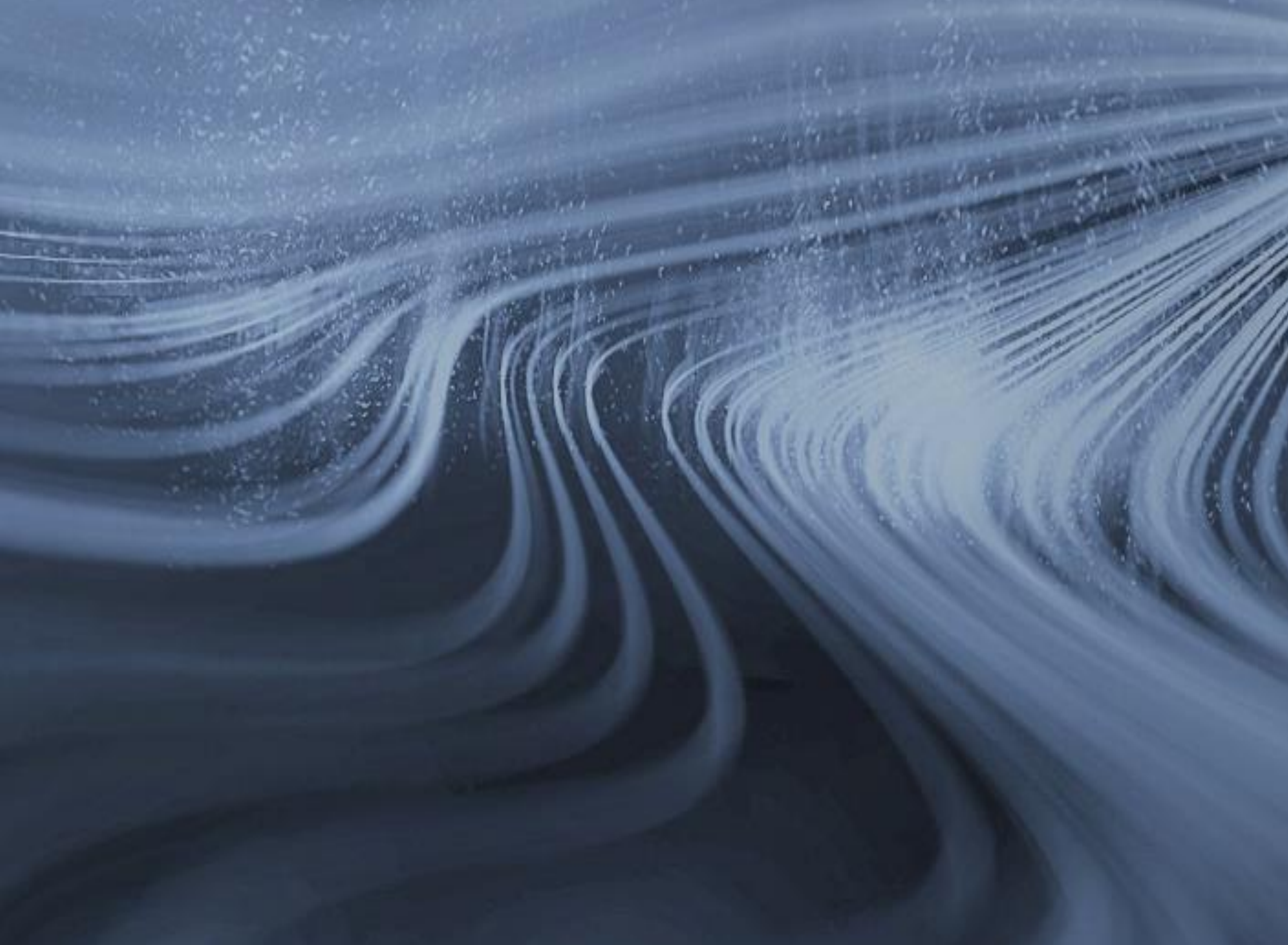
NL AI Coalitie

Responsible data sharing in AI

Original document in Dutch
NL AIC March 2020

Translated document
NL AIC July 2020



Preface

The Netherlands AI Coalition (NLAIC) is advancing in leaps and bounds. At the time of the writing of this preface (March 2020), we already have three hundred organisations as part of our coalition. All in all, a great result, but this is only the beginning. As a proper coalition, we have two main goals: cooperation and results.

This report is a concrete result of true collaboration within the "Data Sharing Working Group". It only happened because of the cooperation amongst participants from different organisations in the working group. Creating this report was not trivial. Only by writing things down, you formulate specific objectives and make choices which are collectively supported. In that way, we establish a connection among participants.

This report is part of our NLAIC proposition, where we realize responsible data sharing as the basis for AI applications working, both with and for our members. Data sharing is a topic of great importance; data is, combined with smart algorithms, the main building block for AI applications, which are also expected to perform adequately and responsibly.

In addition to this report, we offer training courses and a manual on how to get to AI implementation (proof of concept). Henceforth, an underlying knowledge pool is available. This year, we start applying this knowledge in practical use cases, which are proposed by the working groups of the NLAIC in different areas of application. Hereby, we avoid small-scale solutions but strive for end-to-end and scalable solutions, which can be widely employed.

This report was collaboratively produced and supported by the Ministry of Economic Affairs and Climate, TNO, and by the active participants of the Data Sharing Working Group. The working group consists of companies (large and small), universities, governments and interested parties, that read along, co-write, and above all participate!

Kees van der Klauw
Coalition Manager Netherlands AI Coalition
Harrie Bastiaansen
TNO
Frans van Ette
Chair of the Data Sharing Working Group

1. Introduction

1.1. Background

The participants of the Netherlands AI Coalition (NLAIC [1]) have indicated that data sharing is an essential topic for improving the position of The Netherlands in the AI domain. That is a logical conclusion: AI applications need data to allow their algorithms to train, improve and be implemented. Hence, it goes without saying that access to data is crucial.

Ideally, data is freely accessible, but the reality is often different. Organisations sourcing the data make independent choices about who may use their data and for what purposes. Data has inherent value, and sometimes may not or only be partially shared. Also, there may be costs involved in making data available. Moreover, there are regulatory restrictions on data sharing, such as the General Data Protection Regulation (GDPR).

Concerns about trust, safety and lack of control over the use of available data currently hamper large-scale data sharing [2]. All of these slow down the development and introduction of new AI applications, despite the significant advantages that may be achieved when larger volumes and different types of data are ubiquitously available.

Therefore, it is not straightforward to develop and to use AI applications with data that is coming from a multitude of sources. The Dutch Ministry of Economic Affairs and Climate recently published various policy guidelines [3],[4] to share data within and amongst economic sectors and society. These guidelines outlined the economic value of sharing data. The importance of a suitable environment for sharing data was highlighted as a crucial factor. Albeit these policies are not specific to the application of AI, they are, nevertheless, very relevant for AI. Also in the EU policies, data sharing and AI are on the spotlight. Recently, the European Commission released a communication about the importance of AI for Europe [5], a coordinated plan [6] as well as a white paper on AI for Europe [7] together with a communication on the European data strategy [8].

The aim of the workgroup “Data Sharing” within the NLAIC is to reduce the hurdles on data sharing in AI as much as possible. An important starting point is that the owners of data must hold control of their data. The most important term is ‘data sovereignty’. Meaning, it is not about the confidence that “it will be all right”, but instead, an adequate environment is needed that guarantees that data is handled responsibly - under conditions determined by the data owner. There is plenty to consider; however, a lot is already possible. Sometimes difficult choices have to be made, but given the possibilities that AI offers, it is all well worth the effort.

1.2. The strategic importance of data sharing in AI for companies

Since the beginning of the digital age, the role of data technologies within organisations shifted from functional to strategic. Previously, organisations had an IT department that had services delivered to the operational departments. Now, data (and what the organisation can do with it) is often determining the overall organisation strategy.

