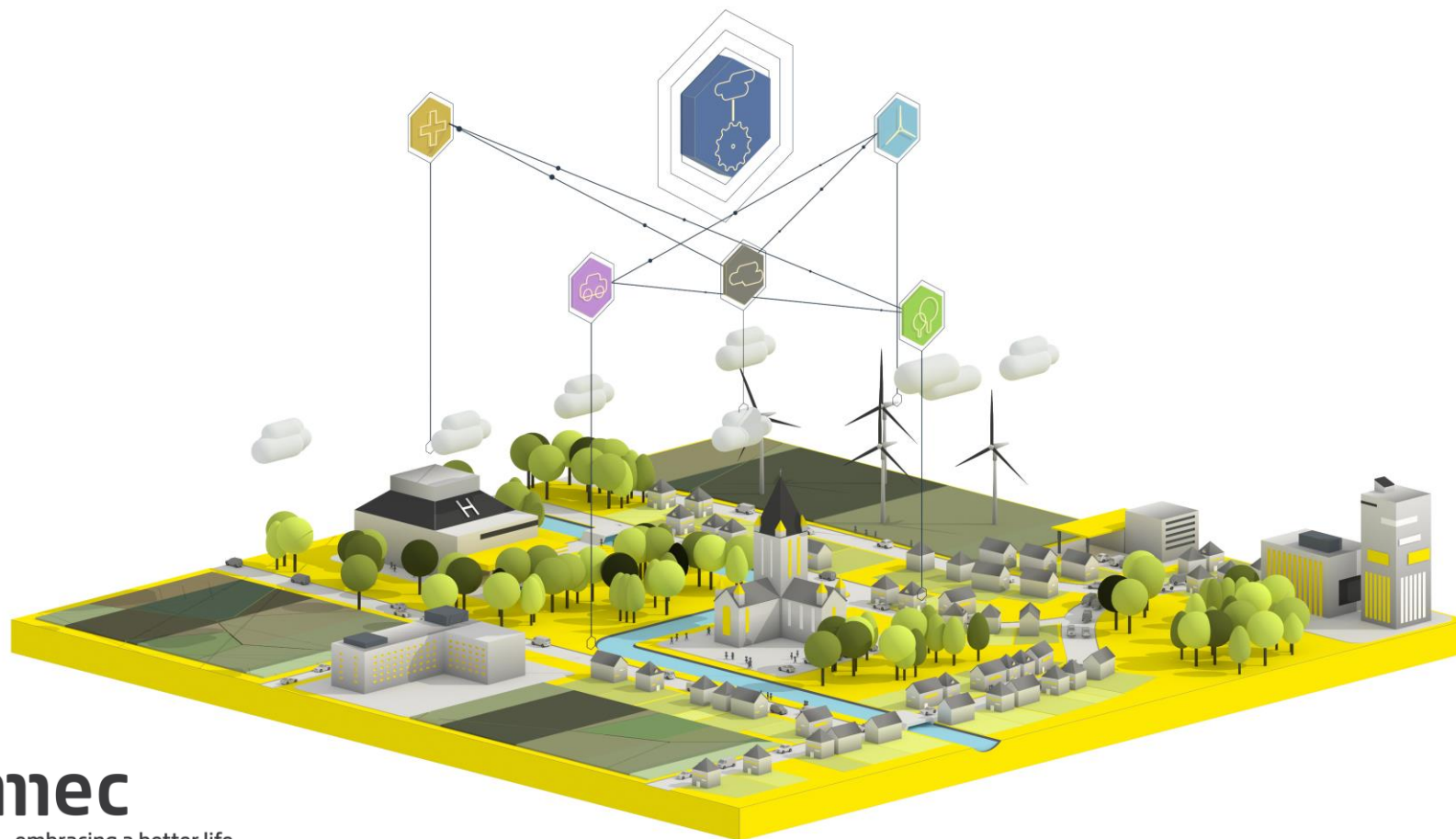


VLOCA

VLAAMSE OPEN CITY ARCHITECTUUR

Data Broker



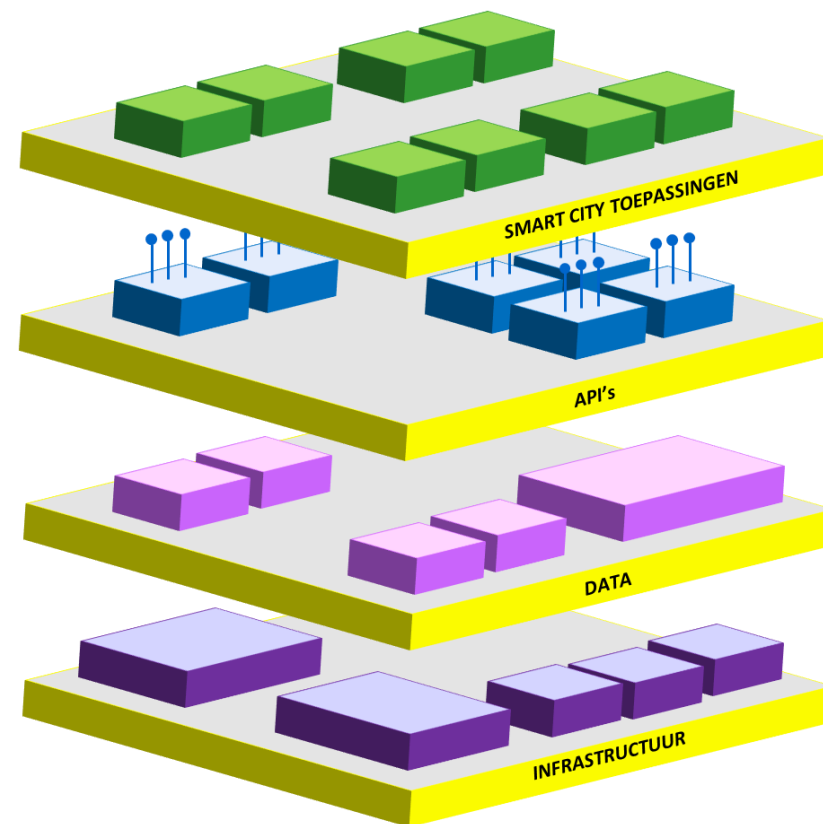
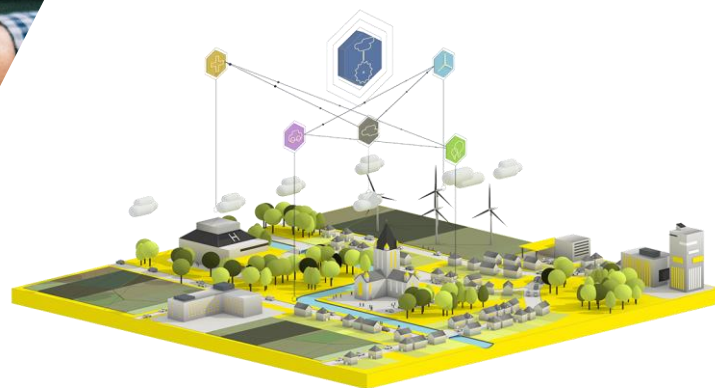
AGENDA

1. Inleiding & reflectie vorige workshop
2. Terminologie
3. Wat is een data broker (component) ?
4. Data Publishing & Data Broker
5. Open Discussie
 1. AS IS
 2. TO BE

