CoE
DSC

Governance design for MPC data collaboration

Case study on the data collaboration in elderly care sector monitoring impact of policies and measures taken on the state of care

Centre of Excellence for Data Sharing and Cloud

Key findings of the case study

1. The Dutch elderly care sector would highly benefit from data collaboration on a regional level to generate statistical insights to measure impact of Dutch policies on elderly care (WMO, WLZ, ZVW), and to provide benchmarks for care providers and municipalities
2. Most of the data used to generate such statistics is privacy-sensitive and therefore it is difficult to achieve trust for a data collaboration
3. Multi-Party Computation (MPC) is selected as technology to organise trust for relatively low costs while safeguarding data privacy
4. Linksight (MPC provider), DSW, Delft Municipality and Pieter van Foreest collaborate through MPC to generate statistics in Delft, Westland, Schieland region, and there is a potential to scale the pilot to other Dutch regions
5. Scaling up however, is expected to create tensions between participants in the network. Namely, Data Stewards on behalf of participants want to have control over the data, while Data Scientists on behalf of participants want to adopt new use cases and generate valuable statistical insights
6. The dynamic cannot be resolved by MPC alone and requires a governance framework to be in place
7. Centre of Excellence – Data Sharing and Cloud (CoE-DSC) supported Linksight in developing governance mechanisms that allow maintaining trust, pace and control on data in the growing network where participants make various requests (e.g. requests to join the network, requests to develop new use cases, to access insights etc.)
8. The proposed governance framework for the MPC data collaboration consists of:
   - Baseline mechanisms per all types of requests (e.g., digital identity procedures, contracting, rules etc.)
   - Additional mechanisms depending on whether participants in the compute group want to (A) exercise direct control, (B) delegate control to a trusted party to maintain pace, or (C) have a tailor-made compromise for control and pace
9. The study group will further discuss the implementation of governance with Linksight and involved stakeholders
10. The developed framework can be re-used as a blueprint for organising governance in other MPC data collaborations

Most important terms used in this document

Term	Explanation
Authentication	The process where the validity of a claimed identity is verified
Authorisation	The permissions or rights of an actor (humans, machines, proxies, etc) to perform an action
Compute group	A group of data providers in a data collaboration that run computational nodes to analyse data through MPC. The compute group operates in accordance with established governance rules and a DPIA
Data collaboration	A collective body for analysing data and sharing insights for some pre-defined purpose (e.g. elderly care monitoring). Typically, consists of a group of participants with various roles and responsibilities
Data Provider	A party in the data collaboration that provides data for generating insights
Delegation	The provision of explicit rights to perform an action on behalf of some party
DPIA (Data Protection Impact Assessment)	A process under the GDPR that includes identifying objectives risks of processing personal data and serves to ensure compliance in any data collaboration, usually this process includes completing several steps and signing an agreement between participants (for more information read here)
Enabling role/component	A role in the data collaboration that supports participants in generating and sharing insights (i.e. MPC provider enables the analysis by setting up technology)
Focal role	An umbrella term for roles of the direct participants in the data collaboration (e.g. a Data Provider, MPC Beneficiary, Data Scientist etc.)
Governance Framework	A trust framework that enables many-to-many transactions though business, legal, operational, functional, and technical agreements, tools, and processes which facilitate trusted transactions between participants in a data sharing context
Governance rules	A set of rules applied in the data collaboration regarding data access, data usage, scope of data analysis and roles and responsibilities of participants
Identification	The process of attributing/issuing an identity to a subject by an authority. This includes issuing a digital identity after physical identity has been verified for example during an onboarding process
Multi-Party Computation (MPC)	A type of privacy enhancing technologies where computations are securely run at each party ensuring that the source data remains private and only insights are shared
MPC Provider	A party in the data collaboration that enables PETs, typically by providing the infrastructure to run computations
MPC Beneficiary	A party that relies on the insights from data analysed using PETs
Privacy Enhancing Technology (PET)	A technical implementation that enables analysis of data in a way that sensitive data remains protected, and secure

Source: CoE-DSC analysis

Table of Contents

1. **Introduction to case study**
2. Proposed governance solution
3. Lessons learned and next steps
4. Appendix

Dutch elderly care sector can improve care provisions by collaborating on data to generate statistical insights

Improving elderly care is high on Dutch political agenda

The Dutch population is aging, which strains the elderly care provisions. Between 2015 and 2020, care costs for municipalities increased by 30% and waiting lists have surged due to the lacking capacity of the system

Binnenlands Bestuur

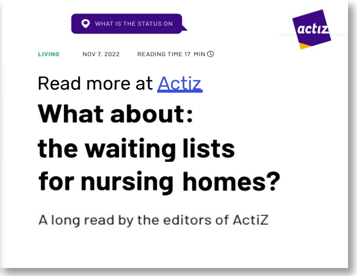
Read more at [Binnenlands Bestuur]

social / news

Wmo costs for municipalities 'explode'

The social domain is developing into an increasingly difficult millstone for municipalities. Wmo costs in particular are rising sharply.

Hans Becker April 11, 2022



WHAT IS THE STATUS ON

actiz

LIVING NOV 7, 2022 READING TIME 17 MIN

Read more at [Actiz]

What about: the waiting lists for nursing homes?

A long read by the editors of ActiZ

Rijksoverheid

Home > Onderwerpen > Kwaliteit van de zorg >

The Integral Care Agreement - working together on healthy care

Accessible, good and affordable care is important. It is becoming increasingly difficult to continue to guarantee this

Read more at [Rijksoverheid]

Data collaboration to generate statistics

- To ensure elderly care remains affordable and accessible in the Netherlands, sector statistics need to be monitored on provision costs, budgets, availability of personnel, treatment effectivity etc.
- Data collaboration would allow continuous monitoring of these statistical insights for care providers and municipalities to benchmark against and measure effectiveness of policies (WLZ, ZVW, and WMO)*

Reasons why MPC is used

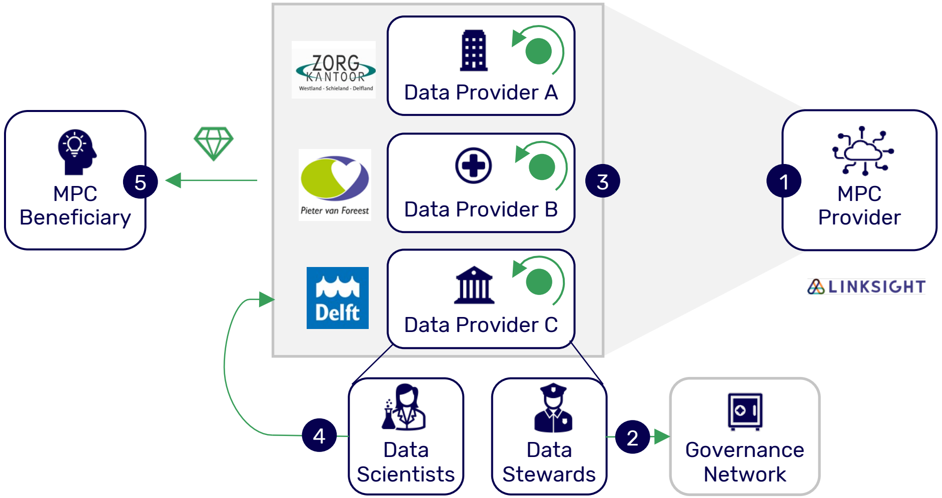
In the elderly care monitoring use case, Multi-Party Computation (MPC) [link to explanation on MPC] elevates several challenges for a data collaboration:

- **Privacy.** MPC ensures privacy for using sensitive patients' data under GDPR
- **Costs.** Through automation reduced high costs of manual analysis to gain performance statistics
- **Fragmentation**. MPC provides technical infrastructure to facilitate computations across fragmented stakeholders involved

Sources: *For more on WLZ, ZVW and WMO see [Ministerie van Volksgezondheid, Welzijn en Sport] and appendix p. 21

Linksight, DSW, Delft Municipality, Pieter van Foreest collaborate through MPC to generate statistical insights on elderly care

Set-up data collaboration



Linksight, DSW, Delft Municipality and Pieter van Foreest collaborate to generate statistical insights in the following way:

1. The MPC Provider supplies the software to the Data Providers, consisting of computational and governance nodes
2. The Data Stewards on behalf of the Data Providers jointly agree on governance rules and store them in the Governance Network
3. The Data Providers form a compute group to generate statistical insights
4. The Data Scientists make queries for the compute group
5. The MPC Beneficiaries receive generated insights in a form of a dashboard (see appendix p. 22)

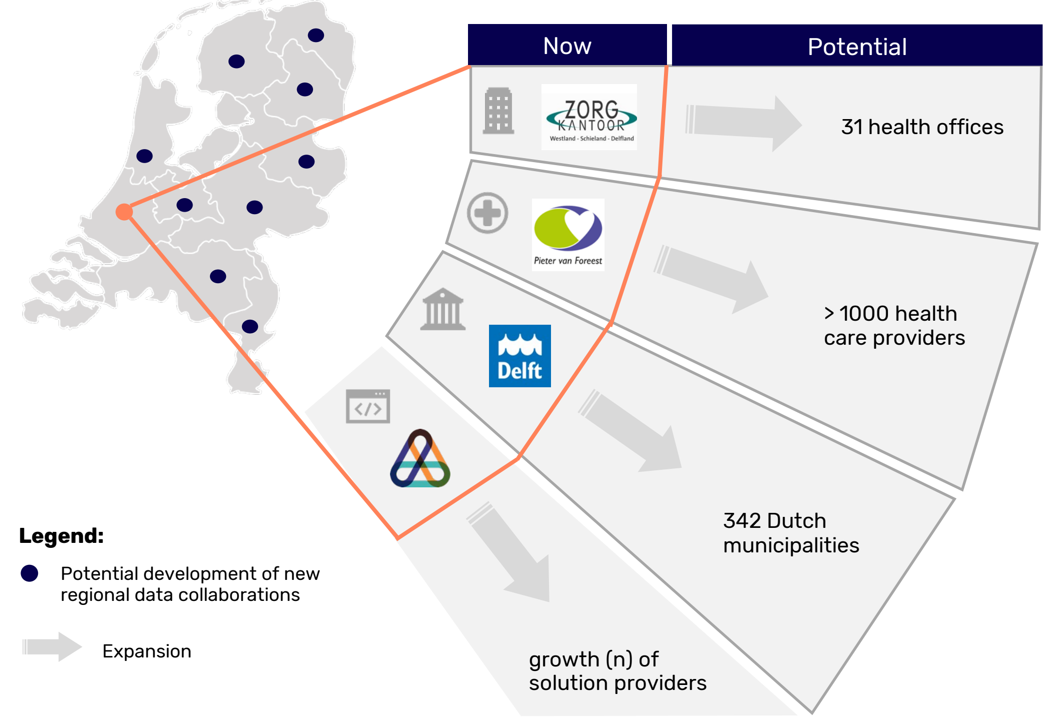
Example statistical insights for MPC Beneficiary:

- € spendings on elderly patients with cardiovascular diseases in Delft region
- € saved due to specific changes in policies WLZ, WMO, ZVW
- € saved due to providing hip protection pads to elderly people
- Avg. # of years elderly patients spend in a long-term care ward
- Avg. # of patients with Parkinson's admitted for short-term care
- Distribution of patients with dementia across nursing homes

Note: One organisation can be both MPC Beneficiary and Data Provider; Data Scientists could also be a trusted third party

There is potential to scale current data collaboration to other Dutch regions, to holistically monitor statistics for elderly care

High-level roadmap


Now
Potential
ZORG KANTOOR
Westland - Schieland - Delfland
31 health offices
Pieter van Foreest
> 1000 health care providers
Delft
342 Dutch municipalities
growth (n) of solution providers

Legend:

- ● Potential development of new regional data collaborations
- ➡ Expansion

Explanation

Scale up data collaborations to other regions is relevant because:

- A regional cooperation is encouraged in the National Care Accord (IZA) to monitor healthcare system performance
- Monitoring care within Dutch regions is needed from the national government perspective, since the policies impacting elderly care (WLZ, WMO and ZVW) are set nationally
- More participants in a sector can be involved and benefit from data collaborations:
  - **31 "Zorgkantoren"** for long term care (Wet langdurige zorg: WLZ)
  - **342 Municipalities** for short term care (Wet Maatschappelijke Ondersteuning: WMO)
  - **1000+ elderly/home/social** care organisations

Source: CoE-DSC analysis; for more on the National Care Accord (IZA) refer to Rijksoverheid

Scaling up collaboration creates tension between participants that MPC technology cannot resolve on its own

Network roles have various interests

Data Scientists and MPC Beneficiaries have conflicting interests with Data Stewards

Data Scientist(s)

I want to generate new insights by doing new analyses with new data

MPC beneficiaries

I want to access new generated insights

Data Steward(s)

I want to keep control over the provided data

Governance is required to deal with interests

- **Participants pose new change requests.** MPC data collaboration is not static, often new requests are put forward by participants. For example requests for new use cases, for joining the network, for new insights
- **Data Stewards manage control and compliance.** Data Stewards review requests to make sure who can access what data, use the data / insights for which purpose, in accordance with regulation
- **MPC technology requires governance.** MPC does not control compliance when changes are made in the data collaboration and thus, data access and data use require additional governance

Current governance setup

- In the current setup, Data Stewards are tasked with assessing requests:
  1. Data Providers request to join the compute group
  2. Data Providers request to leave the compute group
  3. Data Scientists request to run queries*
  4. Data Scientists request to add new use cases
  5. MPC Beneficiaries request to access insights
- Data Stewards vote on requests unanimously to accept changes in accordance with GDPR and established governance rules
- Statistical disclosure control is ensured via governance rules, i.e. under which circumstances results are allowed to be shared and with whom

