



# GEïntegreerde Regionale Data- infrastructuur Achterhoek (GERDA)

# Agenda



Aanleiding  
en achtergrond

Analyse  
en  
ontwerp

Voorbeeld  
decentrale  
data  
analyse

# Aanleiding en achtergrond

# Aanleiding

**Vereniging Digitale Zorg Achterhoek:** Geïntegreerde regionale data-infrastructuur voor beschikbaar stellen en uitwisselen van data vanuit het **primair** proces van zorg (zorgverlener <-> zorgverlener en zorgverlener <-> burger)

+

**8RHK Gezond (kavelmodel):** Geïntegreerde regionale data-infrastructuur voor de uitwisseling, verwerking, analyse en presentatie voor **secundair** gebruik van data over gezondheid en zorg.

=

**GERDA: GE**ïntegreerde **R**egionale **D**ata-infrastructuur **A**chterhoek voor uitwisseling, verwerking, analyse en presentatie van data over gezondheid en zorg, uit zowel primair proces als secundaire bronnen (bijv. registerdata zoals CBS)

# Gezondheid = samenwerken



## Regio-bestuurders

*Hoe staat het er voor in onze regio en binnen ons samenwerkingsverband. Waar liggen kansen en mogelijkheden?*



## Onderzoekers

*Kunnen we onderzoeken of de gezondheidsinterventies in een regio het gewenste effect hebben?*



## (Beleids)adviseurs

*Hoe kunnen we de beslissingen in onze regio zo goed mogelijk ondersteunen met data?*



## Professionals

*Hoe kan ik een goed beeld krijgen van de gezondheid in alle domeinen (niet alleen in de zorg) van mijn patiënten?*

**Maar wie zorgt voor de regionale inzichten?**

# Verschillende vragen...



Gezondheid & gedrag



Zorgkosten- en gebruik



Sociale omgeving



Aanbod en mogelijkheden



Leefomgeving



Fysieke omgeving

# Zomaar een paar vragen...

- Hoe kunnen we inzoomen op de problemen uit het **regiobeeld**
- **Hoeveel mensen** in mijn regio zijn de doelgroep voor interventie X?
- Hoe kom ik erachter in welke **wijken** die mensen zich bevinden?
- Hoe ik aan data (over gezondheid) **buiten mijn organisatie?**
- Hoe **monitoren** we het effect van gezondheidsinterventies?
- Als ik data heb van meerdere bronnen, hoe **koppel** ik die dan veilig?