Organisations mainly create value from data when data comes from the most relevant sources. That means not only their own data but also data coming from outside of the organisation. Typically, external data comes from known suppliers and customers, but may also come from other parties with whom the company has not (previously) cooperated. Sometimes these parties are even competitors. Hence, it is clear for all parties that data is a core asset and organisations care a lot about protecting their valuable data while, at the same time, they try to obtain data from other parties.

The strategic relevance of data sharing has grown throughout a learning path started by digital forerunners. Many tech giants have grown precisely by their ability to collect the most insightful data. They too have started small: companies such as Amazon, first capitalised in existing revenue models, like selling books. Then, analytics showed where inefficiencies occurred in the business processes and how it can be done better. In the meantime, more or less every industry sector is refining its processes.

These forerunners have recognised that comprehensive forms of data analysis based on (shared) data enabled new revenue models. These more comprehensive forms of analyses formed the basis for, then, innovative revenue models, whereby services are sometimes provided free of charge given they yield or supply relevant data. Hence, data has significant value, and we must control access to the right sources.

That said, we arrive at sharing data amongst organisations: vast data sources improve the accuracy of AI systems forecasts and improve their strategic analysis. All of that supports decision-making and innovation to adapt and meet the changing requirements of the end customer. Data shared between organisations, linked to the use of AI, enables organisations to create a customer-centric, integrated business model that provides greater efficiency, flexibility and increased visibility. It may be clear, the business model is defined here in the broadest context, including public services and non-profit organisations. It is, therefore, about all aspects of our economy and the way we organise our society.

There is a rapidly growing number of examples of the use of data sharing in AI for all kinds of applications. The increasing importance of data sharing can be seen in how AI continues to develop step-by-step, enabling new revenue models and services [9].

In this document, we do not further question what we want to do with AI. AI applications are conceived for a reason, which we can assume to be a given. The starting point for this report is that data sharing is a pre-condition for optimally opening up data sources for the benefit of the AI applications.

1.3. Data sharing for the NLAIC: role and context

NLAIC has identified five cross-sectoral themes to stimulate the development and large-scale adoption of AI in the Netherlands. These themes are defined as "across sectors" and are referred to as "horizontals" or "building blocks". The theme of data sharing is one of these building blocks [1].

The working group "data sharing", within the NLAIC has been set up to tackle the challenges in data-sharing ecosystems. This working group plays an essential role in shaping a positive climate and creating a governance model to facilitate data sharing in AI across organisational boundaries, within the Netherlands and in line with international developments.

The context here is AI, and that has its specific challenges. However, data sharing is a vital topic in many different areas and applications. Think of exchanging data for process optimisation across business chains or for tracking products in a supply or production chain. Therefore, various (inter)national initiatives are currently already active in the field of data sharing, which the NLAIC working group "data sharing" should participate with.

An important initiative in this context is the Dutch Data Sharing Coalition (DSC, [10]). The basic principle of DSC is that we must ultimately move towards cross-sectoral data sharing. So not only arranging it between a few organisations for a specific application but arranging it so that data sharing is generic. That is an outcome that also benefits NLAIC members. The costs of construction, implementation, and management decrease as soon as the solutions built in the NLAIC context are re-usable. Thus, the cooperation between NLAIC and DSC is of great importance for both initiatives and the participants involved.

A second crucial powerhouse of data sharing is the data register of the Dutch government [11]. Organised by the Ministry of the Internal Affairs, this portal is the basis to access government data. Much information is already available on the portal. There is a team available that helps to understand the possibilities and what is accessible, Moreover, thinking along to determine the best solutions for the requests. Ultimately, it leads to agreements between the parties involved the sharing of data. This approach is a perfect entry for NLAIC members into the availability of data within the government domain.

These are two prominent examples, but much more is happening in the Netherlands, both at regional and larger levels. Besides, this the NLAIC data sharing initiative is an excellent reason to seek connections and build something together. In this way, the various initiatives learn from each other, and together they can work in cohesion and coordination.

In addition to Dutch initiatives, a lot is also happening internationally. For example, Germany launched the GAIA-X initiative at the end of 2019 [12], which provides a European response to the major US platforms and how Asian (often Chinese) IT systems are developed and operated. Access to data plays a crucial role here, motivating why many initiatives are being developed in an EU context, and investments are being made in solutions where privacy-by-design and data sovereignty are used as fundamental principles.

1.4. Controlled data sharing for the benefit of AI

Organisations increasingly see that data represents an economic asset or a societal interest. Sharing data with other parties offers opportunities for innovation, new forms of collaboration, changing business models and new ways of organising society. Data sharing has thus become a significant (perhaps somewhat unseen) joint ambition of great importance for the earning power of the Netherlands, for social developments and for how people and organisations keep control or, in fact, regain control over their data.

There is an increasing need to understand the conditions for using data if shared with other parties. We want to share data, whilst at the same time abuse should be prevented, giving organisations and people the confidence to share their sensitive data. Therefore, the starting point is "controlled data sharing". Controlled data sharing has three aspects:

- *Data Sovereignty*, i.e. the ability of a natural person or organisation to be completely independent with regards to its data, i.e. the ability for a person or organisation to make exclusive decisions about the use of its data as an economic asset. It requires people and organisations to have control over the circumstances of how their data is shared and how other parties may process it.
- *Trust*, i.e the property of gaining assurance that the entities (persons, organisations or systems) with which data is shared are the entities they claim to be and that they act accordingly. Trust can be achieved through adequate identification, authentication and certification methods.
- *Security* i.e the ability to guarantee that the actual sharing of data is protected against unauthorised use, whether due to malicious or accidental intent. It includes aspects such as encrypted data transport and storage as well as software certification and attestation.

Controlled data sharing is a multidisciplinary process that is not only technology. Business considerations, legal considerations and agreements also play a major role [13]. Controlled data sharing within AI has similarities with data sharing in other domains, such as transaction or operational data sharing - used to improve the effectiveness and efficiency of supply chain processes. However, many features make responsible data sharing for AI also different, more critical or more complex.

1.5. This report: target group, goals and structure

This report was prepared by the working group “data sharing” for all participants in the NLAIC who want to share data for their AI applications.

The goals of this report are:

1. To indicate what makes data sharing for AI applications specific in comparison with data sharing for other applications and what challenges this poses. This goal is addressed in Chapter 2.
2. To present the approach for data sharing for AI, building upon visions, frameworks and technologies that are already being developed in various contexts. This goal is addressed in Chapter 3.
3. To describe the follow-up development process from the current “IST” situation to the future “SOLL” situation, as a process from first-time engineering to operationalisation. This goal is addressed in Chapter 4.

2. Data sharing benefiting from AI

Several characteristics of (controlled) data sharing for AI are different, more critical, or more complex than for data sharing for other applications. Other applications may, for example, be sharing transactional and operational data for process optimisation across the business chains or tracking products in a supply or production chain. In this chapter we identify these specific characteristics of data sharing for AI to take into account in the development and implementation for an adequate data sharing environment.

The following specific characteristics of data sharing in AI are discussed in the subsequent sections of this chapter: the types of data for AI systems and the properties of data sharing in AI.

2.1. Types of data for AI systems

There are different types of AI systems, each with their own needs concerning data sharing. Two main types are highlighted in this section: data-driven and knowledge-driven AI systems. Their common component is that intelligence or knowledge is brought into the AI system so it can perform a complicated task. The difference lies in the way in which intelligence or knowledge is obtained.

Data-driven AI systems automatically learn, for example, about the next steps to be taken through statistical and machine learning methods. Learning often happens using historical data, which as *training data* contains the relationship between properties - features - of data subjects and the resulting "outcome". In a simple case, a machine learning method finds the relationship between a limited number of types of data, for example, based on a simple linear relation. Advanced machine learning methods involve much more data and allow for much more complex relationships. The result is a model which summarises information from the training data in a limited number of parameters (model data).

