this data space?
- What did it cost to start this data space in total?
- How were these costs distributed between partners?
- How did you come to that decision?
- What is the “business case pattern” for the data space (cost sharing, joint innovation. combined forces, shared marketplace, greater common good)?

Future trends

- What is the future of the data space?

What challenges do you foresee for the future?

Appendix 3 – Splitting Investment

Table 5 Elements that influence the level of investment required to set-up a data space

Topics	Low Investment	Medium Investment	High Investment
Use cases	<ul style="list-style-type: none"> • There is one use case running within the data space • Only the data space orchestrator can initiate the creation and implementation of new use cases 	<ul style="list-style-type: none"> • Multiple use cases are operational on the existing platform • Members of the data can create and implement new use cases 	<ul style="list-style-type: none"> • The data space is fully integrated and there is no limit on use cases which could become operational • Every party could create and implement a new use case
Alteration to business-activities end-user	<ul style="list-style-type: none"> • Automation of existing business activities, only changes in task execution by end-user 	<ul style="list-style-type: none"> • The business-activities have altered by the input of the data space, tasks of employees have altered however job descriptions have largely remained the same 	<ul style="list-style-type: none"> • A new way of working is introduced by the use cases in the data space, job descriptions have changed and business-activities have been created or altered
Data quality and availability	<ul style="list-style-type: none"> • The data exchanged only needs minor alterations in its master data • The data only needs to be interoperable between parties on a bilateral basis 	<ul style="list-style-type: none"> • There are significant alterations to the master data each participant needs to do to be able to exchange the data • The data needs to be interoperable within significant portions of the data space 	<ul style="list-style-type: none"> • All master data by all the participants needs to be altered to exchange it • The data needs to be fully interoperable within the entire data space
Governance Framework	<ul style="list-style-type: none"> • Reusable standard contracts are designed and signed by new members 	<ul style="list-style-type: none"> • Several different contracts are created for the participants to join the data space 	<ul style="list-style-type: none"> • Each contract with a new participant is renegotiated
Stakeholder management	<ul style="list-style-type: none"> • Stakeholders to be involved are clear • The stakeholders have worked or are working together • There is a level of trust present 	<ul style="list-style-type: none"> • A group of stakeholders is known however not sufficient to set-up a data space • Some stakeholders have worked together before • There is no trust and no distrust 	<ul style="list-style-type: none"> • The group of stakeholders still needs to be found • None of the stakeholders have worked together before • There is distrust between the stakeholders
Financial organization of data space orchestrator	<ul style="list-style-type: none"> • Third party financing pays to keep the data space operational 	<ul style="list-style-type: none"> • Participants make in kind or financial contributions to keep the data space operational 	<ul style="list-style-type: none"> • The data space needs to be fully self-sufficient in terms of financing
Familiarity to the technical solution	<ul style="list-style-type: none"> • The concept data space is known and in defining the use case the goals is to create a data space • Existing infrastructure or solutions can be used to create the data space 	<ul style="list-style-type: none"> • The concept data space is known to some stakeholders, multiple technical solutions are considered • Customize existing systems or implement new but standard technologies 	<ul style="list-style-type: none"> • The concept data spaces is unknown, in developing the use case the concept is introduced • Developing proprietary infrastructure or specialized systems from scratch

Contact details

If you would like more information on this research, or are interested in exploring how the results obtained through this research can be used in practice, feel free to contact the following TNO researchers.

Geert Lamerichs

✉ geert.lamerichs@tno.nl

Gijs van Houwelingen

✉ gijs.vanhouwelingen@tno.nl

Jisca van Bommel

✉ jisca.vanbommel@tno.nl

Rick Gilsing

✉ rick.gilsing@tno.nl

Frank Berkers

✉ frank.berkers@tno.nl

TNOvector
Centre for Societal Innovation and Strategy

CoE
DSC

tnovector.nl **coe-dsc.nl**