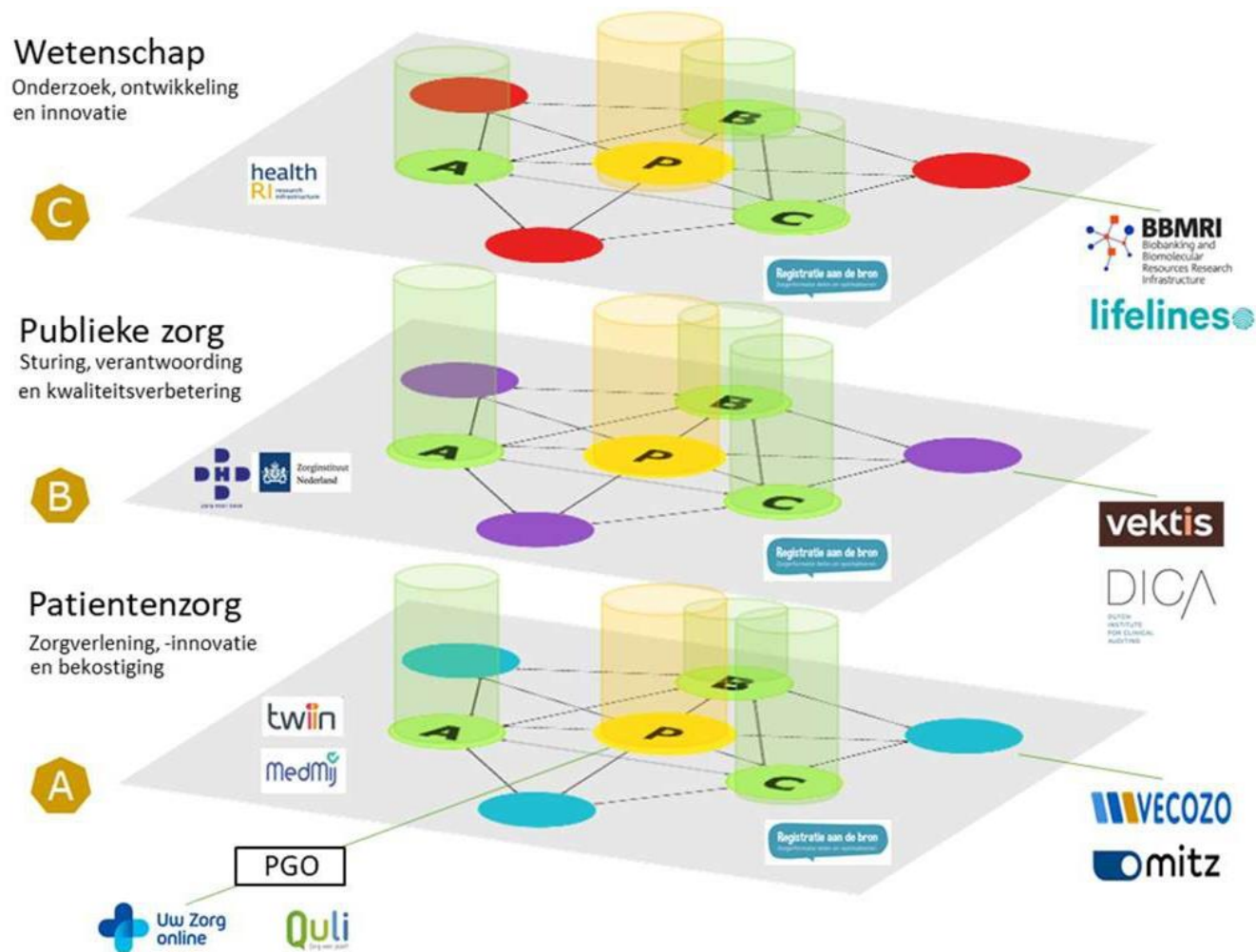
# Uitdagingen

- Meer en complexere vragen voor BI/data specialisten
- Infrastructuur over verschillende use cases beperkt herbruikbaar
- Verschillende typen datagebruikers (met eigen dataverzoeken en output-wensen)
- Samenbrengen en koppelen is niet makkelijk
- Wie van de betrokken organisaties neemt het voortouw?
- Organisaties betrokken bij meerdere regionale netwerken



# Analyse en ontwerp

# Vragen op verschillende niveaus



# Verschillende bronnen...

**GERDA**

## Zorg (primaire proces)

Ziekenhuis, Huisarts, Thuiszorg, GGZ,  
Wijkverpleging, Jeugdzorg, Verzekeraar, ...



## Sociaal domein

Gemeente, Wijkteams,  
Schuldhelpverlening, ...



## Sociale kenmerken

CBS, ...



## Leefomgeving

GGD, RIVM, Kadaster, ...



## Overige gegevens

PGO, eigen data, (sport) verenigingen,  
onderzoek, biobanken, ...

# Verschillende bronnen...

GERDA



Zorg (primair proces)



Leefomgeving



Sociaal domein



Overige gegevens



Sociale kenmerken

Geïntegreerde infrastructuur  
Decentraal  
Meerdere gebruikers  
Veilig koppelen