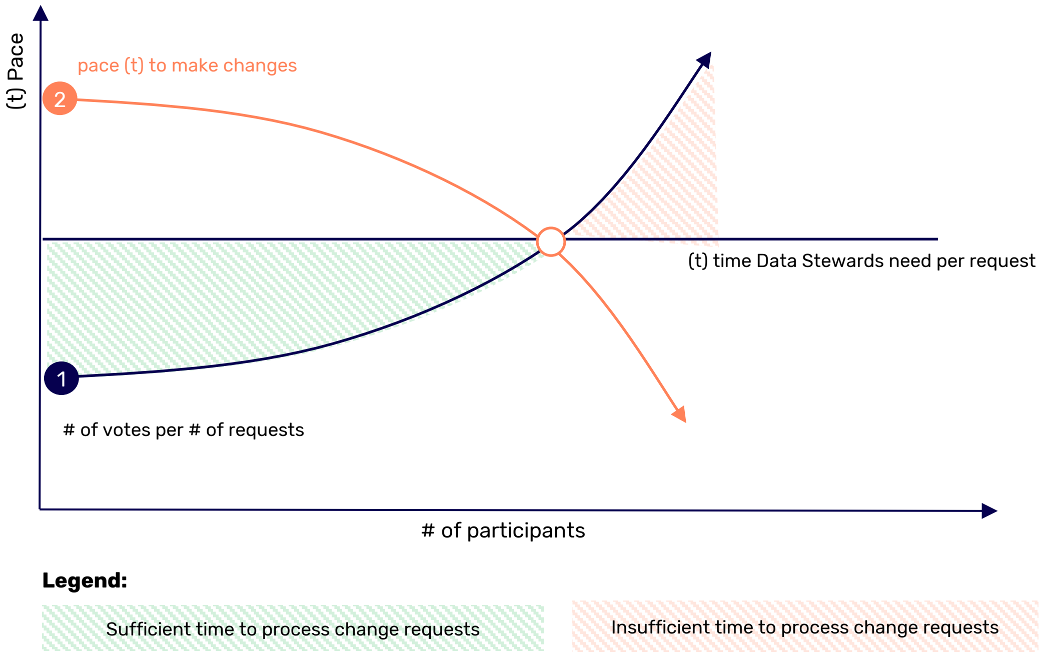
See the process in appendix p. 26

Note: *Allowed queries are covered by governance rules, and are automatically assessed

Source: CoE-DSC analysis

If current governance set-up is not changed, Data Stewards become a bottleneck in scaling the network

Consensus voting on all requests poses bottleneck for Data Stewards



Explanation

Diagram shows two effects:

1. **Network growth means more change requests.** As the number of participants grows, both the number of change requests and number of votes required per change request grow accordingly
2. **Slow pace of change due to Data Stewards handling all change requests.** The result of consensus voting is a fast decline in pace of change, as Data Stewards will individually need the time to process all change requests and thus become a bottleneck

Conclusions

1. Data Stewards voting on all change requests is a governance set up that prevents scaling up
2. A new governance framework is required, possibly including new roles
3. A new governance framework should keep balance between pace and control

Source: CoE-DSC analysis

CoE-DSC supports Linksight, DSW, Delft Municipality and Pieter van Foreest in designing a new governance framework

Key Questions covered by CoE-DSC in the case study

1. How can a governance set-up for MPC data collaboration be arranged?
2. How does the new governance set-up cope with requests made by participants?
    1. Data Providers request to join the compute group
    2. Data Providers request to leave the compute group
    3. Data Scientists request to run queries
    4. Data Scientists request to add new use cases
    5. MPC Beneficiaries request to access insights
3. How does the new governance manage different interests between network participants, in particular “keeping pace in handling requests versus keeping strict control over data”

Source: CoE-DSC analysis

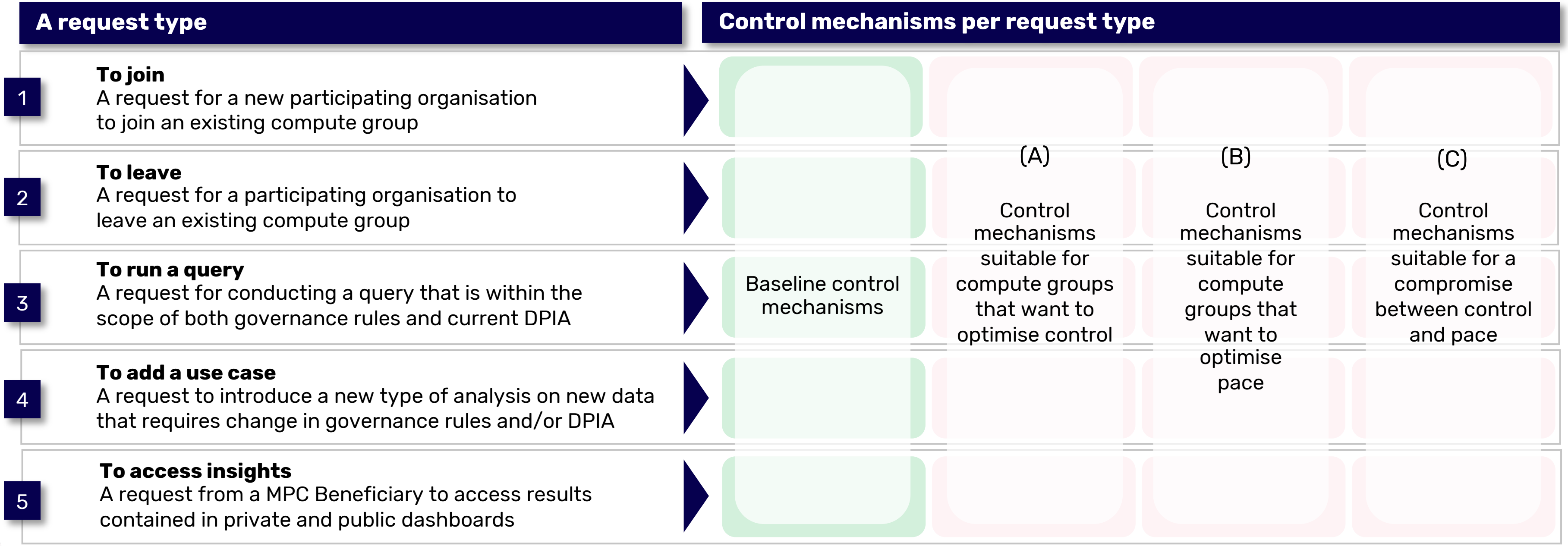
Table of Contents

1. Introduction to case study
2. **Proposed governance solution**
3. Lessons learned and next steps
4. Appendix

Governance includes baseline control mechanisms per change request and additional ones for the specific group preference

Introduction to the new governance

The new governance is composed of a set of mechanisms per request type. For each type there are baseline mechanisms and 3 additional clusters of mechanisms catering for a compute group preferences (A,B,C) varying on control, pace or a compromise between control and pace



A request type

1 **To join**
A request for a new participating organisation to join an existing compute group

2 **To leave**
A request for a participating organisation to leave an existing compute group

3 **To run a query**
A request for conducting a query that is within the scope of both governance rules and current DPIA

4 **To add a use case**
A request to introduce a new type of analysis on new data that requires change in governance rules and/or DPIA

5 **To access insights**
A request from a MPC Beneficiary to access results contained in private and public dashboards

Control mechanisms per request type

Baseline control mechanisms

(A)

Control mechanisms suitable for compute groups that want to optimise control

(B)

Control mechanisms suitable for compute groups that want to optimise pace

(C)

Control mechanisms suitable for a compromise between control and pace

Source: CoE-DSC analysis

Overview of baseline mechanisms ensuring control for requests in the data collaboration