In knowledge-driven AI systems, the knowledge is obtained by consulting domain experts. Gaining expert knowledge is called knowledge elicitation, for which there are many available methods, such as interviewing, case study and role play. Knowledge can be obtained and expressed in different ways: from "if-then" rules to mathematical functions. To make knowledge-driven AI systems possible, shared data must be expressed in a way that is in line with the acquired expert knowledge. When sharing data, requirements regarding the meaning of data must, therefore, be well described and recorded in taxonomies and ontologies, i.e. pre-agreed dictionaries of concepts and the possible relationships between these concepts.

Different types of data play a role in the execution of the data-driven and knowledge-driven AI systems, both in the development and operational phases:

- *Training data*. Data-driven AI systems use historical data to correct themselves. This process is called "training" or "learning", and training data is used for that.
- *Model data.* Training or setting up an AI system based on training data results in a model with several learned parameters. These together are the model data.
- *Knowledge data* In the case of a knowledge-driven AI system, the collected knowledge can also be interpreted as data.
- *Production data*. Production data is the data that is currently being processed by the AI system, and on which an analysis is performed for a decision or prediction. Both knowledge and data-driven systems process production data.

2.2. Data sharing properties for AI

As mentioned earlier, sharing data for AI applications has several specific aspects, which are described in the following paragraphs of this section.

2.2.1. Processing domain: collaboration models

Due to their ability to derive complex relationships, AI systems are ideally suited to analyse simultaneously many different data sources. Some AI systems also work better with vast amounts of data, for example, in deep learning.

The different data sources for AI systems cannot always simply be brought together. Either because the amounts of data are too large, or other reasons like confidentiality. Think of privacy restrictions due to GDPR or company confidentiality, for example. These reasons imply that data should remain with its owner or administrator and are therefore not to be transferred to other organisations.

In short, for various reasons, it may be interesting for AI systems to be distributed: the data is owned by different organisations and is in different places, and AI systems must be able to train using data from multiple sources.

The challenge here is that virtually all existing data-driven AI algorithms currently require data to be available in a single database or dataset. However, techniques are becoming available to design distributed AI systems, in which all the data to be processed doesn't have to be gathered in a single database at the same location. Federated learning is a type of algorithm able to learn from distributed databases. Furthermore, (secure) Multi-Party Computation offers possibilities to have distributed AI systems..

For the domain of processing in the training phase, three basic options (collaboration models) can be distinguished:

- *Analysis-to-data*. The AI system is sent to the data source and runs there. In many cases, several iterations are needed, and the algorithm eventually converges to a solution.
- *Data-to-analysis*. The data is sent to the AI system and processed there, along with data from other sources. In this manner a central data set can be created, supporting the way many traditional AI algorithms work.
- *Data-and-analysis-to-lake.* All data is collected centrally by a Trusted Third Party (TTP), i.e. a "data lake". The AI system is also sent to that same third party. The training then takes place there, and the system communicates the model outputs back. The traditional AI algorithms can therefore also be used in this collaboration model.

Important to say that AI models may contain pieces of training data to a greater or lesser extent. If this training data is confidential, that may be a problem. Therefore, distributing the AI system can be seen as sharing data, for which methods for controlled data sharing are relevant.

The production phase places less strict requirements on the location of the data. In this phase, data sources are in principle approached individually. It is then apparent to apply the "Analysis-to-Data" collaboration model, with the results returning to the AI system. It is of importance that in this phase model data and (or) knowledge data are shared. This data may still be sensitive, requiring mechanisms for controlled data sharing.

Determining the domain of data processing utilising an appropriate collaboration model has consequences for the measures we take to protect data. The collaboration model defines the roles of the parties involved: which supplies data, which processes that data through the AI system, who uses that data, where and when is data transferred between organisations. Adequate measures for protecting the shared data must be set up at each of the organisational transitions.

2.2.2. Data sovereignty: consent management and authorisations

For organisations, data sovereignty is an important starting point when sharing their sensitive data. Data sovereignty means that the data holder determines with whom and for what purposes data is shared. This is designed as an authorisation architecture, determining who has access to which data for what purpose, i.e. "consent management".

The authorisation architecture for AI applications must meet several AI-specific data sharing characteristics:

- Frequently, the data from the data provider is not available on-premise nor on the systems of the data provider itself. For example, thermostat data is sent to the energy supplier and stored in their systems. This process requires that the authorisation architecture allows for delegated permissions and authorisations.

- After data being processed by an AI system, new information is generated based on the source data, which can then be shared with other parties. The question then arises whether the original data provider owns or must have control over this new data. To address that, the authorisation architecture must coordinate the required permissions in the processing chain.
- When applying the different basic collaboration models, as described in the previous section, different types of data are shared between independent organisations at different places in the chain. The authorisation architecture must be able to deal adequately with these different collaboration models.
- In the "Analysis-to-Data" collaboration model, the AI system is implemented in the domain of the data provider. Here, the data provider determines how and whether the AI system is allowed to work with its data. Sovereignty then relates to the AI system. It must be possible to determine whether the AI system is whom it says it is (trust), and the data provider must be able to see what the system does with its data (security). This approach has been elaborated in the "Personal Health Train" [15], in which the AI system is trained on medical data from patients. The algorithm "travels" to the various hospitals, which do not want their data to leave their organisation.

2.2.3. Trust in the AI system and data processing

In addition to data sovereignty, trust in the AI system is of great importance for the willingness of organisations to share data using or through these systems. There are several different aspects to consider:

- *Transparency*. Sometimes it is sufficient to show how the AI system works (which steps and calculations are done: i.e. the algorithm itself). However, transparency and insights into the operation are only as complete as the settings (the way we use the machine learning on the training data: the model) are known. Even if the training data of an AI system is not made available, this can undermine confidence in the AI system.
- *Bias, discrimination and proxies*. To enable data-driven AI, it is necessary to have training data available that covers the necessary information and knowledge as extensively as possible. Or, in other words, the training data must be representative of all data subjects that the AI system makes decisions on. This may seem simple, but it is one of the biggest challenges, especially when it comes to data concerning (groups of) people. For example, historical and social processes may lead to ethnic groups being unjustly underrepresented / overrepresented in training data. We then speak of bias and discrimination. Examples from the criminal justice chain are now notorious [16].

To prevent or at least limit discrimination and bias, GDPR prescribes that specific personal data may not be used without a valid purpose. Consider, for example, gender for recruitment and selection applications. Such information should be taken into account when sharing data. However, due to the power of AI algorithms, it is not enough not to share these attributes from the database with the algorithm. That is because AI algorithms can retrieve information about the applicant's gender from other data, such as the desired number of working hours. The possibility to develop methodologies for these specific pre-conditions requires extra attention.

- *Data quality.* Quality control for data in AI systems is essential, partly because data is processed from different sources, of which the quality can be different. Another reason is that the output of AI systems is often further processed by other systems, which can cause a snowball effect. There is also the risk that humans take over the output of (chains of) AI systems without much critical reflection. A good grip on data quality can be obtained by monitoring from the perspective of self-protection (preventing the data provider from actually producing bad data), or from the perspective of contract monitoring (preventing that the quality of resulting data goes beyond the agreements made with the customer). Data quality can be monitored both on input and output of the AI system. An additional point of attention is that the output of AI systems usually contains uncertainties. How to deal with such uncertainties in the subsequent processing still poses a challenge.

2.2.4. Conformity and compliance

When sharing data with AI systems, data providers can impose conditions on both the access to and the use of the data. The AI system provide accountability to the data providers that have complied with these agreements when processing the data. If the data provider uses generic (public) services from third parties to register and enforce authorisations (consent management), then special attention is also needed for the potential confidentiality of the data generated in the supporting processes (the "metadata"), over which the data provider wishes to keep control and over which the processing third party must be accountable [17].

It is essential to be able to respond to questions and address complaints about the delivered AI services and products. To do so, you need to be able to find out exactly how the delivered results came about and on which data the results have been based, i.e. "traceability". Traceability can range from transparent information that provides high-level insight, through a detailed description of all data, AI models and configurations, to the (automatic) reproduction of the delivered result. Traceability is essential for the accountability of a system (can you determine which parts of the system have done what and who may be responsible) and therefore mainly has to do with trust.