# Welke oplossingen hebben we onderzocht?

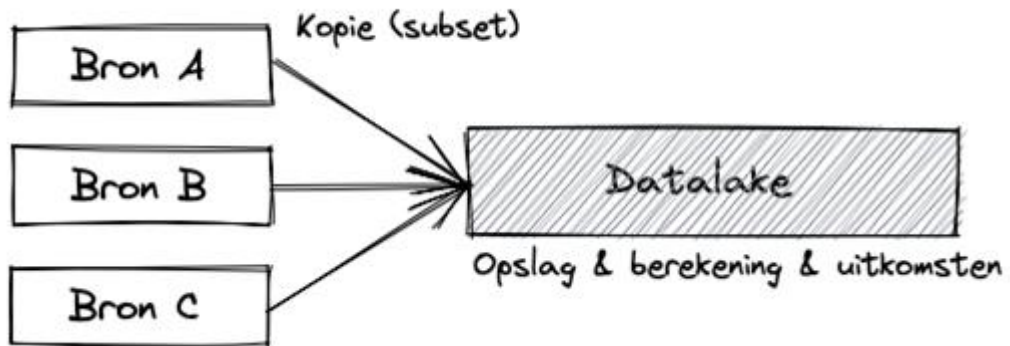
- Nationaal datalake
- Regionaal datalake
- Data virtualisatie
- Privacy preserving analytics
- FAIR data station

# Centraal of gefedereerd?

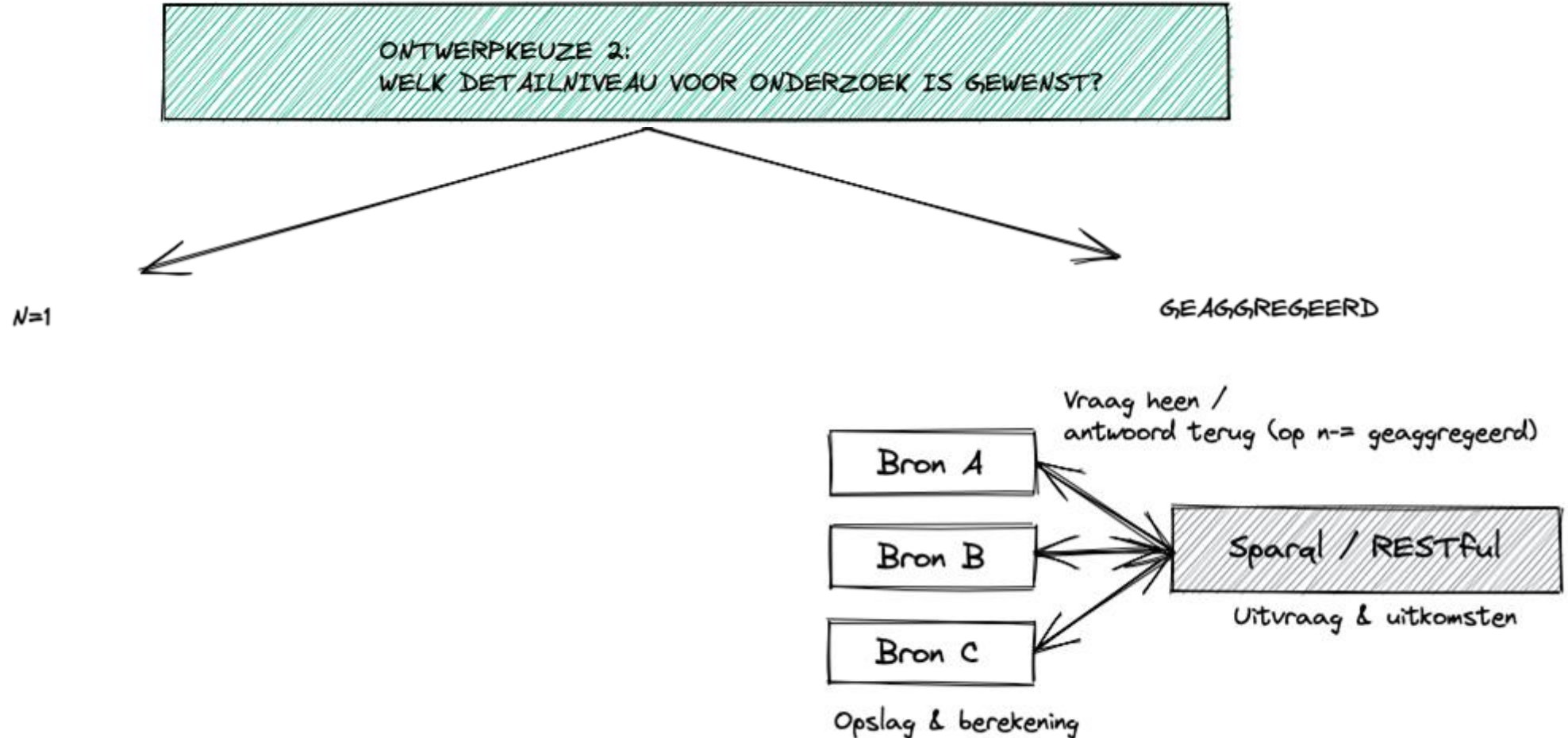
ONTWERPKEUZE 1:  
WELKE INRICHTING, TEN AANZIEN VAN DATA-OPSLAG, IS GEWENST?