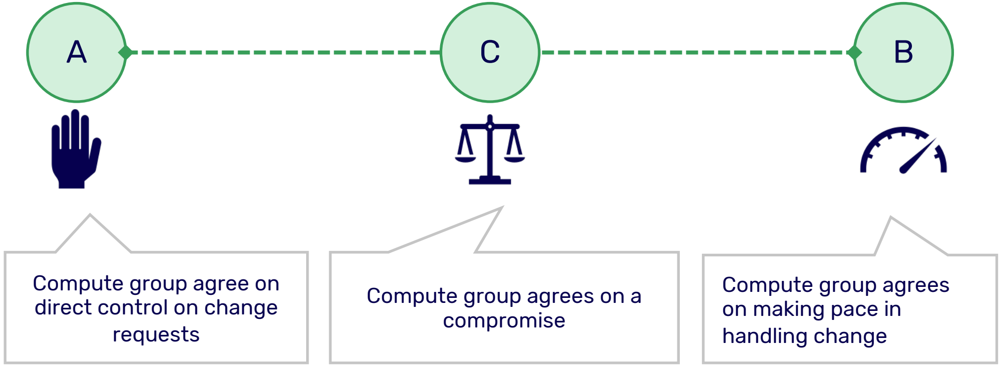
Request type	Selected control mechanisms	Facilitated by
1 **Request to join**	• Any IAA mechanisms to onboard, authenticate and verify identities of participants (e.g., MS365/Google authentication means, eHerkenning, eIDAS login means)	• Identity Provider *
	• A participant contract with an MPC Provider confirming conditions for network participation	• MPC Provider
	• A joint controller agreement with the Collaboration Authority, instead of bilateral contracting	• Collaboration Authority *
	• A signed DPIA to comply with GDPR regulation	• Data Steward
2 **Request to leave**	• A period of notice arranged by the Collaboration Authority to allow other participants in a compute group to finalise ongoing computations and prepare for closing/forking of the compute group	• Collaboration Authority *
	• An emergency button for force majeure circumstances allowing a Data Provider to halt their MPC node to stop provision of the data	• MPC Provider
3 **Request to run a query**	• An automatic procedure in the software that checks if a query is within the established governance rules	• Governance Network
	• A consent button sent to Data Scientists for each query as part of terms and conditions	• Governance Network
	• A rulebook for handling misconduct and incident management to prevent data misuse and fraud	• Collaboration Authority and collaboration/compute group Administrator(s) *
	• Any penalties/liabilities for Data Scientist in case of misconduct (e.g., malicious queries that go beyond purpose, circumventing established governance rules)	
4 **Request to add a use case**	• A Collaboration Administrator chairs the discussion meetings where participants collectively make decisions, and set timeframes for changing process of governance rules and/or DPIAs	• Collaboration Administrator(s) *
	• A Data Stewards creating new governance rules, and/or make amendments in their Data Provider's DPIAs	• Data Steward
	• An automatic procedure tracks and saves changes made to the governance rules (e.g., audit trails)	• Governance Network
5 **Request to access insights**	• The login credentials and authorisation mechanisms for MPC Beneficiaries are in place, for example: • Re-using digital identity means issued during the onboarding of a participant • Issuing specific credentials for the access to non-public dashboards	• Identity Provider
	• A procedure for creating public dashboard that does not contain any sensitive data and is up-to-date	• Dashboard Administrator(s) *

Source: CoE-DSC analysis

*Note: new roles are introduced to facilitate control mechanisms, see more on p. 16

On top of baseline, a compute group selects preference for additional mechanisms to balance ‘control’ and ‘pace’

Three options on top of baseline



Compute group agree on direct control on change requests

Compute group agrees on a compromise

Compute group agrees on making pace in handling change

Description

For additional governance mechanisms, compute groups can select one of three options:

A. Optimising direct control. In option A Data Stewards take direct control over change requests. Compute groups tend to select A when participants have low trust and familiarity among each other, and deal with sensitive heterogenous datasets

B. Optimise pace: In option B participants delegate control to someone in the group, to make sure change requests are quickly processed. Compute groups tend to select B when participants have high trust, familiarity, shared interests, and deal with homogenous datasets

C. Compromise: In option C both direct control and delegation are used by Data Stewards. Compute groups select C when interests vary, and data is sensitive and heterogenous.

See pp. 27-28 on rationale for compute group preferences

Source: CoE-DSC analysis

Overview of specific mechanisms under options A, B, C for compute groups to choose

Request type	A. Optimise direct control	B. Optimise pace of change	C. Compromise	Facilitated by
1 **Request to join**	**Applies to 1, 4, 5:** • **Full consensus** voting to accept requests and agree upon new governance rules • **Automated push notifications** to remind Data Stewards of their voting duties • **Temporary stop for a particular participant or computation** until a decision is made about a request. There is transparency in the status of decision making (e.g., clear deadlines to formalise the progress)	**Applies to 1, 4, 5:** • All participating organisations **delegate voting to a trusted third party** (i.e. Collaboration Administrator) • **Fast-lane procedures for requests that meet pre-set requirements** (e.g., a 'white-list'). This is managed by a Collaboration Administrator • **Temporary forking of a compute group** with participants who mutually agreed to changes and need to continue operations (audit trails are kept at systems level for traceability)	**Applies to 1, 4, 5:** • **Some** participating organisations **delegate voting** to a trusted third party (i.e. Collaboration Administrator), while **some require direct control** vote themselves • **Majority voting rule** (if 70% agree to a change request, Data Providers can choose to follow the majority decision) • **Fast-lane procedures for requests that comply to pre-set requirements** (e.g. a 'white-list'). This is managed by a Collaboration Administrator	Data Steward(s) and Collaboration Administrator(s) are involved in all A, B, and C
4 **Request to add a use case**				
5 **Request to access insights**				
2 **Request to leave**	**A, B, C preferences are not applicable for request 2 and 3, because:** • **(2)** Anyone can leave a compute group as covered by baseline mechanisms, preference for A, B and C does not generate new requirements • **(3)** Making a query is a standardised process and uses automated checks as long the query is within the scope of governance rules. If it's beyond the scope, then request for a new use case (type 4) shall be made instead.			
3 **Request to run a query**				

Source: CoE-DSC analysis

Governance framework should be executed by both existing and new roles in the network

Role	Description of responsibilities
Collaboration Authority	- A trusted party that represents the data collaboration, with whom each participant needs to sign an agreement - Acts as an enforcement body within the data collaboration
Collaboration/ Compute group Administrator(s)	- A party that manages and facilitates revisions of governance rules and DPIAs. - Facilitates a process for fast-lane admission of new participants. - Aids in detecting misconduct and handles incident management
Dashboard Administrator	- A party that manages access to the dashboard by MPC Beneficiaries
Data Steward	- A party that sets the governance rules for data sharing and data access on behalf of the Data Provider and votes to accept changes - Inputs governance rules in the governance and audit interface
Governance Network	- A group of trusted parties responsible for storing governance rules and audit trails, and managing governance and audit interface
MPC Provider (Infrastructure Administrator)	- During onboarding, activates computational nodes for the new participants - Ensures that participants' systems conform to the technical standards of the infrastructure, and ensures that nodes run properly
Identity Provider	- A trusted third party responsible for issuing digital identities to participants as a part of the onboarding procedure (e.g., trusted service providers under eIDAS or eHerkenning)

Source: CoE-DSC analysis

Table of Contents

1. Introduction to case study
2. Proposed governance solution
3. **Lessons learned and next steps**
4. Appendix

Next steps for CoE-DSC and the case study group

For CoE-DSC

- **Share outcomes for CoE-DSC participants.** Ensuring that proposed governance reaches other use cases this case study should be presented in CoE-DSC community
- **Embed results in CoE-DSC programme.** Re-use the governance framework for MPC data collaborations in future use cases and projects

For case study group

- **Develop a decision making tool for compute groups to assess their governance preferences in terms of control vs. pace:** In the future, the participants in the compute group would need the tool to aid a decision making process for assessing their preferences for desired governance and control mechanisms (see p. 28).
- **Share insights with potential participants:** To ensure that the implementation of the proposed governance is most useful to the data collaboration participants, this case study should be discussed by Linksight with their involved stakeholders.
- **Organise awareness workshops among the data collaboration participants:** In the long run, participants need to be aware of the changes brought by the data collaboration growth. Having workshops ensures that parties understand their needs for data control and changes in pace of processes given the scaled landscape.