In addition to characteristics of data sharing for the benefit of AI itself, there are also legal reasons why the use of AI in the context of data sharing is specific. For example, GDPR deals with automated decision-making. In such cases, transparency is absolutely required. That decision-making often uses some form of AI. That is why transparency is a further important aspect of data sharing in AI.

GDPR also sets requirements for the processing and sharing of personal data, which in particular has an impact on working with (central) training data. And of course, ethical and legal discrimination should not be allowed as indicated in the previous paragraph. Again, for AI systems, it is not enough not to share the discriminatory data, because often proxies for those data are present.

Furthermore, the principle of ‘goal limitation’ applies. This principle means that the processing organisation and the AI algorithm may only request, use, and share the (results of) the data for specific, explicitly defined and justified purposes, for which the data provider has granted permission.

Finally, privacy must be enforceable. When giving consent and in the authorisation architecture, it should be possible for the data provider to indicate for what purpose its data is made available and whether anonymity is guaranteed in the further processing and distribution of its data.

3. Approach: building on visions and frameworks

As indicated in section 1.3, the topic of data sharing is currently in the spotlight. Therefore, the NLAIC data sharing working group can (and must) build upon existing and emerging visions and frameworks on data sharing. Meaning,. This chapter describes some of these visions and frameworks. They provide input for the development process as detailed in the next chapter.

3.1. Challenges to sharing data in the AI domain

The Big Data Value Association (BDVA) is an association of research institutions focusing on data research and development in Europe. Data sharing is an essential topic for this group. In response to the European Commission Communication [18], the BDVA has released a position paper [19] identifying the critical technical challenges for data sharing from the users' perspective. These challenges arise from the ambition: "*to create a cross-border, cross-sectoral sharing data space and enabling platforms to process mixed proprietary, personal and open public data introduces new technical challenges*". The challenges are not just about the data itself, but also the metadata, the models and the use of algorithms. Although these challenges are not specially formulated with AI in mind, they are all relevant for AI. The challenges, specific to AI, are shown in Table 1.

The Strategic Research Innovation & Deployment Agenda (SRIDA, [20]) prepared by the BDVA and euRobotics identifies the main challenges to be tackled at EU level. The corresponding proposed actions are:

- To create the conditions for the development of trusted European data-sharing frameworks, building on existing initiatives (data platforms, i-spaces, big data innovation hubs). It concerns different types of data: personal, non-personal, open, closed and proprietary data.
- To promote the use of open datasets and open benchmarks for AI algorithms, especially for quality validation from software engineering and functional points of view.
- To identify specific measures aimed at including data sharing at the core of data life cycle management for better access to data, encouraging collaboration between data chain actors in both directions along the chain and across different sectors.
- To provide support for European companies to implement new technologies, practices and policies safely.
- To coordinate and harmonise Member States' efforts and realise the potential of European digital AI services in the face of global competition.
- To lead and influence standards related to data sharing tools, privacy, quality control, collaboration and interaction.

- To promote standardisation at European level but continue to work with international AI initiatives worldwide.

Table 1: Data sharing challenges in AI

Challenge	Relevance for AI
Reliability of data and quality of AI	Because AI is applied "ever deeper" in decision making, reliability is becoming ever more critical. Validation of (the origin of) data is essential. This also includes bias, especially in training data sets.
Protection of sensitive data, privacy	Because AI can extract more information from data (including implicit sensitive information), dealing with sensitivity and privacy is even more important.
Sovereignty, ownership of data	Both companies and individuals "provide" data. Given the more significant opportunities that AI provides to extract value, the importance of sovereignty under AI only increases. Secure access control is of great importance, especially with decentralised processing. Besides, AI may process data from multiple sources.
Data life-cycle management	Data may not have been generated for sharing. That is why it is extra important to pay attention to the maturity of data services (cleaning, aggregation). The distinction between "training" and "production" is of great importance for data life-cycle management. It is of great importance for the quality of AI systems that they are not trained on outdated data.
Open data	Open data provides opportunities for AI because many AI algorithms have to perform tasks that require much data (think of imagenet for training image recognition). However, it is also a potential threat because it makes it possible to re-identify anonymised data sets.
Verification and provenance	The reliability of the data itself and knowing its origin is essential for reliable results. Moreover, it also has a liability component.
Decentralised processing, access to data	Architectures to learn AI on distributed data, where access to data is even more critical because of the "data hunger" of many AI algorithms. Privacy, bandwidth and size of data play a role in consideration of decentralised processing.
Data sharing frameworks, European and international coordination	Agreement frameworks that pursue the righteous goals collectively are crucial for AI.

	Especially now that there are discussions about a "European" version of AI that differs from "capitalist" AI (US) or "totalitarian" AI (China).
Interoperability	Because AI is not only data-driven but also knowledge-driven, the semantics of data is becoming more critical; interoperability is the basis for proper semantics.

3.2. From a (closed) hub model to an (open) network model

The view on data sharing is developing. Concepts such as data sovereignty are continually receiving more attention. All of that leads to changes in the current model in which data is being shared.

Data sovereignty is an essential condition for data providers (owners) to share their potentially sensitive data. From the perspective of the providers, sharing data applies to a potentially large number of data recipients with whom they would like to share data. However, this poses a significant challenge, as the concepts for reliable data sharing are currently mainly offered from specific environments and platforms based on specific solutions. We refer to that as the hub model (Figure 1) [21]: a diversity of data sharing environments with their specific data sovereignty solutions. The hub model is often used for sector-specific, closed communities.

For the data providers, that leads to a multitude of data sharing environments to interconnect to with an associated lock-in threat to their specific data-sovereignty approaches, resulting in major integration efforts.

A single access point for the data provider with standard and agreed protocols for defining and enforcing data sovereignty can give data providers clear operational benefits regarding the efficiency and effectiveness of the management of their data sharing connections. Therefore, the network model approach is currently receiving much attention as alternative for the hub model. It provides generic infrastructure building blocks for controlled data sharing, as will be elaborated in section 3.3. It gives the data provider a single point of access to shared and agreed protocols for controlled data sharing. A network model approach has been successfully applied in the banking and telecommunications sector.

Figure 1 illustrates the transition from a solution specific hub model approach to an open, generic network model approach with infrastructure building blocks for data sovereignty.

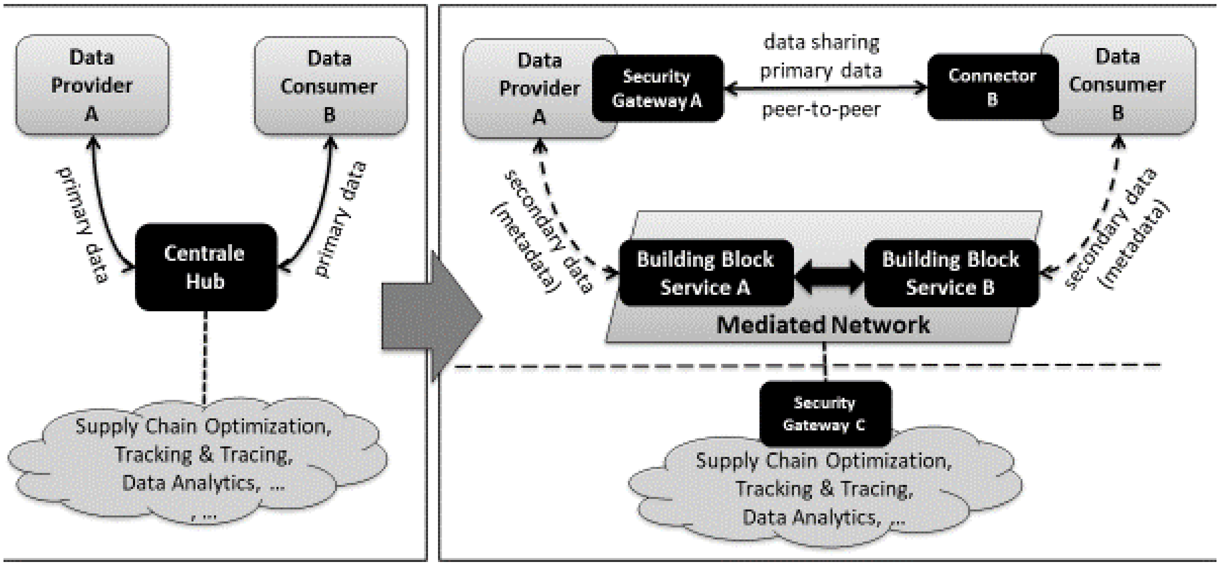


Figure 1: The hub model approach (left) and open network model approach (right) for data sharing [21].