CENTRAAL

GEFEDEREERD



# n = 1 of geaggregeerd?



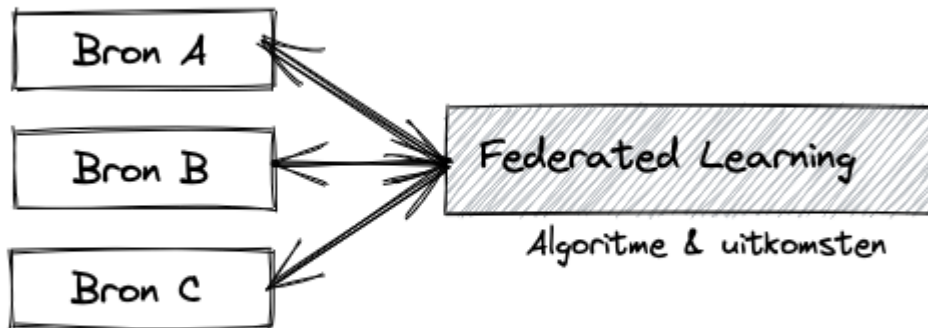
# Horizontaal of verticaal?

ONTWERPKEUZE 3:  
WELK SOORT ONDERZOEK IS GEWENST?

HORIZONTAAL  
(samenvoegen van mutually exclusive doelgroepen)

VERTICAAL  
(samenvoegen van overlappende datasets)