Inleiding en reflectie workshop 1

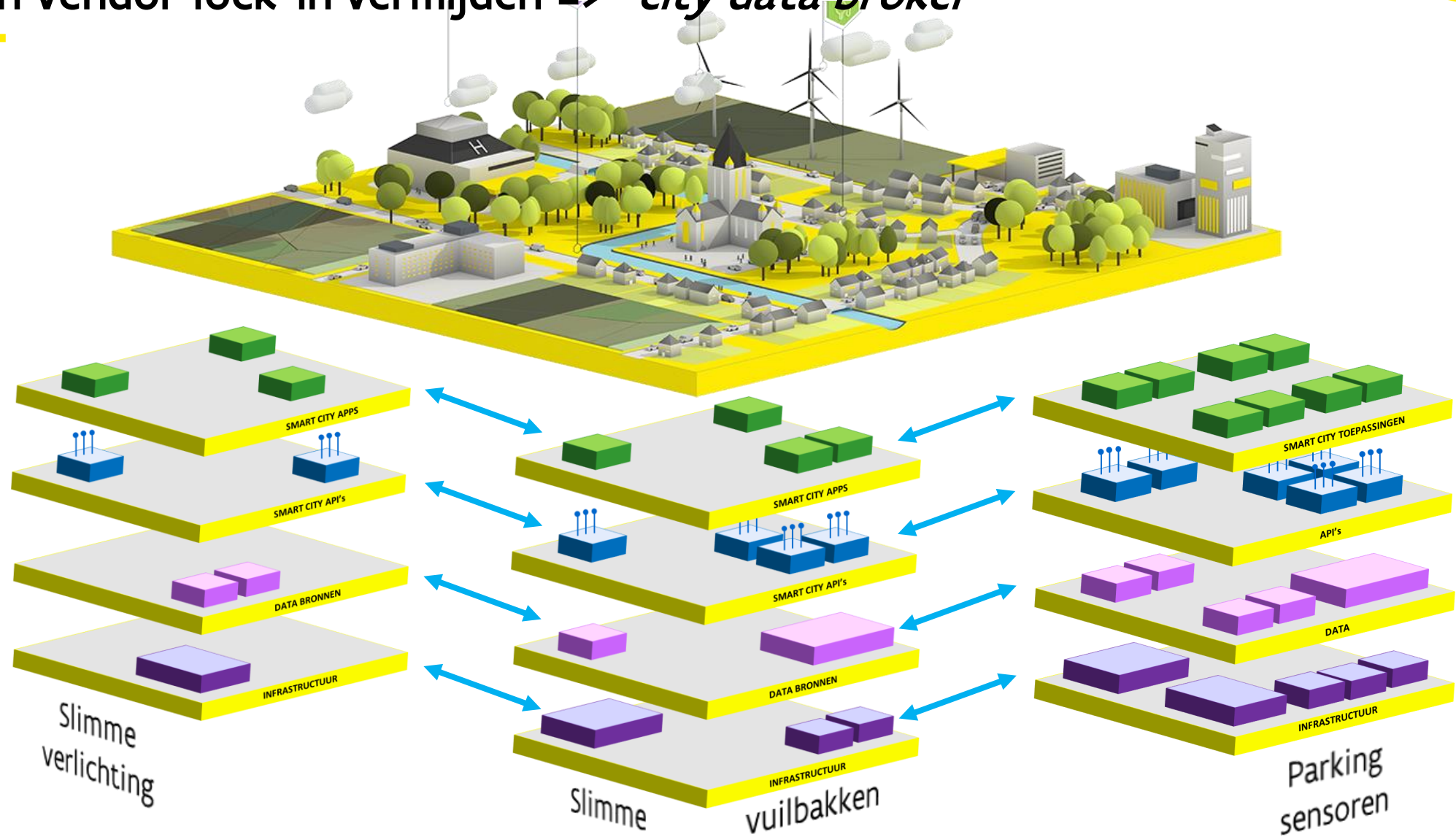
Selecteer Smart City bouwstenen

die voldoen aan VLOCA standaarden



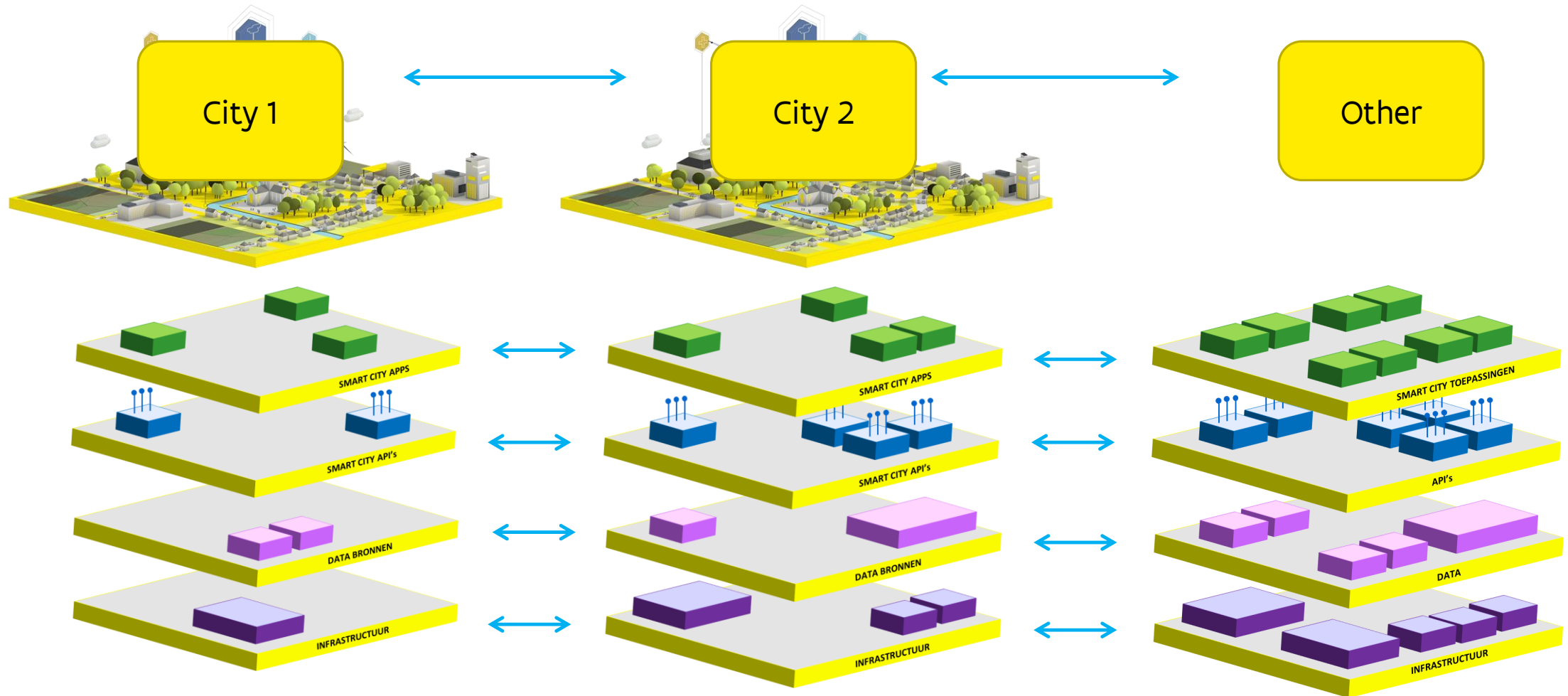
Geconnecteerde smart city oplossingen

silo's en vendor lock-in vermijden => “*city data broker*”