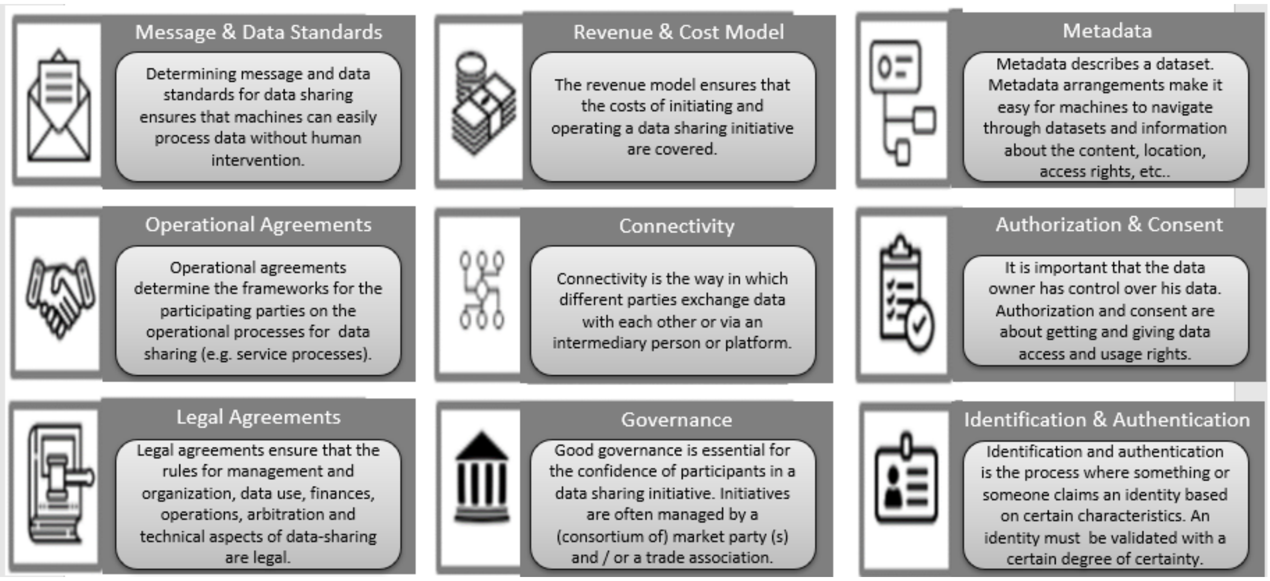
Reliable data sharing based on an open network model approach with data sovereignty is gaining momentum. The technological concepts and components to enable such a network model are currently maturing and becoming available.

The iSHARE and IDS initiatives as further explained in section 3.5 of this report, are examples of network model approaches for data sharing, which are currently receiving significant attention.

3.3. Controlled and reliable data sharing: the basic building blocks

A set of essential building blocks for enabling controlled, reliable data sharing has been identified and described in [3], as shown in Figure 2.

The essence of these data-sharing building blocks as a foundation for the emerging data economy is that they enable organisations to leverage the potential value of their data and take advantage of it.



Message & Data Standards

Determining message and data standards for data sharing ensures that machines can easily process data without human intervention.

Revenue & Cost Model

The revenue model ensures that the costs of initiating and operating a data sharing initiative are covered.

Metadata

Metadata describes a dataset. Metadata arrangements make it easy for machines to navigate through datasets and information about the content, location, access rights, etc..

Operational Agreements

Operational agreements determine the frameworks for the participating parties on the operational processes for data sharing (e.g. service processes).

Connectivity

Connectivity is the way in which different parties exchange data with each other or via an intermediary person or platform.

Authorization & Consent

It is important that the data owner has control over his data. Authorization and consent are about getting and giving data access and usage rights.

Legal Agreements

Legal agreements ensure that the rules for management and organization, data use, finances, operations, arbitration and technical aspects of data-sharing are legal.

Governance

Good governance is essential for the confidence of participants in a data sharing initiative. Initiatives are often managed by a (consortium of) market party (s) and / or a trade association.

Identification & Authentication

Identification and authentication is the process where something or someone claims an identity based on certain characteristics. An identity must be validated with a certain degree of certainty.

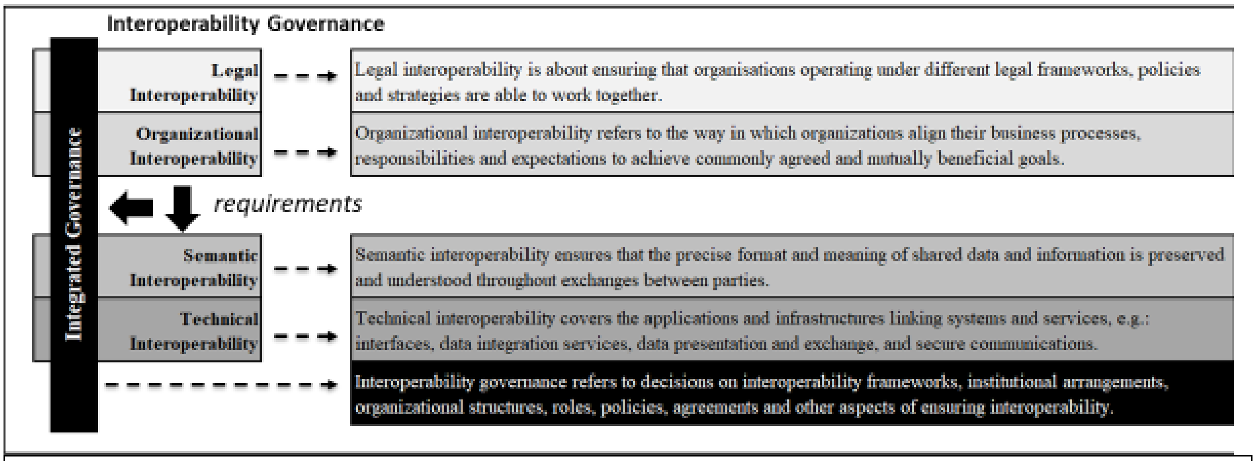
Figure 2: Essential building blocks for controlled and reliable data sharing [3].

3.4. Interoperability: governance, legal, organisational, semantic and technical

Responsible and controlled data sharing present a multidisciplinary challenge. Network model approaches for sharing data across AI applications, organisations, and sectors can be characterised as federated environments (or 'systems-of-systems'), in which a variety of specific intermediary orgnizations bundle their capabilities to enable controlled data sharing while preserving data sovereignty. Interoperability is essential for broad adoption and easy integration.

Different frameworks have been developed to achieve interoperability for such federated environments. An interoperability approach that is often used is provided by the (new) European interoperability framework, as developed by the European Commission [22]. As depicted in Figure 3, it distinguishes four interoperability levels to be implemented under an overarching governance approach: legal, organisational, semantic and technical interoperability.

The levels of interoperability should be addressed in an overarching approach for controlled and reliable data sharing for those cases where interoperability between sectors, applications, countries or jurisdictions is pursued.


Interoperability Governance
Integrated Governance
Legal Interoperability
Organizational Interoperability
requirements
Semantic Interoperability
Technical Interoperability

Legal interoperability is about ensuring that organisations operating under different legal frameworks, policies and strategies are able to work together.

Organizational interoperability refers to the way in which organizations align their business processes, responsibilities and expectations to achieve commonly agreed and mutually beneficial goals.