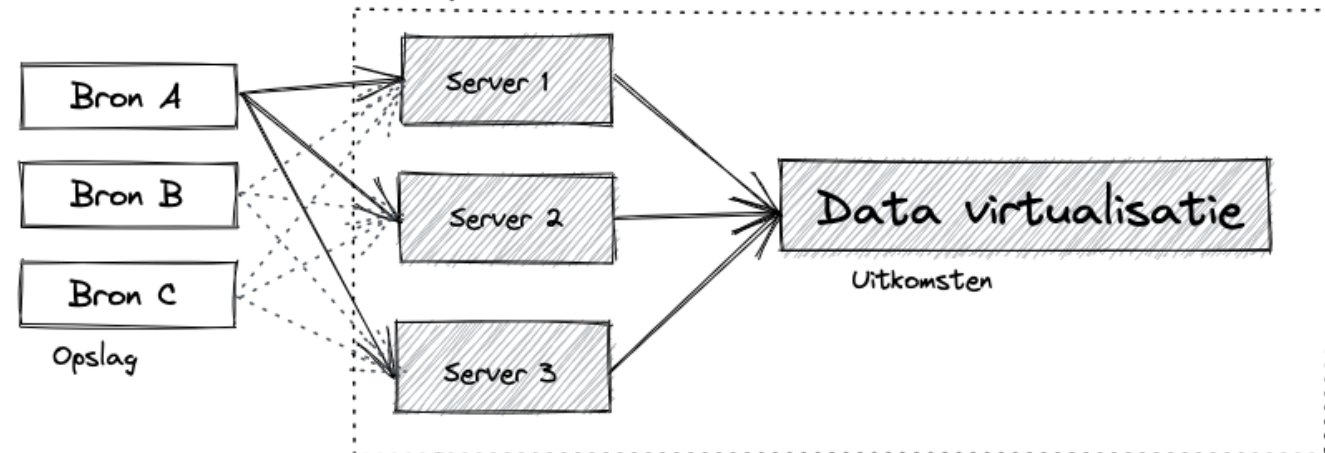
## Personal Health Train



Opslag & berekening

Interoperabiliteit in individuele databron

## Privacy Preserving Analytics



Berekening

Interoperabiliteit over databronnen heen



**Conclusie:**

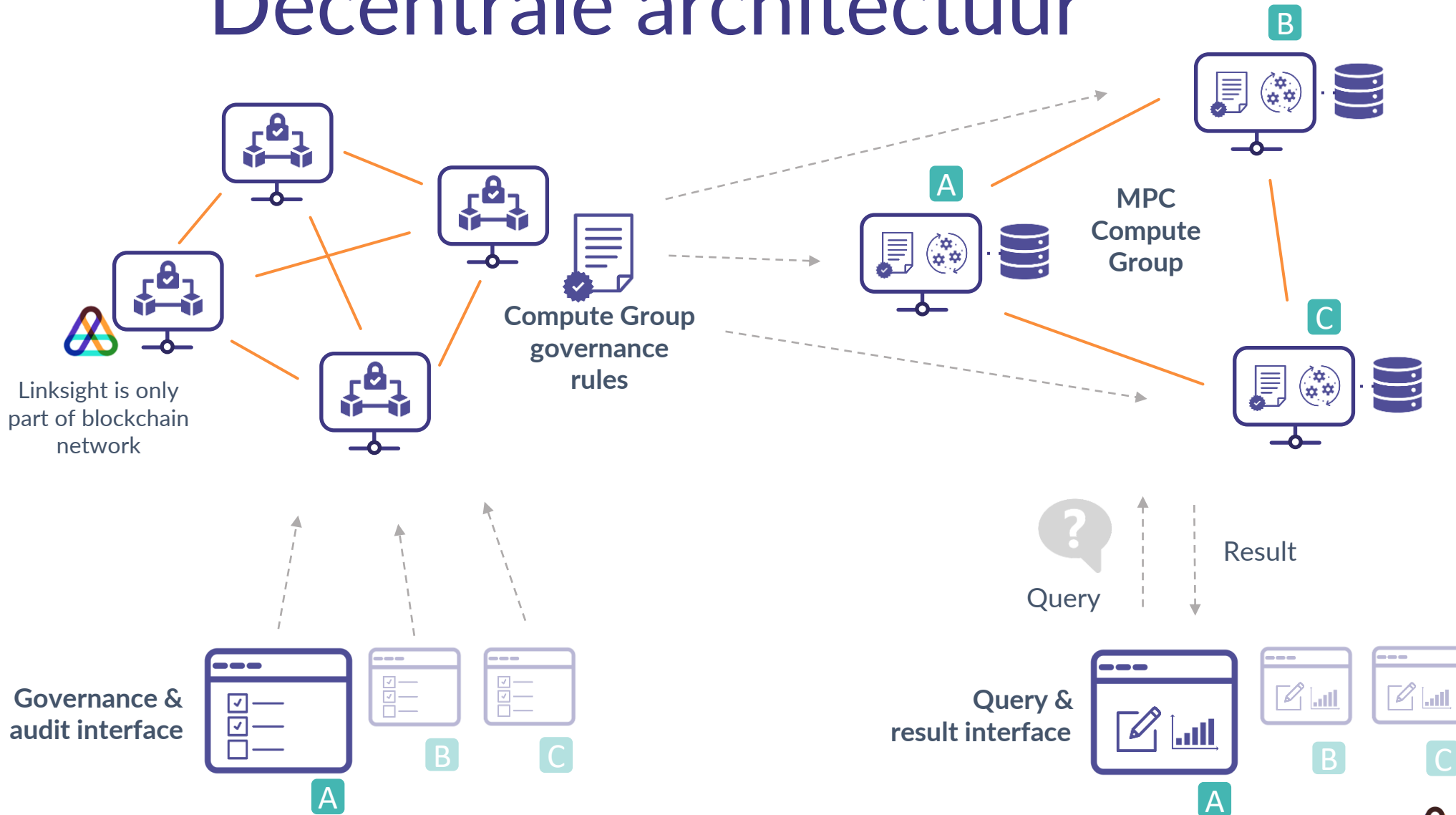
Gefedereerd

$N=1$