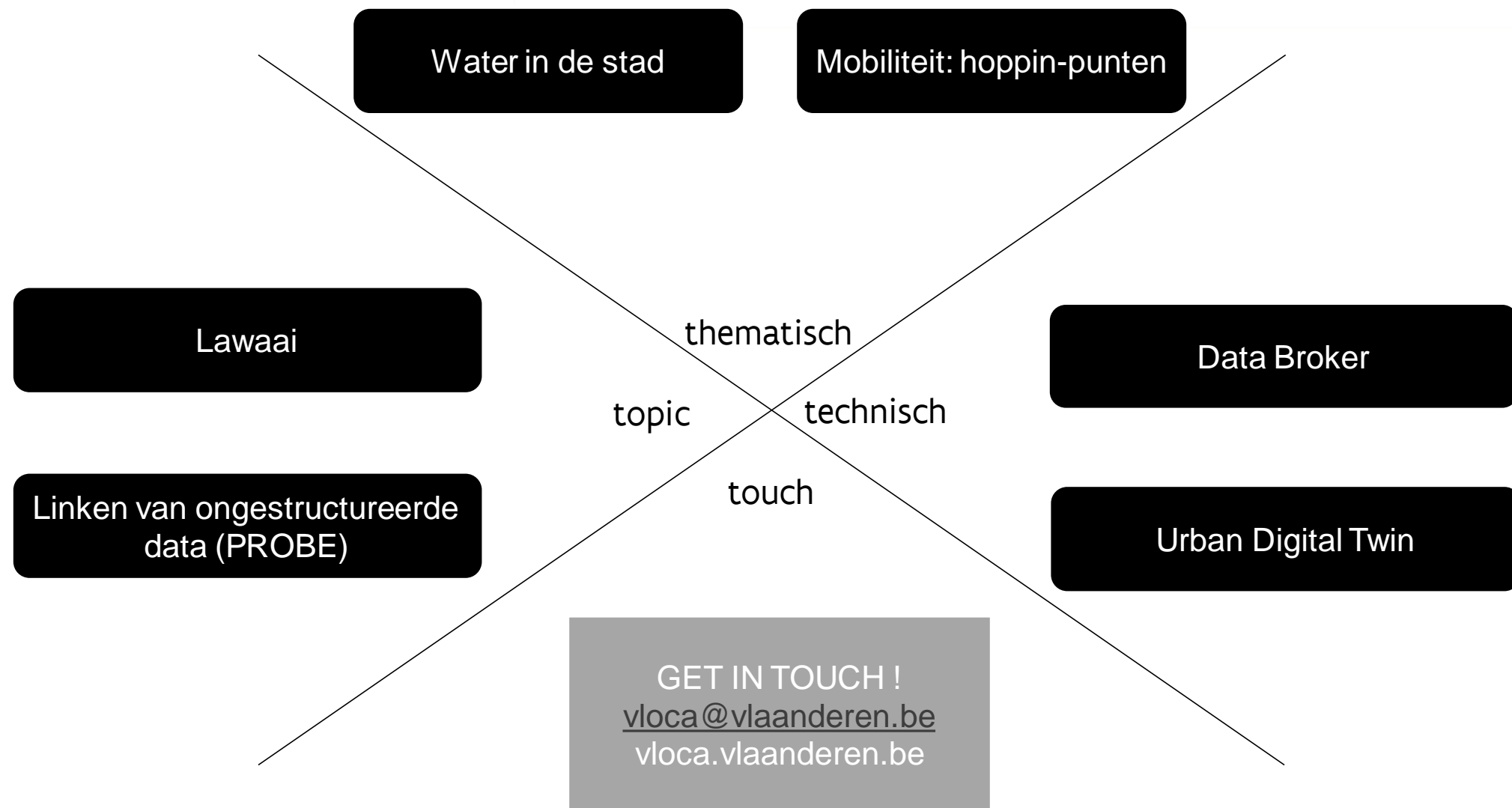


Bouw je eigen smart city

op maat van je lokaal bestuur => “*regional data broker*”



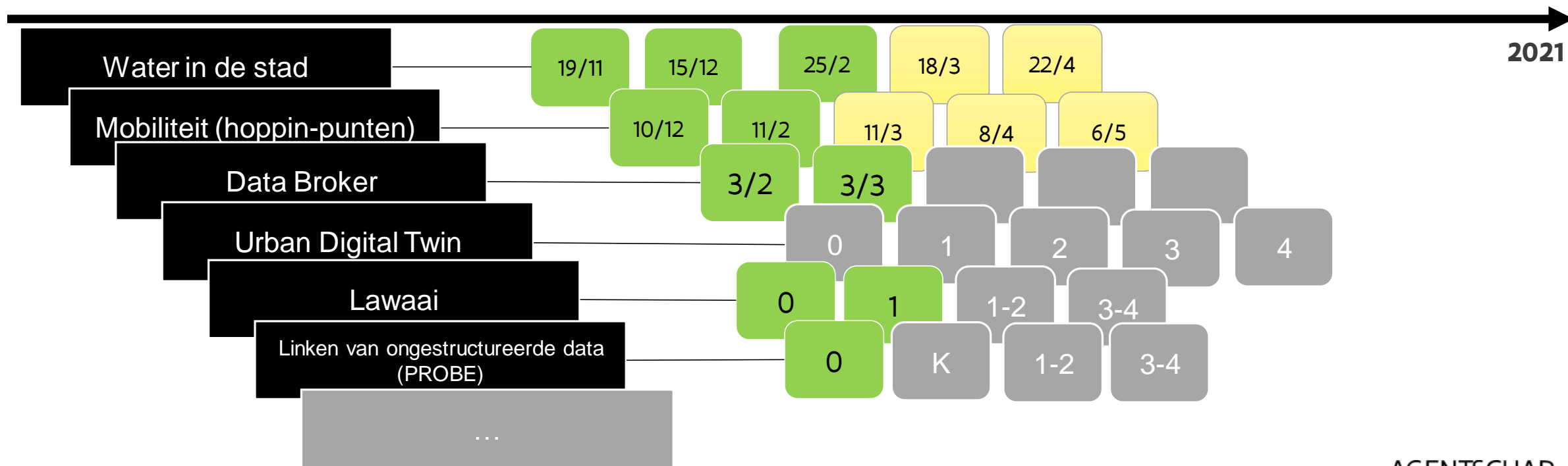
Actieve VLOCA trajecten (03/2021)



Lopende VLOCA trajecten (02/2021)

Input kennishub & co-creatie

Consolidatie in standaarden en draaiboeken



0: voortraject
K: kennismakingssessie
1-4: co-creatie sessies

AGENTSCHAP
BINNENLANDS
BESTUUR

Reflectie vorige workshop

Welke rol speelt een databroker?

- In Gent Data Broker gaat het over een **City (Of Things) Data Broker**, met aandacht voor het economische, technische, legale en functionele plaatje. Daarbij zijn specifieke keuzes gemaakt voor bijvoorbeeld het gebruik van Linked Data.
- Aanbesteding Brugge gaat over een “**Smart City Data Platform**”
- In Antwerpen heeft ACPAAS ook het doel om te evolueren naar een **generiek data platform**.

Reflectie vorige workshop

Aangehaalde uitdagingen / vragen :

- Software/data engineering en onderhoud :
 - Open source is goed, maar vraagt ook onderhoud.
 - Evolutie van standaarden : hoe managen ?
 - Data governance & architecture ?
- Waar kan Vlaanderen de steden ontzorgen ? Kan er sprake zijn van “**totale ontzorging**” ?
 - Meer genuanceerd : welke elementen worden beter op regionaal niveau opgelost en welke lokaal ?
- Hoe kunnen de steden aligneren met de regionale ontzorging ?
 - En bijvoorbeeld zich groeperen om de **economies of scale** te krijgen ?
 - Conflicteert een initiatief van B/R/L met het relance plan ? (sensor platform bijvoorbeeld)

Reflectie vorige workshop

Reeds enkele richtinggevende **antwoorden**:

- Relance plan gaat GEEN monoliet platform in de markt zetten
 - Focus op
 - standaarden
 - open source componenten en tools
 - Drempelverlagend voor publicatie van open linked data
- Het DNB kan zeker een rol spelen in realisatie van het relance plan
 - Maar is nog in volle definitie
- Data Governance zou een GEDEELD probleem moeten zijn (regionale, lokale overheden ism industrie)
- Ontzorging van steden op de “goeie” plaatsen is de uitdaging

Koppeling naar deze workshop

Wat gaan we in deze workshop doen ?

- Proberen verder te definiëren waar ontzorging belangrijk is, adhv een meer architectural plaatje van waar een data broker zich situeert binnen een gefedereerd data platform.
- Aantonen dat stedelijke platform ambities (vb. Brugge aanbesteding) en realisaties met roadmap (vb. ACPAAS) zich perfect kunnen aligneren met die regionale ontzorging (dat is de key doelstelling).
- We willen ook aantonen dat smart city data management een holistische benadering vraagt :
 - Wisselwerking op gemeentelijk/stedelijk niveau
 - Data behoefte op bovenlokaal niveau



Terminologie:

**nood aan een lingua franca voor het data
ecosysteem**

De eigenschappen en karakteristieken van (big) data

- De 3 V's, 5 V's, 10 V's, 14 V's, 17 V's, ... van (big) data
- Eenduidige definities zijn toch niet evident, en dit blijft evolueren omdat data steeds complexer wordt om te ontginnen.
- Maar die zijn wel nodig om te kunnen nadenken over data beheer en IT oplossingen daarvoor (VLOCA doelstelling)
- Dus, we gaan die alvast moeten definiëren in VLOCA
- <https://vloca-kennishub.vlaanderen.be/vloca-kennishub/v>

VLOCA:Data Terms & Concepts

Lijst met data eigenschappen (niet exhaustief) [bewerken]