Semantic interoperability ensures that the precise format and meaning of shared data and information is preserved and understood throughout exchanges between parties.

Technical interoperability covers the applications and infrastructures linking systems and services, e.g.: interfaces, data integration services, data presentation and exchange, and secure communications.

Interoperability governance refers to decisions on interoperability frameworks, institutional arrangements, organizational structures, roles, policies, agreements and other aspects of ensuring interoperability.

Figure 3: Interoperability model as defined in the 'new European Interoperability Framework' [22].

3.5. Data sharing technologies: overview and relevance to AI

There are several relevant technologies available for the technical design of infrastructures for controlled and reliable data sharing for AI applications. Four categories may be distinguished:

- Data sharing architectures with generic building blocks according to the network model.
- Security techniques to protect data while sharing it.
- Security techniques to protect the data before or after the AI application.
- Adjacent techniques that are relevant, but that do not directly protect the confidentiality of data.

Table 2 shows several examples of technologies for each of these categories.

Table 2: Overview of data-sharing technologies

Data Sharing architectures	**Adjoining techniques**
International Data Spaces (IDS) A European initiative to enable (controlled) data sharing in a standardised manner based on a reference architecture for security gateways and accompanying intermediary roles.	*Distributed ledger* A decentralised way to store large amounts of transactions (contracts, documents, etc.), guaranteeing that the stored information is immutable. Blockchain technology is an example of such a technique.
iSHARE	*Ontology-Based Access Control (OBAC)*

A Dutch initiative stemming from the logistics sector, consisting of a set of aligned agreements for identification, authentication and authorisation.	OBAC is a tool to get structure in extensive data sets with high diversity. It provides options to control access to the data.
Security techniques during sharing	**Security techniques before/after AI**
Secure Multi-Party Computation (MPC) A collection of innovative cryptographic technologies that allows multiple parties to work together with data, as if they have an extensive database to work on, but without accessing each other's data.	*Differential privacy* A way to prevent the results of statistical analyses from being traced back to individuals, by adding noise to personal data. It can be used in combination with federated learning or MPC to protect the output of AI algorithms.
Federated learning A form of machine learning on distributed data, in which the algorithm is designed in such a way that the data exchanged between parties is less sensitive.	*Anonymize and pseudonymize* Techniques to prevent data from being traced back to individuals, by removing or scrambling identifiable information in the data.

4. Development process: from “IST” to “SOLL”

Data sharing architectures are in an emerging phase. Many organisations need guidance to use their advantages, especially with newly available technologies. In the context of the NLAIC, it is possible to develop new ideas, architectures and concepts on data sharing for AI.

This chapter describes the development process from the current “IST” situation to the future “SOLL” situation. To this end, the following subsequently sections describe the role model / ecosystem approach for the design of the generic data-sharing infrastructure, the process from first-time engineering to operationalisation and the initial, representative, application scenarios and proof-of-concepts (PoC’s) as the first step in the follow-up process of the NLAIC working group on data sharing.

4.1. Ecosystem for data sharing: building blocks and roles

There is no “one-size-fits-all” approach for data sharing in AI. Different objectives, interests and perspectives mean that different data-sharing approaches will be followed. It is important that the various approaches can be supported by a diversity of data sharing functions. By making these available as generic and reusable building blocks in an (open) network model, we avoid dependence on (closed) platforms with potential lock-in, see section 3.2.

An ecosystem of roles is defined to make the diversity of approaches to data sharing in AI and solutions transparent and manageable. This ecosystem can jointly supply the required building blocks, see Figure 4. The roles are developed and validated together in coherence.

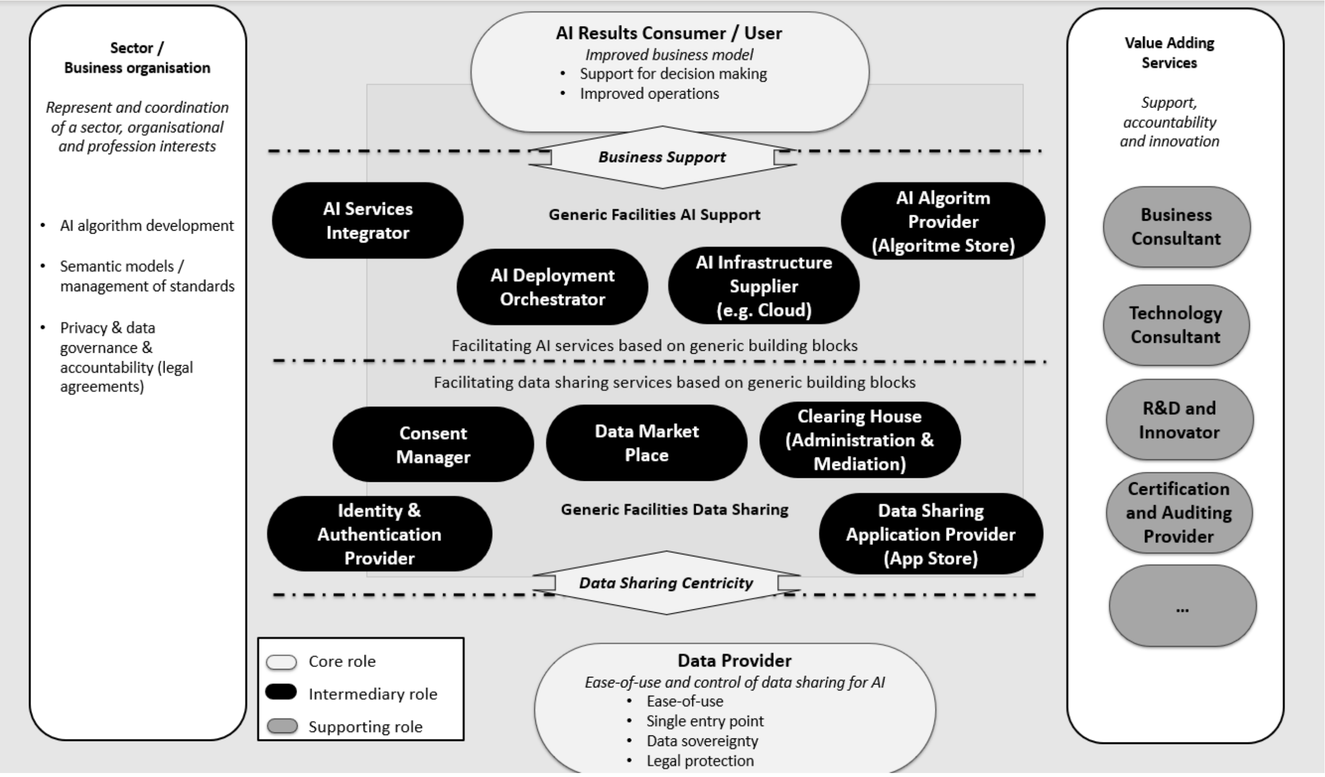

Sector /
Business organisation
Represent and coordination of a sector, organisational and profession interests
AI algorithm development
Semantic models / management of standards
Privacy & data governance & accountability (legal agreements)
AI Results Consumer / User
Improved business model
Support for decision making
Improved operations
Business Support
AI Services Integrator
Generic Facilities AI Support
AI Algoritm Provider (Algoritme Store)
AI Deployment Orchestrator
AI Infrastructure Supplier (e.g. Cloud)
Facilitating AI services based on generic building blocks
Facilitating data sharing services based on generic building blocks
Consent Manager
Data Market Place
Clearing House (Administration & Mediation)
Identity & Authentication Provider
Generic Facilities Data Sharing
Data Sharing Application Provider (App Store)
Data Sharing Centricity
Core role
Intermediary role
Supporting role
Data Provider
Ease-of-use and control of data sharing for AI
Ease-of-use
Single entry point
Data sovereignty
Legal protection
Value Adding Services
Support, accountability and innovation
Business Consultant
Technology Consultant
R&D and Innovator
Certification and Auditing Provider
...