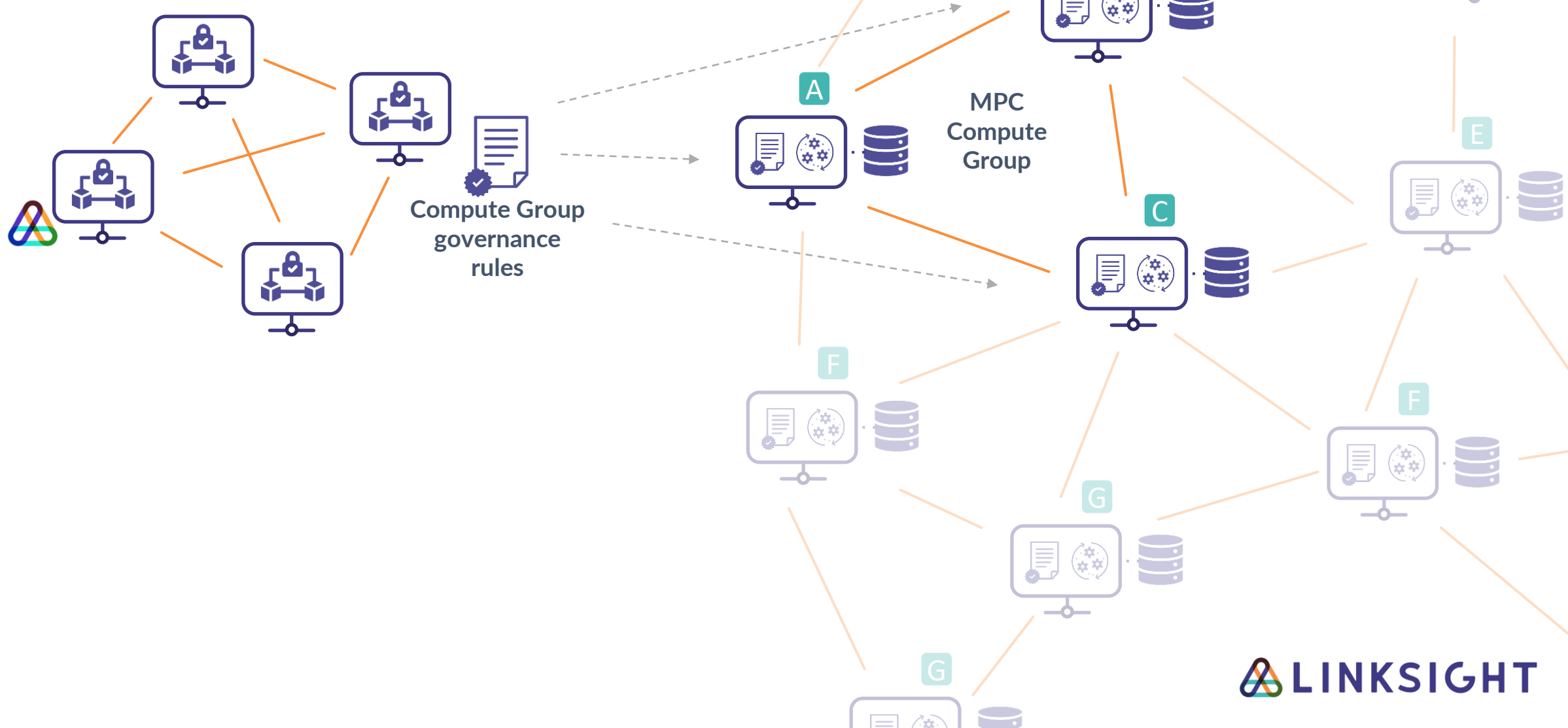
Verticaal

# Decentrale data-analyse

# Decentrale architectuur



# Decentrale architectuur



# Schaalbare en herhaalbare aanpak



# Governance interface: manage decentralized modules



**LINKSIGHT**

Dashboard  
Beheer compute groups  
Beheer infrastructuur  
Organisaties  
Instellingen