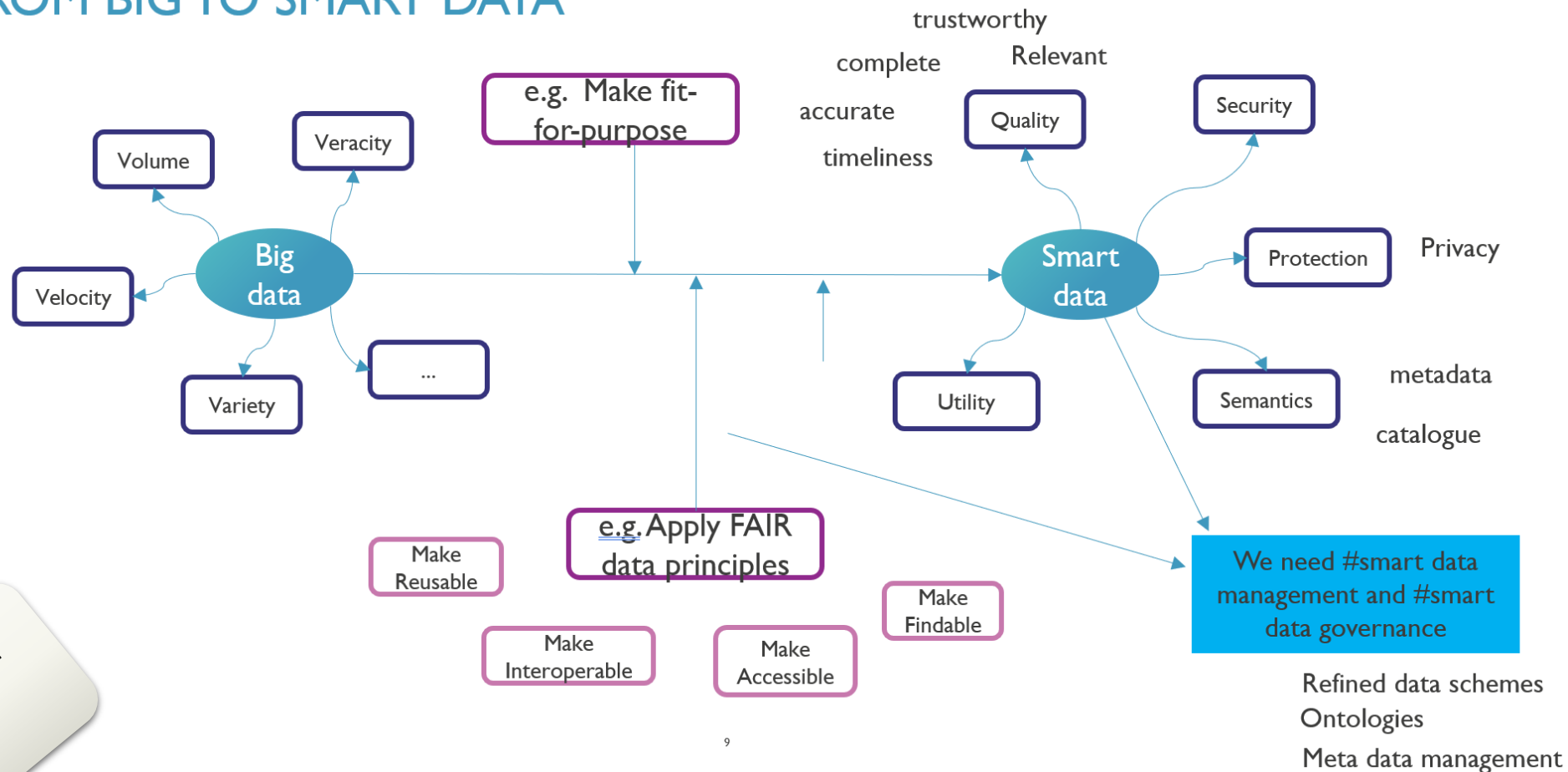
- data **velocity**
- data **quality**
- data **privacy**
- data **security**
- data **ownership/openness**
- data **variety**
- data **volume**
- data **veracity**

Open voor co-
creatie

**BINNENLANDS
BESTUUR**

Wat is “smart data”?

FROM BIG TO SMART DATA



Open voor co-
creatie

AGENTSCHAP
BINNENLANDS
BESTUUR

Wat houdt het beheer van Smart Data in ?

Lijst met data management capabilities (niet exhaustief) [\[bewerken\]](#)

- Versioning
- Querying
- Verifying
- Aggregation
- Archiving
- Merging
- Joining
- Storing
- Publishing
- Brokering
- Registering
- Authentication
- Licensing
- Discovery
- Annotation
- Modeling
- Governance
- Architecture



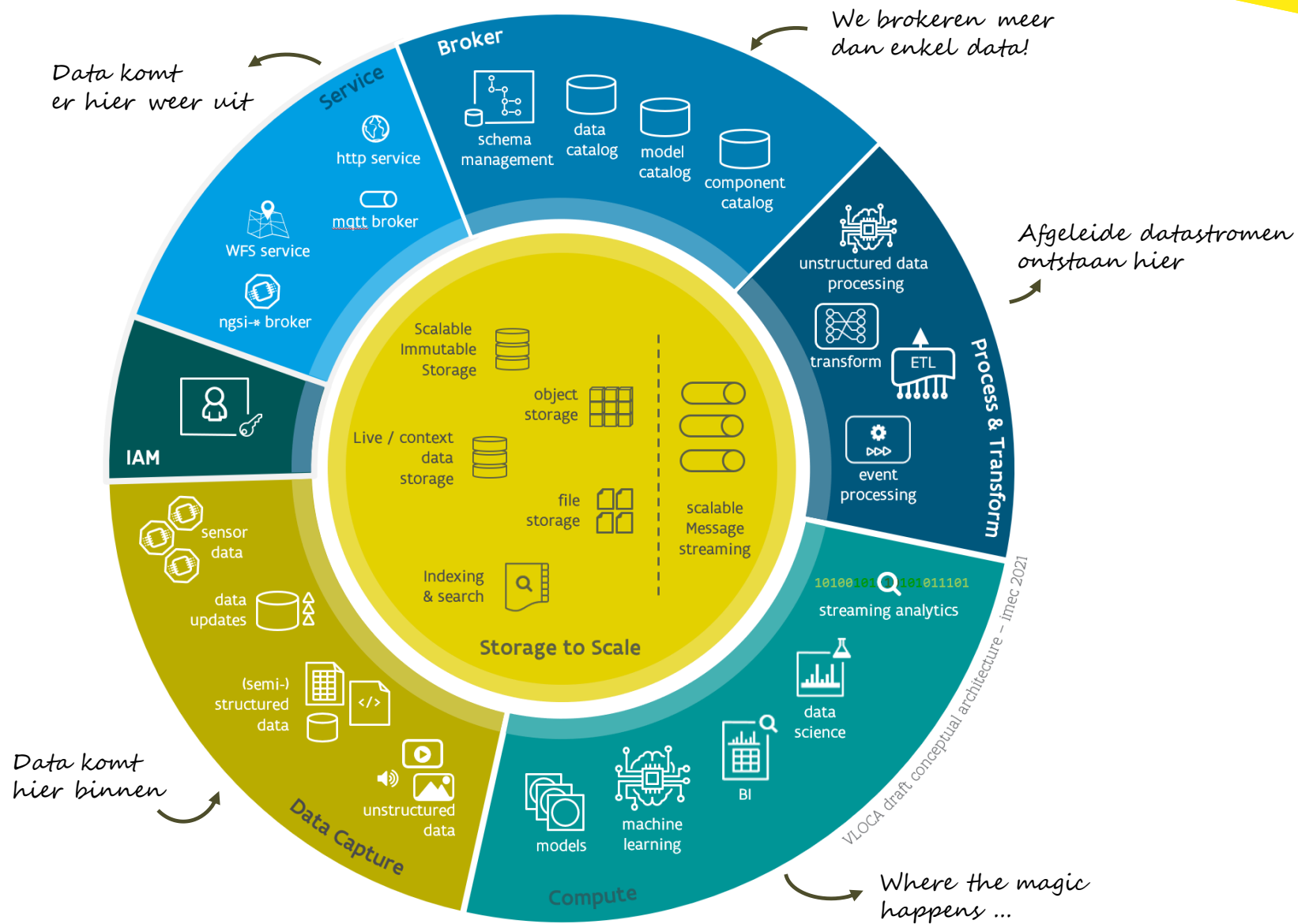
Wat is een data broker ?

Definitie op de VLOCA pagina

Data broker - user's Wiki! (vlaanderen.be)

- Een vlag die vele ladingen dekt:
 - Enkel het broker gedeelte van een dataplatform
 - ... of het hele platform
 - ... of iets ertussen?
- Een kadering in het smart data landschap