Figure 4: The role model (ecosystem) for data sharing for AI

As Figure 4 indicates, the data-sharing ecosystem for AI distinguishes between "core roles", "intermediary roles" and "supporting roles":

- *Core roles.* The core roles are the direct stakeholders in data sharing. These are the parties between which primary data is shared, i.e. the data providers and the AI users.
- *Intermediary roles for generic facilities*. The intermediary roles provide the support functions as reusable building blocks that facilitate data sharing between the core roles. As Figure 4 shows, we distinguish intermediary roles for data sharing and intermediary roles for AI support. These roles should not have access to primary data. They may process the secondary data required and generated by the support processes, also known as "metadata". The intermediary roles are usually fulfilled by trusted organisations, i.e. so-called "Trusted Third Parties" (TTPs).
- *Supporting roles.* These roles do not perform any activity in the actual data sharing process but may be necessary for the system to function accordingly.

Concerning these points, it is good to refer to some international initiatives currently related to the approach described in this section. Google offers its AI platform [23] that provides functions and building blocks for AI support. Deutsche Telekom provides the concept of the “Data Intelligence Hub” [24] with functions and building blocks for both AI support and data sharing. In addition, the GAIA-x initiative [12] is currently receiving significant attention. This initiative was launched at the end of 2019 in Germany to develop a European cloud infrastructure as an alternative to the major platforms in the US and Asia.

4.2. From first time-engineering to operationalisation

It is essential to work on shared infrastructures for sharing data in AI to (i) strengthen the power of the Dutch AI landscape, (ii) to increase its potential and possibilities and (iii) to prevent vendor lock-in. The aim is to create a robust AI infrastructure for the Netherlands, in which organisations can quickly develop impactful applications through collaboration. To this end, this section outlines the process of how companies can share data for AI, from experimental (“first-time engineering”) phase, to a phase of daily practice (“operationalisation”). A more detailed elaboration is included in a separate report by the data sharing working group of NLAIC [25].

The process from first-time engineering to operationalisation is based on the development process referred to as the “strategic options model” - shown in Figure 5. It describes the development for digital services from experimental first-time engineering (indicated in grey) to operationalisation (indicated in black). Here, we distinguish between the generic technical infrastructure (the “mediated network”) that is required to support (data sharing for) AI and the applications of the new technology itself. In addition, the generic technical infrastructure contains both the data sharing and the AI-facilitating layer as depicted in Figure 4, which contain reusable building blocks that are required for the safe and responsible sharing of data with various parties and supporting AI applications. This infrastructure can contain both physical components (e.g. hardware) and abstract elements, such as an agreement system or software.

For some parties, it may be extra valuable to gain knowledge and experience in developing, running and facilitating the organising infrastructure, while others are looking for experience in developing new AI applications that can use that.

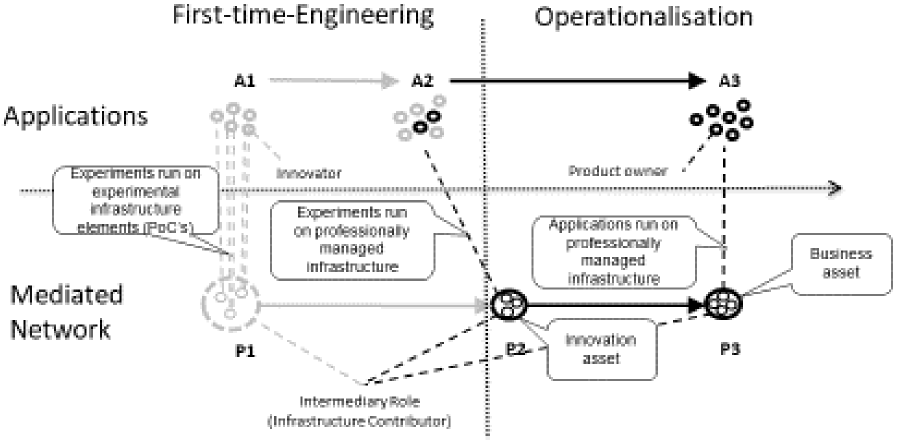

First-time-Engineering
Operationalisation
A1
A2
A3
Applications
Innovator
Product owner
Experiments run on experimental infrastructure elements (PoC's)
Experiments run on professionally managed infrastructure
Applications run on professionally managed infrastructure
Business asset
Mediated Network
P1
P2
P3
Innovation asset
Intermediary Role (Infrastructure Contributor)

Figure 5: Strategic Options Model. The infrastructure layer contains both data sharing and facilitating AI functions which jointly enable AI applications.

The role of "infrastructure contributor" is reserved for parties who make building blocks available as a service and together make the infrastructure into a functional whole. Think of building blocks in the data-sharing layer as well as building blocks in the AI-facilitating layer, as shown in Figure 4. This role can be assumed by a multitude of parties' side by side.

In the process towards operationalisation, parties can start developing phase "I1": one experimentally shared infrastructure for running multiple experimental applications. Based on several initial use cases, this can be designed as a collaboration of various parties ("front-runners") with a common goal to develop AI applications with specific technical use cases and the technical infrastructure in mind.

It is essential for the speed and quality of joint AI development in the Netherlands that the initial use cases are not only there to learn from each other, but also as an impetus to be able to continue building towards operationalisation. So, a use case provides insight into the application as well as guidance for an initial (reference) architecture and implementation for reusable building blocks.

As shown in Figure 5, as part of the process, three phases are distinguished in infrastructure and AI applications, as described in Table 3.

Infrastructure (I)	**AI-applications (A)**
I1: An experimental infrastructure for sharing data, based on generic building blocks in the PoC, to learn how to set up and maintain such an infrastructure. This also includes the deployment and management of AI applications based on this data. Developers of AI applications have to take into account that this infrastructure will not yet meet all their requirements, such as availability, stability, or security.	A1: Experimental AI applications, primarily designed to develop skills for data sharing.
I2: A data sharing infrastructure that is sufficiently mature to be offered as a service to users who want to experiment with developing and running AI applications. These AI applications place new demands on data sharing such as access, control, security and feedback. It is an innovation asset that enables innovative AI applications based on shared data.	A2: Experimental AI applications, mainly intended to gain experience in developing AI applications in preparation for real-life production.
I3: This is an operational, generic, data sharing infra-structure for exploitable AI applications. It is a business asset, which must exceed the requirements of I2 because exploitable applications depend on it. Other business risks and agreements also apply – not covered by I2.	A3: Exploitable AI applications that are reliable and stable enough to be available to end-users in exchange for something of monetary value.

FTable 3: Phases for the generic infrastructure (“I”) and AI Applications (“A”) in the process from first-time engineering to operationalisation

In the process, the partners of NLAIC actively participate in the joint task of drawing up a roadmap for going through these phases for infrastructure and AI applications. The following three aspects are distinguished in the roadmap:

- *Technology.* Designing the architecture necessary for sharing data, defining interfaces and information models and demonstrating and realising use cases.
- *Organization (governance).* Focuses on the (further) development of the proposed approach, architectures, interfaces and standards. Includes both the process design for a governance and organisational structure, and setting up a change (advisory) board for the technical roadmap.
- *Ecosystem.* Aimed at the adoption by organisations and market parties. Since we are dealing with a multi-party implementation, we are talking about a complete

ecosystem. It is crucial to facilitate ecosystem dynamics as much as possible in preparation for growth. By scaling up both the number of parties that share data responsibly with each other and the AI applications themselves, the positive impact for the end-users is increasingly being felt.

Figure 6 shows a conceptual roadmap for the development process from first-time engineering to operationalisation. The individual activities are further described in [25].