**Beheer compute nodes**

- Costs 2012-2019 (192.168.2.40) - Online (eHealth evaluation, SkinVision)
- Fraud signals (192.168.2.42) - Online (KvK fraud detection, Erasmus MC)
- + Voeg compute node toe

**Beheer query interfaces en dashboards**

- eHealth dashboard (192.168.2.52) - Online (eHealth evaluation)
- Statistics (192.168.2.50) - Online (eHealth evaluation, SkinVision)
- Fraud dashboard (192.168.2.54) - Online (Fraud detection)
- Fraud monitor (192.168.2.55) - Online (Fraud detection)
- + Voeg query interface toe

**Beheer blockchain node**

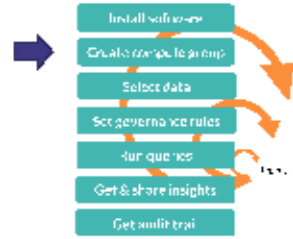
- Blockchain node (192.168.2.44) - Online

Tjerk Heijmans Visser  
CZ

The infrastructure management interface allows organizations to manage all their decentralized modules, regardless whether they run on-premise, in a private cloud or as SaaS at Linksight.

NB: Running a blockchain node is optional, not required.

# Create compute group



The screenshot shows the LINKSIGHT interface. On the left is a navigation menu with options: Dashboard, Beheer compute groups (selected), Beheer infrastructuur, Organisaties, and Instellingen. The main area is divided into two panels. The left panel, titled 'Actieve compute groups', lists several groups with their respective member organizations: 'Analyse Sananet 1' (CZ, Zuyderland, Carbrina, Sananet), 'Analyse Sananet 2' (CZ, Zuyderland, Carbrina, Sananet, Haga, Elkerliek, Cam), 'Analyse Luscii 1' (CZ, Carbrina, Luscii, Haga, Elkerliek), 'SkinVision kostenvergelijking' (CZ, SkinVision, Erasmus MC), 'eHealth app evaluatie (IBDCoach)' (CZ, Zuyderland, CBS), and 'Fraude detectie' (CZ, K&K, Achmea, Ibbobank). The right panel, titled 'Nieuwe compute group', is a form for creating a new group. It includes fields for 'Naam' (filled with 'Analyse eHealth'), 'Omschrijving', 'Analyse protocol' (set to 'Basic statistics'), 'Uit te nodigen organisaties' (CZ, Zuyderland, CBS), and a toggle for 'Gebruik een bestaande compute group als template' (checked) with a dropdown menu showing 'eHealth app evaluatie (IBDCoach)'. At the bottom right are 'Annuleren' and 'Afmaken' buttons.

In the governance interface organizations can create a compute group, select the algorithm and invite other organizations to participate.

Organisations can copy the settings of previous compute groups to make setting up a new one very easy.





# Setup governance rules (2)



The platform allows for various business rules to be set and technically enforced. Examples of such business rules that we used in the pilot are:

“Er mogen niet meer dan <X> queries per tijdseenheid <T> worden uitgevoerd.

Invulling:

<X>: 20

<T>: 24 uur”

“Het resultaat van een query mag niet opgeleverd worden als het aantal personen in de query dat voldoet aan de filter kleiner is dan <N>. De waarde van <N> is onafhankelijk van de attributen die in een query gebruikt worden.”

Invulling:

<N>: 30”

“De volgende organisaties mogen queries uitvoeren: <Y>  
De volgende organisaties mogen query-resultaten zien: <Z>”

Invulling:

<Y>: CBS, CZ, Zuyderland

<Z>: CBS, CZ, Zuyderland”

# Run queries



The screenshot shows the 'New query' dialog box in the LINKSIGHT application. The 'Select' section contains one entry: 'MEANSDEV' (Operation) applied to 'StarkinsuranceCorp: Fysio costs 2017' (Attribute). Below this is a '+ Add Select' button. The 'Where' section contains two conditions: 'ApertureStats: Education == 2' and 'GothamCityHospital: Used coach == Yes'. Below these is a '+ Add Where' button. A 'Close' button is located at the bottom right of the dialog box. The background shows a list of existing queries with their IDs and SQL snippets.