Smart Data Management

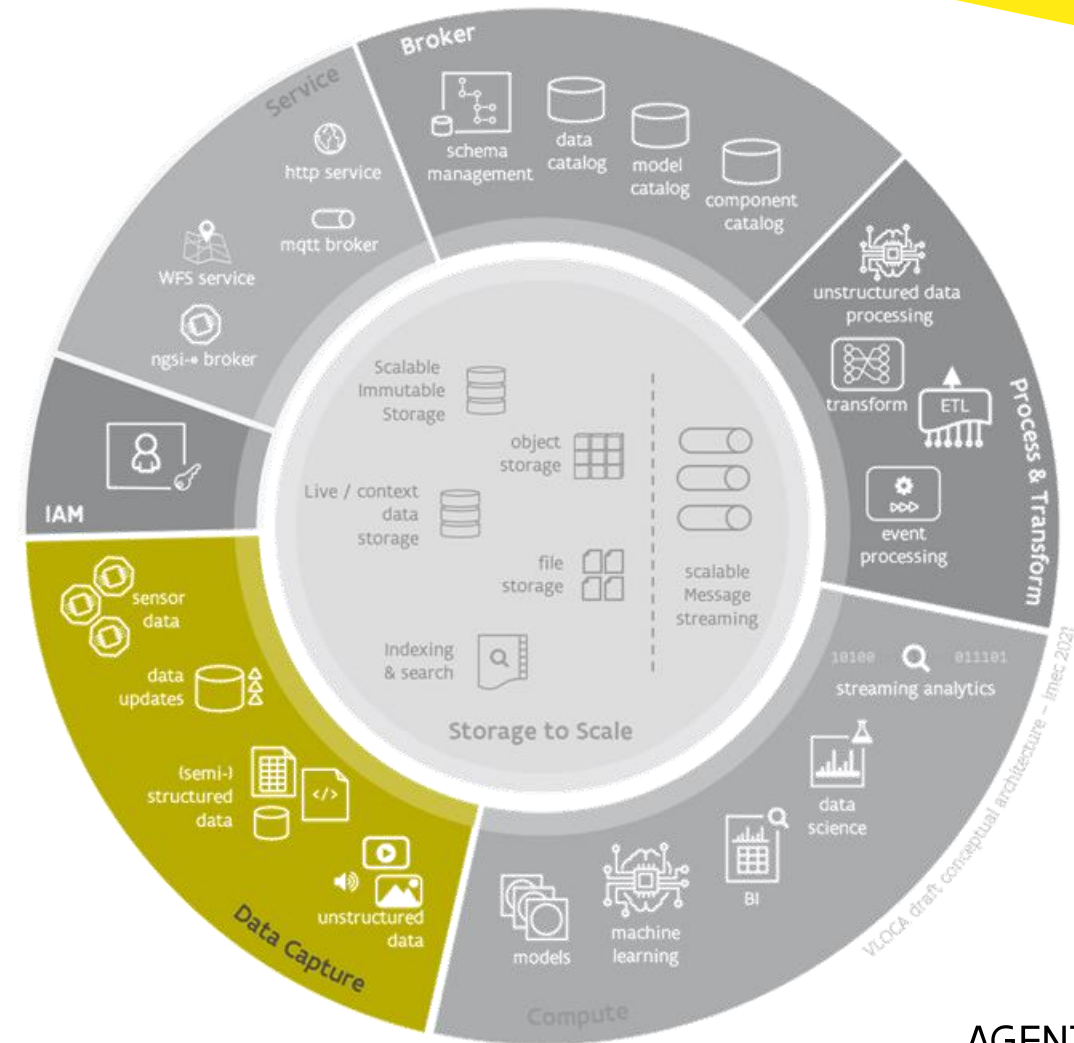
Wat ontbreekt hier nog?

- GIS data
- Low latency data
- Personal data management

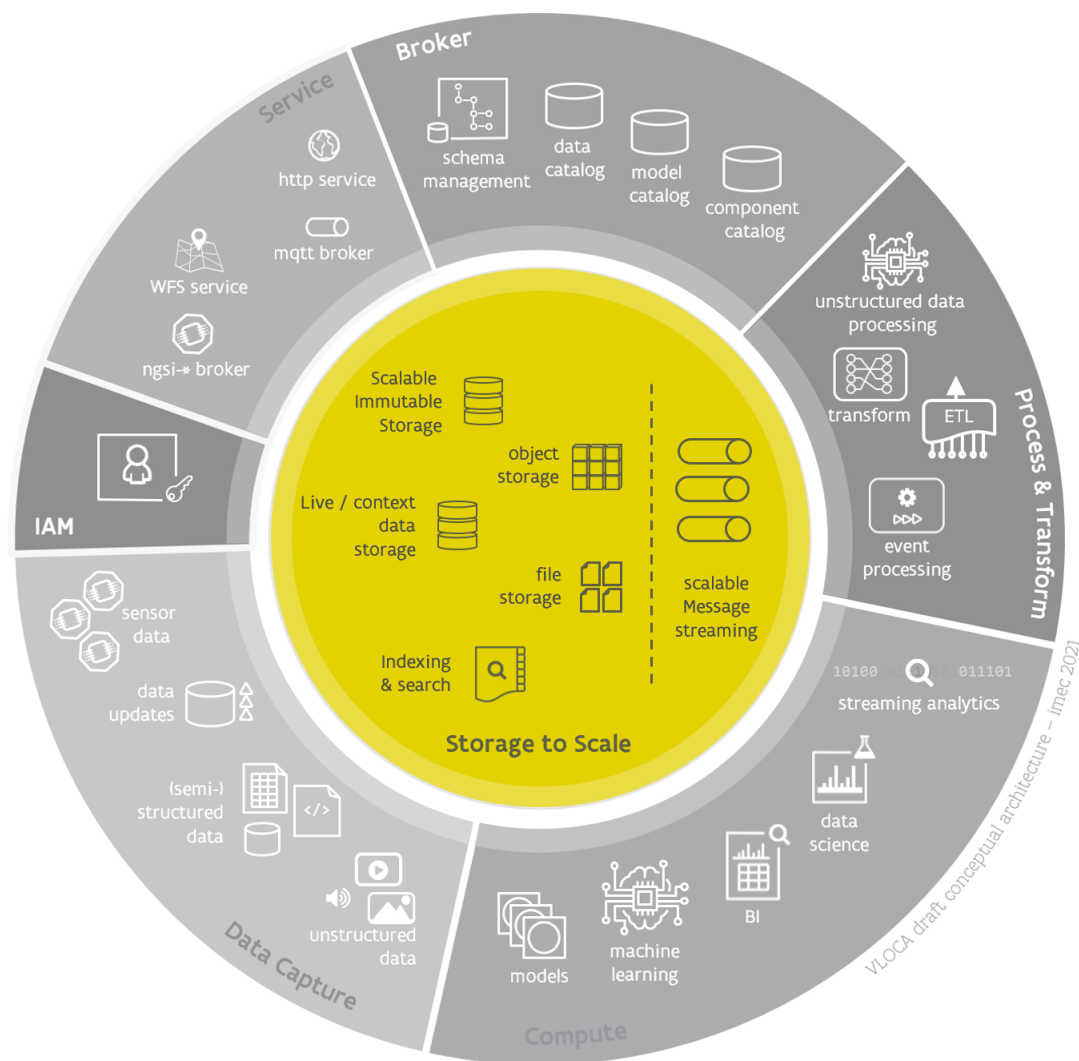
Data Capture

Instroom van data op het platform

- Sensor data
- Data transacties
- Gestructureerde data (data set)
- Semigestructureerde data
- Ongestructureerde data
 - Video/audio feed
 - Video/audio/tekst



Storage to scale



Opslag van data

- In functie van behoeften
- Opslag op schaal en fit for purpose
- Aandacht voor retentie en archivering

Data streaming

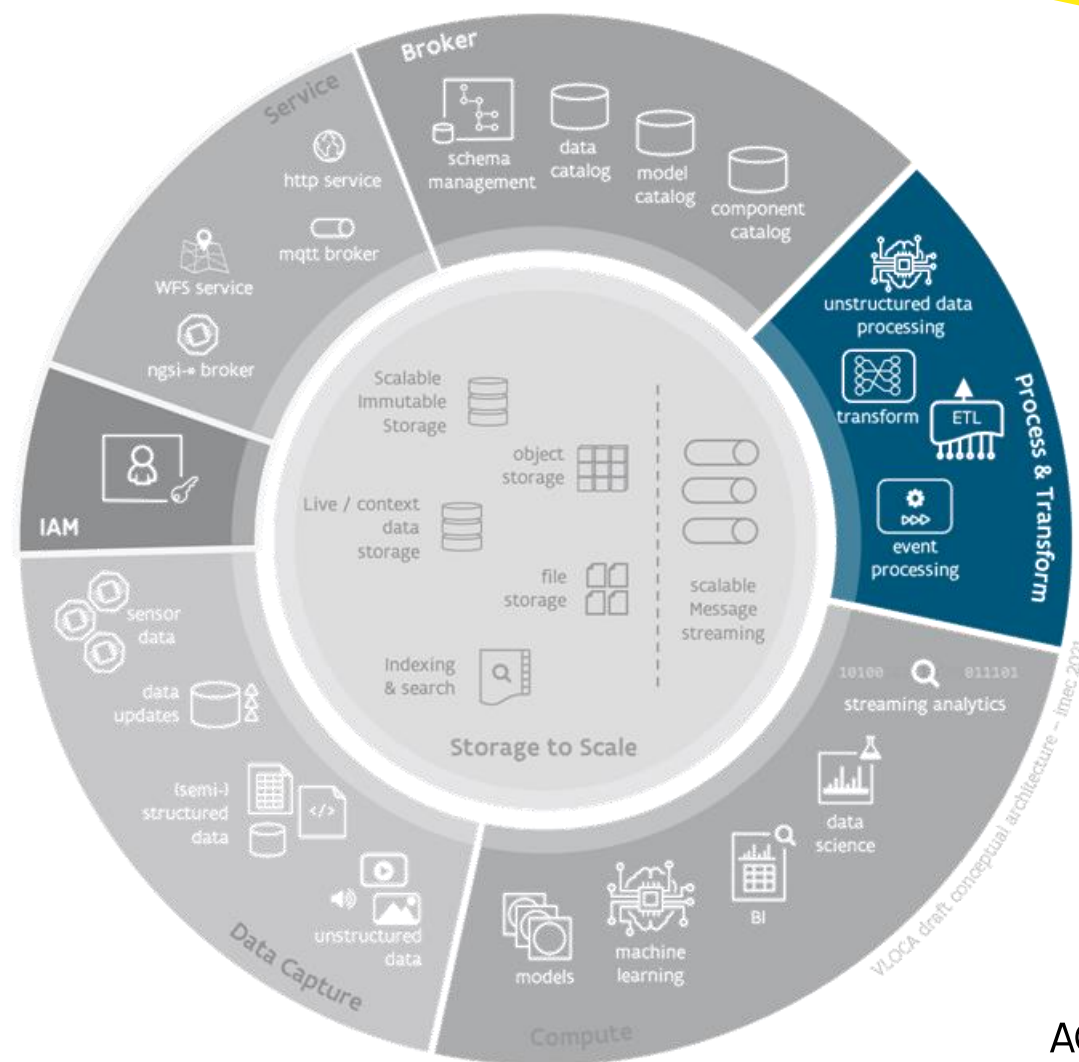
- Maakt mogelijk om data end-to-end te laten stromen
- Faciliteert data projecties (afgeleide stromen)