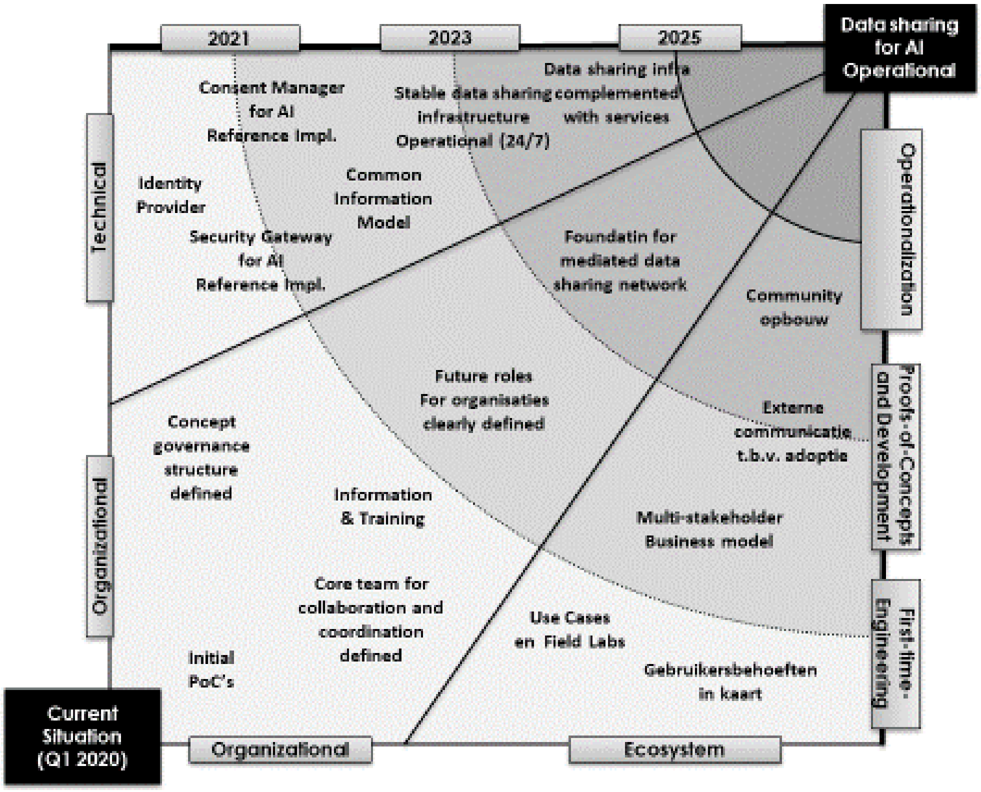

2021
2023
2025
Data sharing for AI Operational
Consent Manager for AI Reference Impl.
Stable data sharing infrastructure Operational (24/7)
Data sharing infra complemented with services
Technical
Identity Provider
Common Information Model
Security Gateway for AI Reference Impl.
Foundatin for mediated data sharing network
Operationalization
Community opbouw
Future roles For organisaties clearly defined
Proofs-of-Concepts and Development
Externe communicatie t.b.v. adoptie
Concept governance structure defined
Organizational
Information & Training
Multi-stakeholder Business model
Core team for collaboration and coordination defined
First-time-Engineering
Use Cases en Field Labs
Initial PoC's
Gebruikersbehoeften in kaart
Current Situation (Q1 2020)
Organizational
Ecosystem

Figure 6: Concept roadmap for the development process from first-time engineering to operationalisation.

4.3. Application scenarios and proofs-of-concept

In the trajectory towards operationalisation, the ideas, architectures and concepts around data sharing for AI are shaped using a selection of application scenarios and technical Proofs-of-Concept (PoCs). The application scenarios and the PoCs are aimed at demonstration and validation of a typical data-sharing infrastructure. The goals of the PoCs are to:

- Further develop and technically validate the approach to data sharing according to the role model / ecosystem, as shown in Figure 4, with generic roles and building blocks for both the “Generic Facilities for Data Sharing” layer and for the “Generic Facilities for AI Support” layer. Preferably, this is done independently of sector or specific application, so that the infrastructure can be offered in a cross-sector manner.

- Determine whether this approach contributes to the willingness of organisations to come to a way of collaborating and thereby sharing (sensitive) data for the benefit of AI applications.

Together with the sector representatives in the NLAIC, AI application scenarios are chosen for which the data sharing scenarios are developed as PoCs. AI application scenarios with significant impact are chosen. The industry organisations are involved in the preparation and realisation of the AI application scenarios and PoCs. The results are shared with the members of the coalition.

There is no one-size-fits-all approach for structuring data sharing in the AI domain. It is, therefore, not the intention and possible to define a single application scenario that covers the entire playing field of (controlled) data sharing for AI. Hence, an approach is chosen based on several illustrative and representative application scenarios and associated technical PoCs in which the PoCs technically demonstrate the different perspectives on the role model as shown in Figure 4 and its challenges. This approach is illustrated in Figure 7.

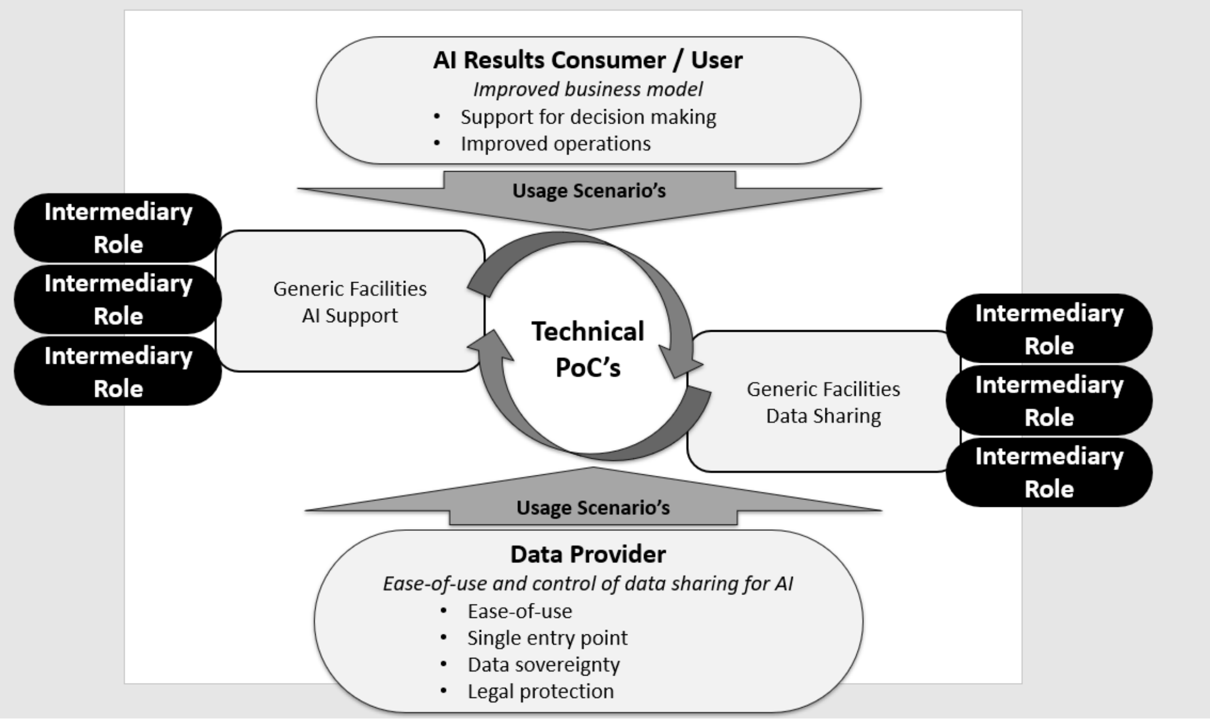


Figure 7: Identification and definition of technical PoCs to support the development process from first-time engineering to operationalisation.

Colophon

F.T.A. (Frans) van Ette, MSc
H.J.M. (Harrie) Bastiaansen, PhD
P.J.M. (Thijs) Veugen, PhD
C.J. (Cor) Veenman, PhD
F.W. (Freek) Bomhof, MSc
Prof. dr. ir. D. (David) Langley
S. (Somayeh) Djafari, LL.M
L. (Laura) van den Aarssen, PhD
M.J. (Marc) van Lieshout, MSc

NL AI Coalition
Coalition desk – info@nlaic.com
NLAIC.com

Translation
Eduardo Wilde Barbaro (IBM)