Source: CoE-DSC analysis

Two important lessons learned from the case study on governance design for MPC data collaborations

Lessons learned

Devise DPIA's with a long-term purpose in mind

- In accordance with the GDPR, each participant of a data collaboration is required to devise the Data Protection Impact Assessment (DPIA)
- In a growing data collaboration, DPIAs that cover narrow data sharing contexts need to be revised very often
- Participants are encouraged to devise DPIAs that have a long-term relevance. For this, participants need to have alignment discussions on the scope, purpose, and risks for collaborating

Establish scalable contracting via an authority (*derdenwerking*) instead of bilateral contracts

- In the Netherlands *derdenwerking* is an example of scalable contracting, where an authority of the data collaboration makes contracts with individual participants binding the contracts chain (see pp. 29-30)
- This eliminates unscalable bilateral contracts that need to be (re-)made between all participants if somebody joins the network
- In addition, through contracts a collaboration authority functions as monitoring and enforcement body within the collaboration

Source: CoE-DSC analysis; see Data Sharing Canvas Section 7.3.2 on contracting

Appendix

- Summary of Dutch health care system - WMO, WLZ, ZVW
- Dashboard example for elderly care monitoring
- Explanation of Multi-Party Computation technology
- Full interaction model and roles
- Process design participants' request
- Dimensions to assess compute group preferences
- Examples of scalable contracting (derdenwerking)
- Examples of eHerkenning digital identity means

Dutch Healthcare system explained: WMO, WLZ and ZVW

	WMO: Wet Maatschappelijke Ondersteuning	**WLZ**: Wet Langdurige Zorg	**ZVW***: Zorgverzekeringswet (*scope: care at home)
Main goal of law	WMO ensures that people can continue to live at home for as long as possible	WLZ regulates heavy, intensive care for frail elderly people, people with disabilities and people with mental illnesses	ZVW mandates all Dutch citizens to have health insurance and mainly covers medical care costs
Organisation approving/financing care	Municipality	Approve: CIS Finance: Zorgkantoor	Approval (for 'wijkverpleging'): Thuiszorginstelling Finance: Health Insurer
Housing situation	At home ('NL: zelfstandig')	Includes right to move to care provider ('NL: zorginstelling')	At home ('NL: zelfstandig')
Time of care provided	Fixed periods	Lifetime	Fixed periods

Getting insights into combined performance of WMO, WLZ and ZVW is main focus of this use case. These insights can be used to holistically optimise healthcare policies

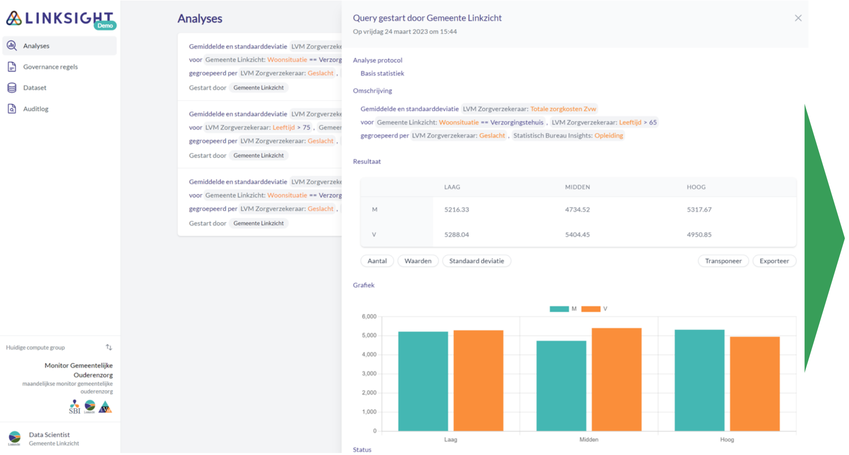
Source: Ministerie van Volksgezondheid, Welzijn en Sport

The insights for monitoring Dutch elderly care are created by Linksight as the result of the MPC-based data collaboration

Example visual of a statistical dashboard for monitoring Dutch elderly care

Simplified
For illustrative purposes only

Insights created in the Query and Result Interface . . .



Huidige compute group

Monitor Gemeentelijke Ouderenzorg
maandelijkse monitor gemeentelijke ouderenzorg

Data Scientist
Gemeente Linkzicht

Analyses

Gemiddelde en standaarddeviatie LVM Zorgverzeke
voor Gemeente Linkzicht: Woonsituatie == Verzorg
gegroepeerd per LVM Zorgverzekeraar: Geslacht ,
Gestart door Gemeente Linkzicht

Gemiddelde en standaarddeviatie LVM Zorgverzeke
voor LVM Zorgverzekeraar: Leeftijd > 75 , Gemeer
gegroepeerd per LVM Zorgverzekeraar: Geslacht ,
Gestart door Gemeente Linkzicht

Gemiddelde en standaarddeviatie LVM Zorgverzeke
voor Gemeente Linkzicht: Woonsituatie == Verzorg
gegroepeerd per LVM Zorgverzekeraar: Geslacht ,
Gestart door Gemeente Linkzicht

Query gestart door Gemeente Linkzicht
Op vrijdag 24 maart 2023 om 15:44

Analyse protocol
Basis statistiek

Omschrijving
Gemiddelde en standaarddeviatie LVM Zorgverzekeraar: Totale zorgkosten Zvw
voor Gemeente Linkzicht: Woonsituatie == Verzorgingstehuis , LVM Zorgverzekeraar: Leeftijd > 65
gegroepeerd per LVM Zorgverzekeraar: Geslacht , Statistisch Bureau Insights: Opleiding

Resultaat

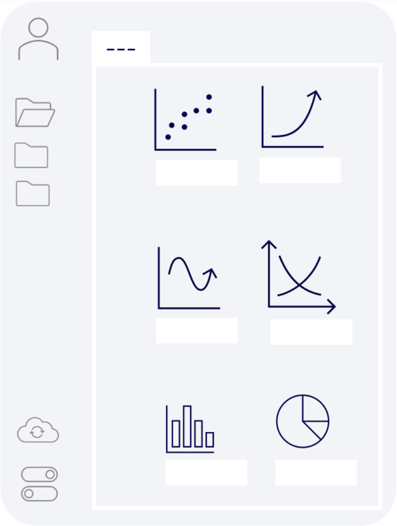
	LAAG	MIDDEN	HOOG
M	5216.33	4734.52	5317.67
V	5288.04	5404.45	4950.85