Process & transform

Mogelijkheid om data te 'projecteren', bijvoorbeeld:

- Verwerken
- Transformeren
- Verrijken
- Pseudonimiseren
- Aggregeren
- ...

Verwerking kan ook op streams door bijvoorbeeld machine learning toe te passen op een video-feed en om te zetten naar gestructureerde data (detectie personen, ANPR, ...).



Compute

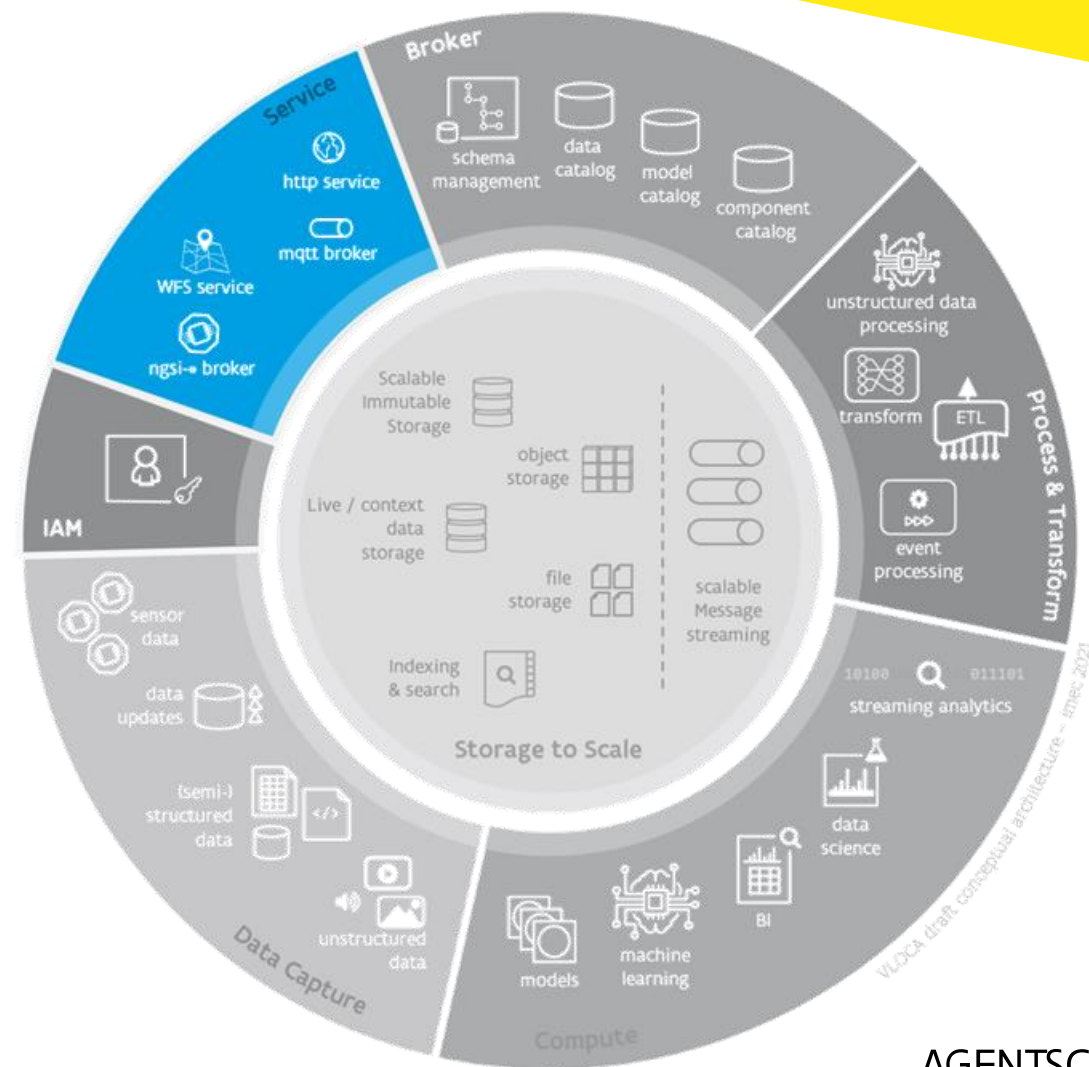
Modellen en algoritmes die met de data werken om tot nieuwe inzichten te komen en om processen te automatiseren

- **Modellen:** interoperabele en/of portabele modellen die kunnen worden ingezet voor simulatiedoeleinden
- **Machine learning:** machine learning en hun training ten behoeve van het verwerken van datasets en datastromen
- **Streaming analysis:** de mogelijkheid om datastromen al dan niet met behulp van complex modellen te analyseren (e.g., kalibratie)
- **Data science:** Tools om statistische analyse uit te voeren op databronnen om verbanden te ontdekken en computationele modellen samen te stellen
- **BI & reporting:** Tools om de data voor te bereiden voor rapportering of analyses zodat bij het opvragen van de rapporten deze data onmiddellijk beschikbaar is

Service

Publiceer dezelfde data met verschillende protocols en in verschillende formaten.

- Streams voor low-latency levering met fit-for-purpose formaten (bv. Parquet)
- Standaard interfaces voor maximale interoperabiliteit (Fiware, OGC, ...)
- De kracht van Http voor efficiënte data-transfer