Query interface: building and submitting a query.

In this example:  
“what is the mean and standard deviation of physiotherapy cost for people with HBO/WO education who have used the IBD-coach in 2017?”

# Get the insights (2)



LINKSIGHT | Queries | Auditlog | Businessrules | Approval | Version

Query 13e20a successfully created.

## Queries

New query

Status | Sender | Search | 20 items per page | Sort: Newest first

13e20a	SELECT MEANSTDEV(StarkInsuranceCorp:Fysio costs 2017) FOR ApertureStats:Province == U, GothamCityHospital:Used coach == No	Aborted X
--------	--	-----------

**Query aborted**  
Reason: Fewer results than minimum result size

**Query details >**  
Query ID: 13e20a8d44833721ea8b7535abc2ccd57a6b24579d9cfe04cc080630aacbf469  
Sender: 0xa5bf461440753e0ee9cad9f32724f0aa1c9d40de  
MinResultSize: 30  
MinResultPercentage: 1  
MaxResultPercentage: 75  
PercentageConstraints: None  
StdevConstraints: None  
Groupkey: None  
DestinationNodes: None  
Reviewer: ApertureStats

When a query does not comply with the governance rules, a fault message is displayed here, and no result is calculated.

# Audit trail



**LINKSIGHT** < Queries Auditlog Businessrules Approval > Version

## Auditlog

Timestamp	Kind	
02/12/2021, 21:06:51	Query Status	13e20a
02/12/2021, 21:06:43	Query Approved	13e20a

**Sender:** 0xa5bf461440753e0ee9cad9f32724f0aa1c9d40de  
**Query:**  
ID: 13e20a8d44833721ea8b7535abc2ccd57a6b24579d9cfe04cc080630aacbf469  
aggregates:  
[ "MEANSTDEV(StarkInsuranceCorp:Fysio costs 2017)" ]  
filters:  
[ "ApertureStats:Province == U", "GothamCityHospital:Used coach == No" ]  
**Metadata:**  
MinResultSize: 30  
MinResultPercentage: 1  
MaxResultPercentage: 75  
PercentageConstraints:  
[]  
StdevConstraints:  
[]  
Groupkey:  
DestinationNodes:  
[]  
Reviewer: ApertureStats

All parties in a compute group can check the audit log and see what actions have been performed (queries, governance rules proposals).

# Q&A

*Nog vragen, interesse in het whitepaper of ervaringen?*

Maarten den Braber - [maarten@populationhealthdata.nl](mailto:maarten@populationhealthdata.nl)

Ilona Oude Nijhuis - [ilonaoudenijhuis@room-to.nl](mailto:ilonaoudenijhuis@room-to.nl)

# Bijlagen

# Plan van Aanpak

ONTWERP

## DATA ENGINEERING

Vaststellen voor alle bekende regionaal geduide use cases (w.o. acute zorg kwetsbare ouderen) welke inrichting er benodigd is om aan de huidige en de toekomstige onderzoeksvragen te voldoen.



## DATA SCIENCE & ANALYTICS

Onderzoeken welke infrastructuur nodig is om gekoppelde data te verwerken tot inzichten. Welke tools en outputs zijn hiervoor nodig? En welke type?



## DATA GOVERNANCE

Het doel is het definiëren van de kaders waarbinnen we de regionale data-infrastructuur op een duurzame wijze inrichten.

## ROAD MAP

Doel is om op basis van de uitkomsten van de werkpakketten 1-4 te komen tot een plan van aanpak van wat we moeten doen en met wie om de GERDA te realiseren

USE CASE

## PROTOTYPE

Het doel is om echt aan de slag te gaan met het daadwerkelijk koppelen van data in een beveiligde omgeving. We beantwoorden de vraag: Hoe gaat dit en waar lopen we tegen aan? We proberen een proof of concept op te leveren van een dashboard acute zorg. De inzichten uit dit traject verwerken we in bovenstaande ontwerpen.

REALISATIE