Aantal · Waarden · Standaard deviatie · Transponeer · Exporteer

Grafiek

Status

. . . serve as input for the Dashboard



Explanation

- The MPC analyses run in the compute group based on the Data Scientists' queries. As a result the anonymised statistical insights are generated
- Those insights are then translated into a dashboard with the help of a Dashboard Administrator. The dashboard is accessed by MPC Beneficiaries (e.g. health care providers, municipalities)

Non-exhaustive examples of the dashboard insights include:

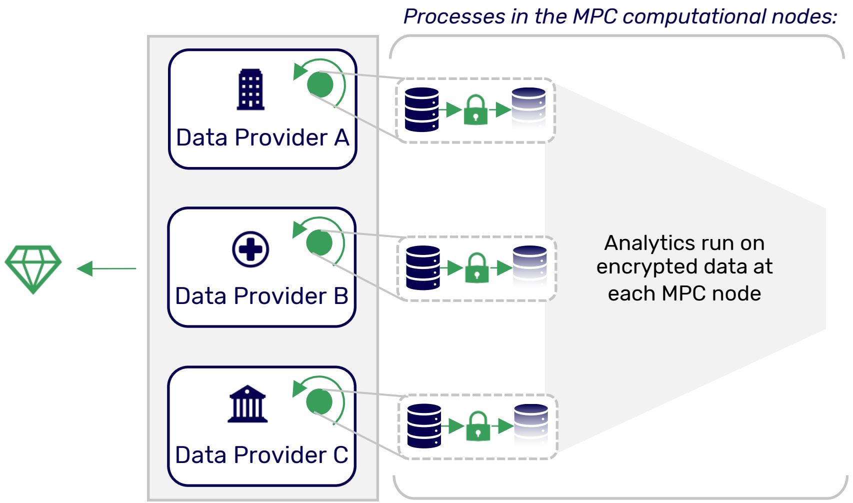
- Basic statistics on budgets and spendings per different patient groups
- Statistics on the performance of care provision across policies (e.g. differences of patient groups under WMO, WLZ, ZVW)
- Measured impact on patient groups due to a specific change in policy or some controlled intervention (e.g. effect from providing hip protection pads to elderly in the Delft region)

Source: Linksight

MPC strengthens trust and privacy in a data collaboration, how it is implemented is subject to participants' rules and agreements

Multi-Party Computation (MPC) in action

Simplified


Processes in the MPC computational nodes:
Data Provider A
Data Provider B
Data Provider C
Analytics run on encrypted data at each MPC node

Legend:

→ Process	● Computational node	Focal role	Raw data	Compute Group
Insights	Data encryption	Enabling component	Data with minimised PII	

Explanation of MPC for ensuring privacy and data control

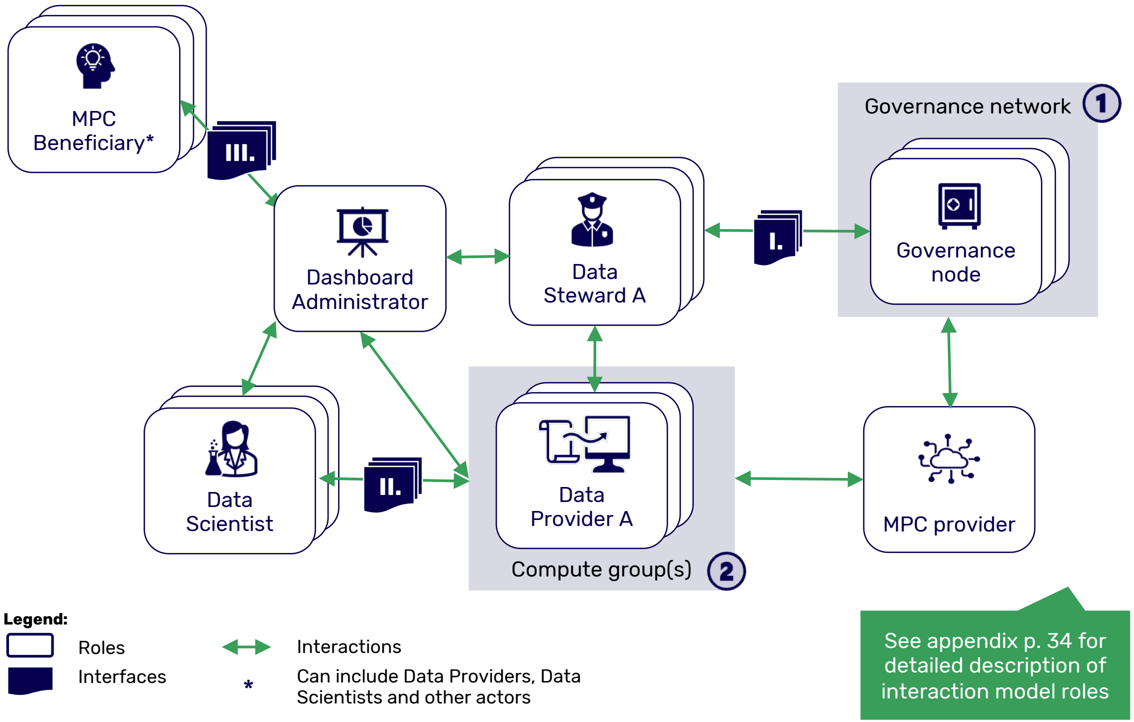
- Multi-Party Computation (MPC) is a cryptographic technique that enables multiple parties to perform computations on data in a way that insights are generated while keeping all input data private
- It is realised by setting up computational nodes at each Data Provider to secure and encrypt the data, then the decentralised analysis is run across the nodes
- Hence, MPC ensures that sensitive data is not shared and remains private. Only anonymised insights are made available as the result of the analysis
- Additional (governance) rules are needed to ensure compliance from all the participants on which analyses are allowed to be performed

Source: CoE-DSC analysis **Note:** Here the model shows what occurs on a compute group level

Current MPC data collaboration for monitoring Dutch elderly care consists of various roles and software interfaces

Interaction Model (simplified)

Note: participants can fulfil multiple roles



See appendix p. 34 for detailed description of interaction model roles

Explanation

Current data space consists of two parts:

1. One Governance Network - is a managing body consisting of fixed number of members (max 5), each with an active governance node
2. Many Compute Groups - flexible in their size and number, where each data provider has an active computational (MPC) node

Participants of the data collaboration intract with the following interfaces:

I. **Governance and audit interface** contains governance rules and audit trails

II. **Query and result interface** contains results to the requested analysis

III. **SaaS dashboard interface** contains aggregated anonymised insights

Source: CoE-DSC analysis

Data collaboration roles and their responsibilities

Main roles and responsibilities of participants in the network

Note: participants can fulfil multiple roles

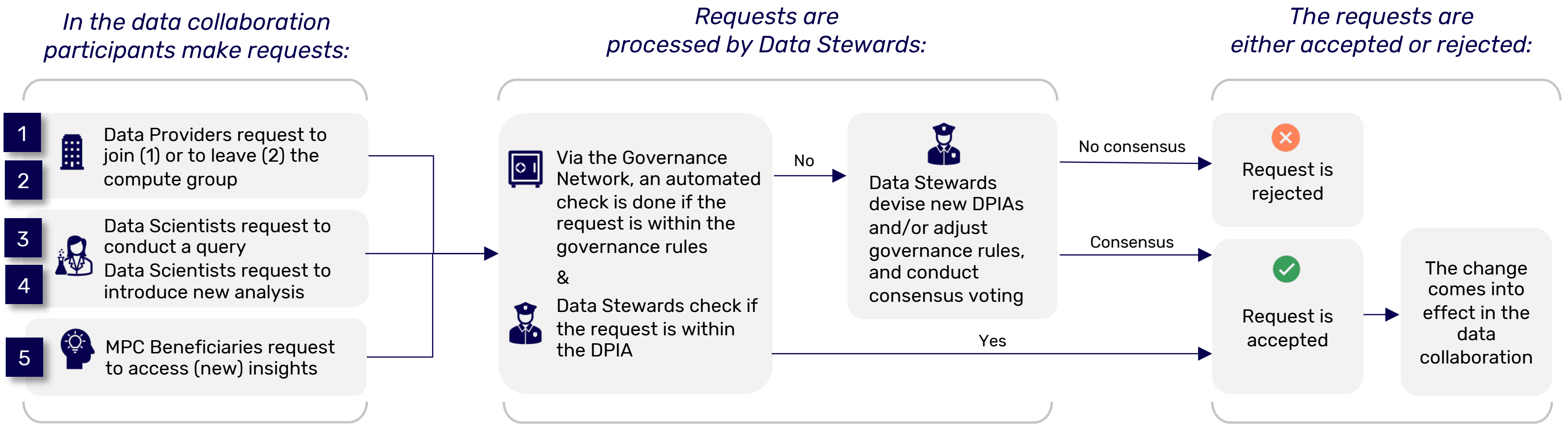
MPC Beneficiary	"Use the results"	• Interacts with the SaaS dashboard to access the insights from MPC computations
Data Provider	"Provide data for PET computation"	• Runs a computational node for MPC calculations • Note, the data remains at a source
Governance Network	"Manage rules"	• Manages the governance & audit interface • Upkeeps rules for the compute group
Compute group(s)	"Compute while complying to rules"	• Runs queries and computes the results • Manages the query and result interface and SaaS dashboard
Data Steward (1 per data provider)	"Create rules"	• Sets the rules for data sharing and data access • Inputs rules in the governance and audit interface
Data Scientist (>1 per compute group)	"Make queries"	• Creates data queries for making computations • Uses query and result interface
Dashboard Administrator	"Facilitate the dashboard"	• Combines results from the data scientists' queries • Reminds data providers (if needed) to provide their data in the right quality • Grants PETs Beneficiaries the access to the dashboard
MPC Provider	"Facilitate the network"	• Currently, Linksight supports the governance network and compute groups as a PET software provider

Source: CoE-DSC analysis

Processing of participants’ requests ensures that Data Providers remain in control over their data even when changes occur

Simplified

Processing map



In the data collaboration participants make requests:

1, 2 Data Providers request to join (1) or to leave (2) the compute group

3, 4 Data Scientists request to conduct a query
Data Scientists request to introduce new analysis

5 MPC Beneficiaries request to access (new) insights

Requests are processed by Data Stewards:

Via the Governance Network, an automated check is done if the request is within the governance rules

&

Data Stewards check if the request is within the DPIA

No → Data Stewards devise new DPIAs and/or adjust governance rules, and conduct consensus voting

The requests are either accepted or rejected:

No consensus → Request is rejected

Consensus → Request is accepted

Yes → Request is accepted

Request is accepted → The change comes into effect in the data collaboration

Legend:

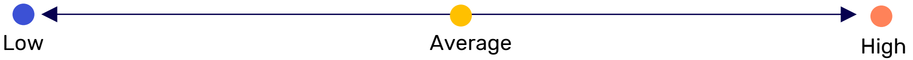
→	Process flow
▭	Process steps
🏢	Data Provider
💡	MPC Beneficiary
👮	Data Steward
🧑‍🔬	Data Scientist
▣	Governance Network

Compute groups can select scenarios by scoring their size, familiarity, sensitivity and homogeneity of the data

Indicative

Dimensions[1]	Description	Rationale
1 **Size of the compute group**	Number of the participating organisations in the analysis	Size of the compute group determines whether participants are likely to keep direct control on incoming change requests, or delegate control (e.g., the lower the group size, the easier to keep direct control)
2 **Degree of familiarity**	Degree to which compute group shares similar objectives and characteristics	Degree of familiarity determines whether participants are likely to keep direct control on incoming change requests, or delegate control (e.g., the higher the familiarity, the easier to delegate control)
3 **Sensitivity of data**	Degree to which data has private, confidential information (e.g. PII of patients is highly sensitive)	Degree of data sensitivity determines whether participants are likely to keep direct control on incoming change requests, or delegate control (e.g., the higher the data sensitivity, the participants are more inclined to opt for direct control)
4 **Homogeneity of data**	Degree to which data has same attributes during integration of datasets[2]	Degree of homogeneity determines whether participants are likely to keep direct control on incoming change requests, or delegate control (e.g., when data with new attributes are merged so data sets become heterogeneous, participants are likely to keep direct control)

Each dimension lies on a three-point scale:



Sources: 1. Open Data Institute: Federated learning an introduction (2023); 2. Mihaylov, I. et al. (2019). https://doi.org/10.1186/s13062-019-0249-6

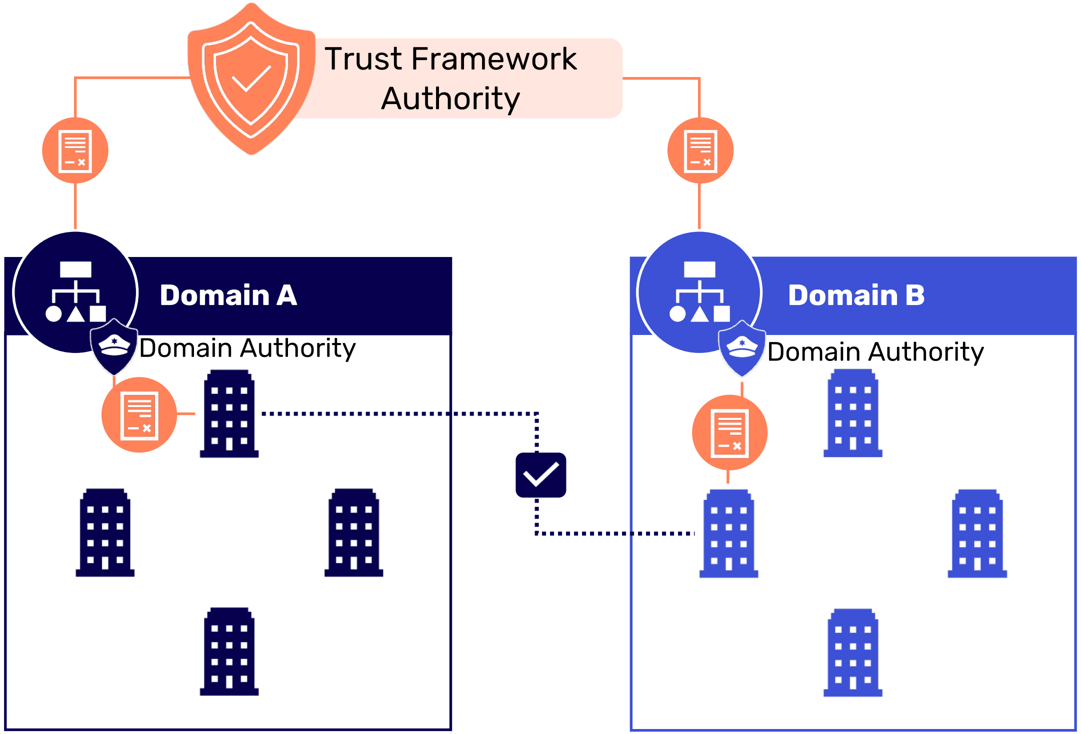
Dimension scores could be used to aid in deriving preferred governance strategy (and associated mechanisms)