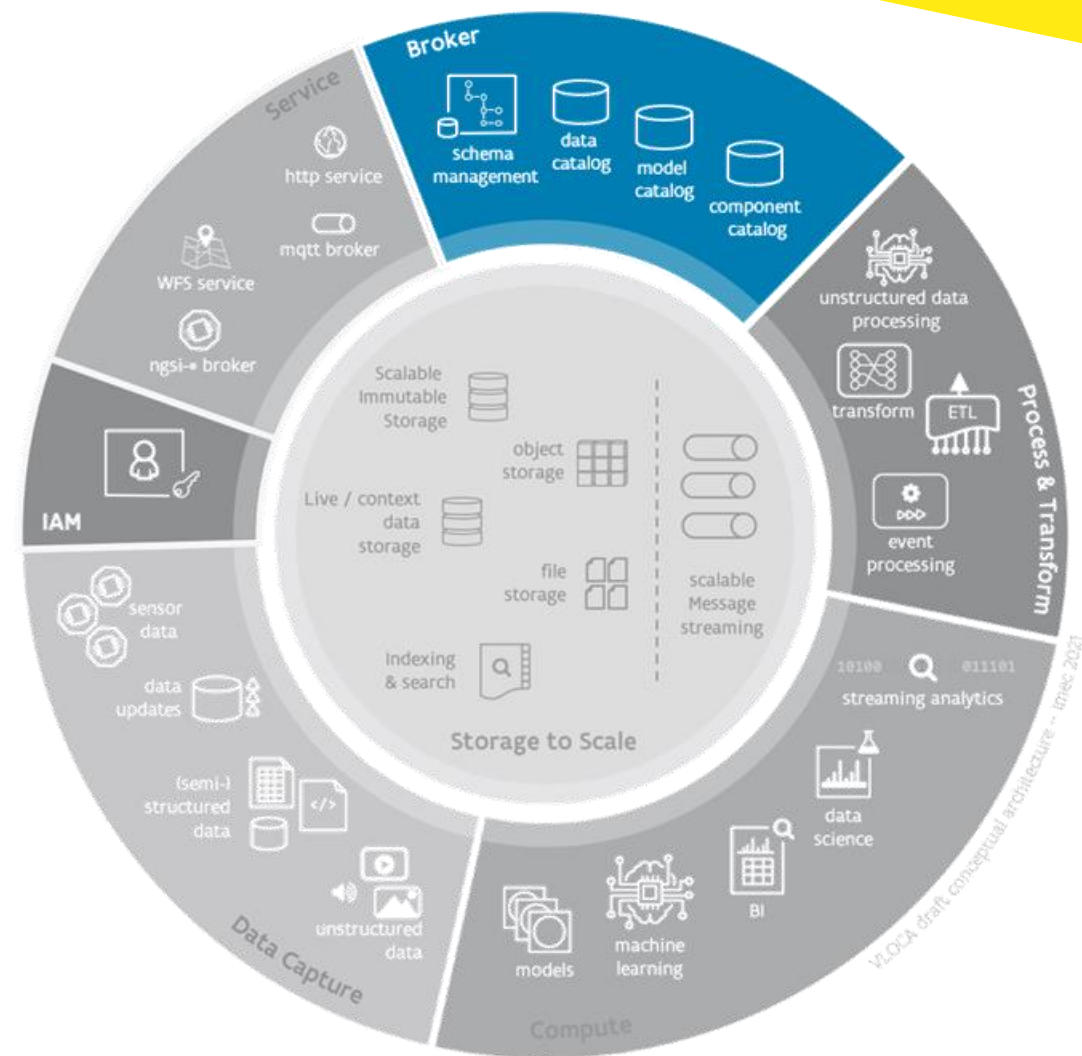
Broker

Catalogus van databronnen met metadata
Beheert *smart data*

- Refined Schema
 - Semantics
 - Syntax (schema + format)
 - Version
- Metadata

Catalogus van modellen: model-as-a-service,
mobiele modellen, koppelbare modellen

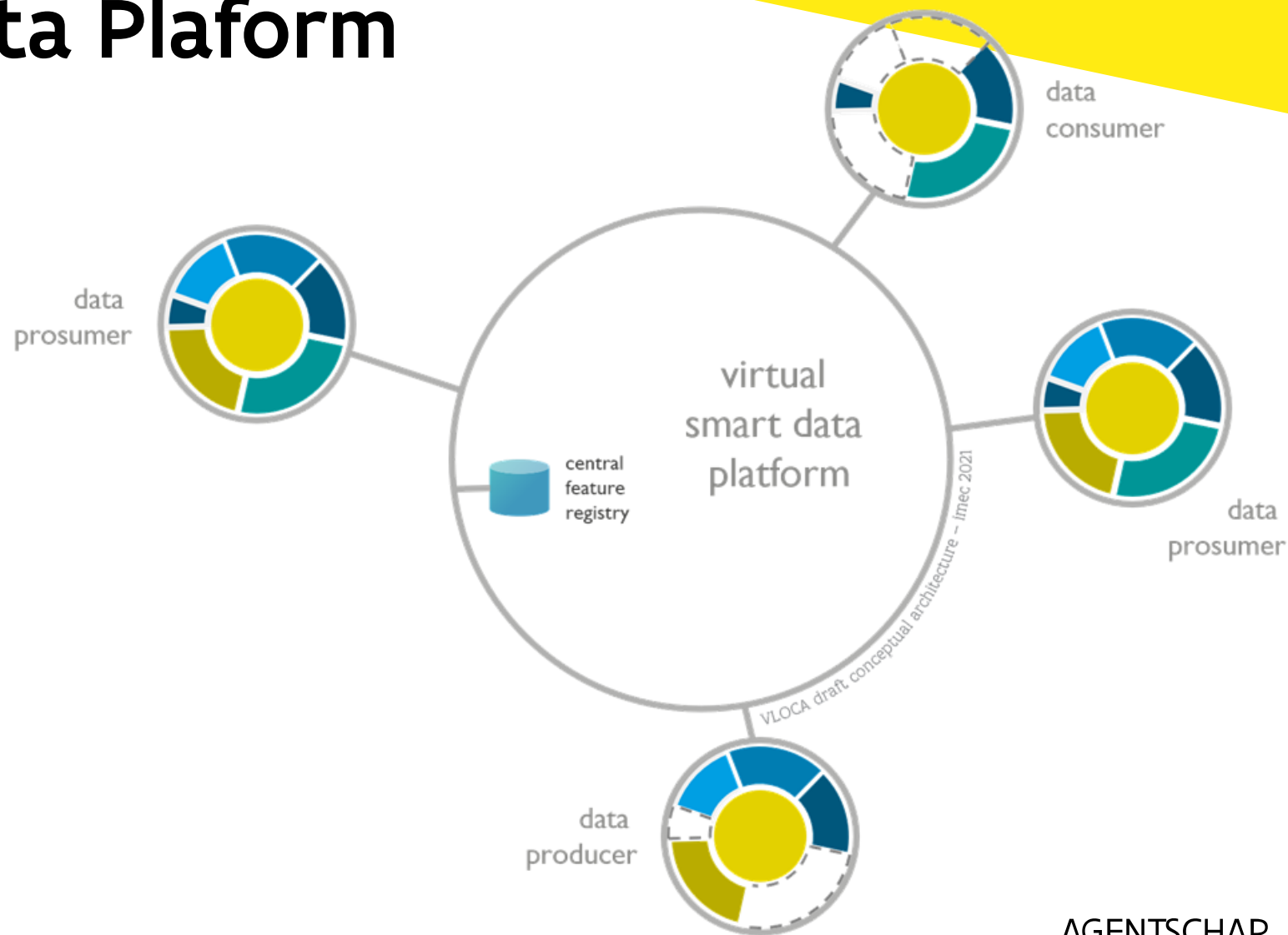
Catalogus van herbruikbare componenten: (RML)
mappings, pipeline componenten, services en
clients, ...



Virtual Smart Data Plaform

Een virtueel platform bestaande uit verschillende *tenants*

- Locale en bovenlokale besturen
- Vlaamse agentschappen
- Onderzoeksinstellingen
- Intercommunale dienstverleners
- Bedrijven



Linked Data Event Streams

bouwstenen voor een open sensor data platform

**DIGITAAL
VLAANDEREN**



**Vlaamse
overheid**



Recap vorige toelichting vanuit Relance: Probleemstelling

Hoe maken we van real time sensordata een belangrijke hefboom voor duurzame groei van de Vlaamse data-economie?

- huidige platformen zijn niet altijd geschikt om grote volumes sensordata aan te bieden op een kostenefficiënte manier
- smart city toepassingen zijn vandaag nog te veel maatwerk en onvoldoende schaalbaar
- sensordata ontstaan in silo's en bieden op zichzelf onvoldoende meerwaarde voor nieuwe business modellen, niet gestandaardiseerd
- een dreigende vendor lock-in op sensordata en geconsolideerde smart city markt vormt een sterke rem op de innovatie

Hoe? En welk effect beogen we?