Indicative

Dimensions[1]	How do dimensions relate to governance perspectives? (examples)		
1 Size of the compute group	HIGH 🟠	LOW 🔵	HIGH 🟠
2 Degree of familiarity	LOW 🔵	HIGH 🟠	AVERAGE 🟡
3 Sensitivity of data	HIGH 🟠	LOW 🔵	AVERAGE 🟡
4 Homogeneity of data	LOW 🔵	HIGH 🟠	AVERAGE 🟡
Mapping	**A Optimised direct control** **Explanation:** The scores indicate that Data Stewards should exercise direct control because there are a lot of participants with different interests, diverse data attributes and high data sensitivity	**B Fast pace of change** **Explanation:** The scores indicate a small group with similar objectives, low sensitivity of data and homogenous data attributes. Here delegation of control allows for fast processing of requests	**C Compromise** **Explanation:** The scores indicate a variety in the compute group when it comes to characteristics, sensitivity and homogeneity of data they expose. The compromise should ensure that parties contributing vital data for the analysis remain in the group

Sources: 1. Open Data Institute: Federated learning an introduction (2023); 2. Mihaylov, I. et al. (2019). https://doi.org/10.1186/s13062-019-0249-6

Derdenwerking is an example of scalable contracting, where Domain Authorities make contracts with individual participants

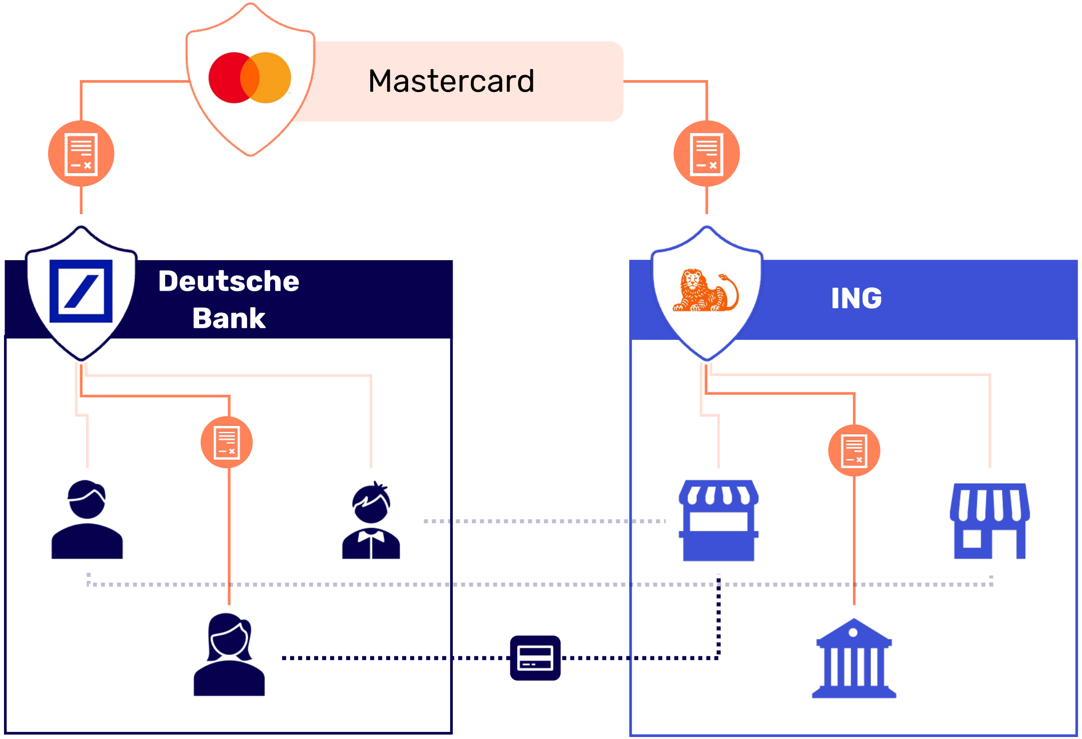


Explanation of derdenwerking

- Domain authorities are needed to aggregate the chain of contracts connecting all organisations in each domain
- In addition, the Domain authority functions as monitoring and enforcement body within the Domain (concerning the rules set within Domain schemes)
- The Domain authority could be executed by the domain scheme

Source: CoE-DSC analysis; see Data Sharing Canvas Section 7.3.2 on contracting

Example from Mastercard: A chain of contracts binds all actors within the ecosystem to enable payments between actors



Explanation

- Deutsche Bank has a contract with Mastercard to enable them to issue Mastercard branded credit cards
- Deutsche Bank issues Mastercard branded credit cards to their customers, who all have a contract with Deutsche Bank
- ING has a contract with Mastercard to enable them to facilitate accepting Mastercard payments at their merchants
- ING functions as an acquiring bank for their merchants, who all have a contract with ING
- Payments are facilitated between all Deutsche Bank customers and ING merchants

Source: CoE-DSC analysis; see Data Sharing Canvas Section 7.3.2 on contracting

eHerkenning

Example of eHerkenning: trusted identity providers issue eHerkenning login credentials for parties in an organisation

Steps 1-3 cover decisions for applying for eHerkenning

Step 1:

Decide on service providers you want to log into

More than 500 different service providers (i.e. governmental and private organizations) allow you to login with eHerkenning.

Check this list to see which service providers you intend to connect to with eHerkenning.

Step 2:

Decide who will represent the company using eHerkenning

Single eHerkenning means with the accompanying authorisations are linked to one individual only.

You need to issue eHerkenning means individually for each representative. But it is possible to apply for them in bulk.

Step 3:

Decide on the needed level of assurance out of 4 levels:

EH2 **EH2+** **EH3** **EH4**

The service provider determines the LoA required for their online services.

If you intend to use multiple services, better opt for the highest level.*

*Note: if needed the level can be upgraded later on

Steps 4-6 describe the actions for obtaining eHerkenning

Step 4:

Authorise each individual representative

The authorisation specifies for which service providers, and for which services, an individual can log into on behalf of their organisation.

Two people grant an authorisation:

Authorised signatory

Authorisation manager

Step 5:

Select a trusted supplier and apply for eHerkenning

6 official suppliers of eHerkenning can identify you and provide login means based on the assurance level you choose.

Z LOGIN

kpn

we·ID

QuoVadis

Digidentity

Reconi

Step 6:

Activate eHerkenning means and start using them

Once you have purchased eHerkenning login means*, you can activate and use them.

*Overview of login means per assurance level:

EH2	Username & password
EH2+	2FA
EH3	2FA
EH4	PKI certificate or 2FA

Source: CoE-DSC analysis based on https://eherkenning.nl/en/applying-eherkenning

CoE DSC