Sensordata zijn een groeimarkt voor de data-economie van morgen

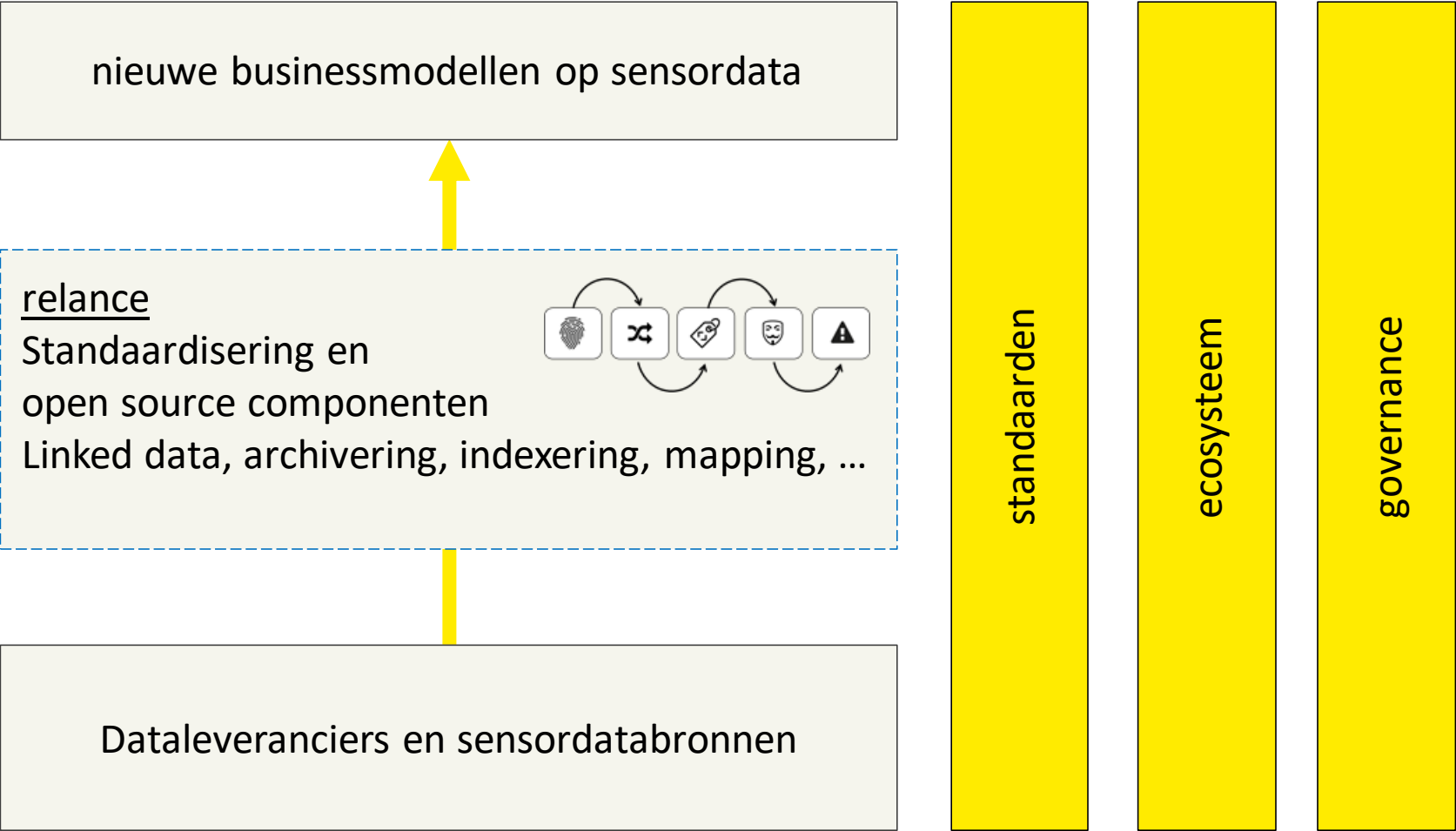
Doelstellingen

- Herbruikbare **open source componenten** aanbieden aan de markt die kostenefficiënte publicatie van grote volumes sensordata in real time toelaat
 - interoperabiliteit
- Inzetten op **open standaarden** zodat sensor data sensor-agnostisch wordt en 'loskomt' van de leverancier vd sensor en vlot gerelateerd kan worden aan slow moving data (vb. **basisregisters**)
- **Ecosysteem** van 'samenwerkende' sensor dataplatformen tot stand brengen

Business voordelen

- Innovatiesnelheid en time to market van Vlaamse bedrijven gevoelig verhogen. Dit is een voorwaarde voor de **veerkracht** van onze Vlaamse bedrijven
- **Realtime** en **gepersonaliseerde** data zijn een vereiste voor digitale transformatie geworden
- Self Service by design en betaalbaar te gebruiken door de **gehele markt** (incl KMO)

Sensor data platform



Wat we beogen...

Publicatie

Moet toelaten om sensordata vlotter duurzaam te publiceren op een interoperabele manier

- Afkomstig van vele input formaten
- Verder bouwen op Oslo standaarden
- Hergebruik toelaten van reeds uitgevoerde mapping en optimalisaties door andere partners in het ecosysteem
- Een oplossing bieden voor gestandaardiseerd archiveren van data

Hergebruik

Moet toelaten sensor data en context data te raadplegen

- Voor mens en machine
- Schaalbaar
- Puntbevragingen en batch operaties: Makkelijk bevroagbaar en combineerbaar over domeinen heen

Hoe?

Standaarden

- We ontwikkelen nieuwe standaarden en implementeren deze.
- We werken een Europese specificatie uit voor Linked Data Event streams

Ecosysteem & Community

- We haken in op bestaande ecosystemen en communities
- We zetten een brede ecosysteem werking op poten tussen publieke, privaat en academische wereld (incl de developer community) → adoptie
- We supporteren het ecosysteem met een testbed/incubator, specificaties en bouwblokken

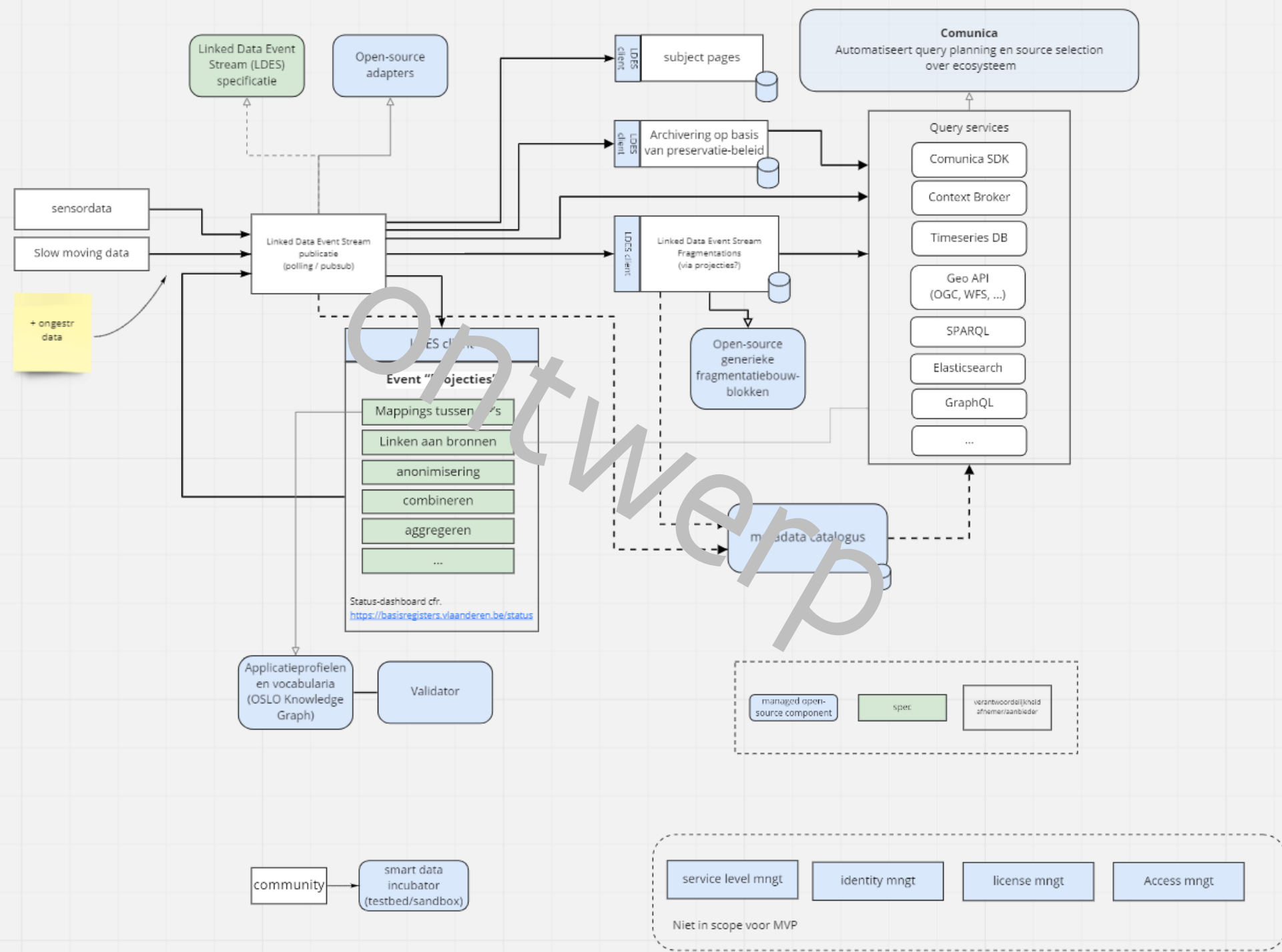
Governance

- Governance van standaarden en bouwblokken
- Bouwblokken marktplaats

Bouwblokken

- Uitwerken bouwblokken als managed services (sleutelbouwblokken zelf uitwerken) en onderhouden deze.

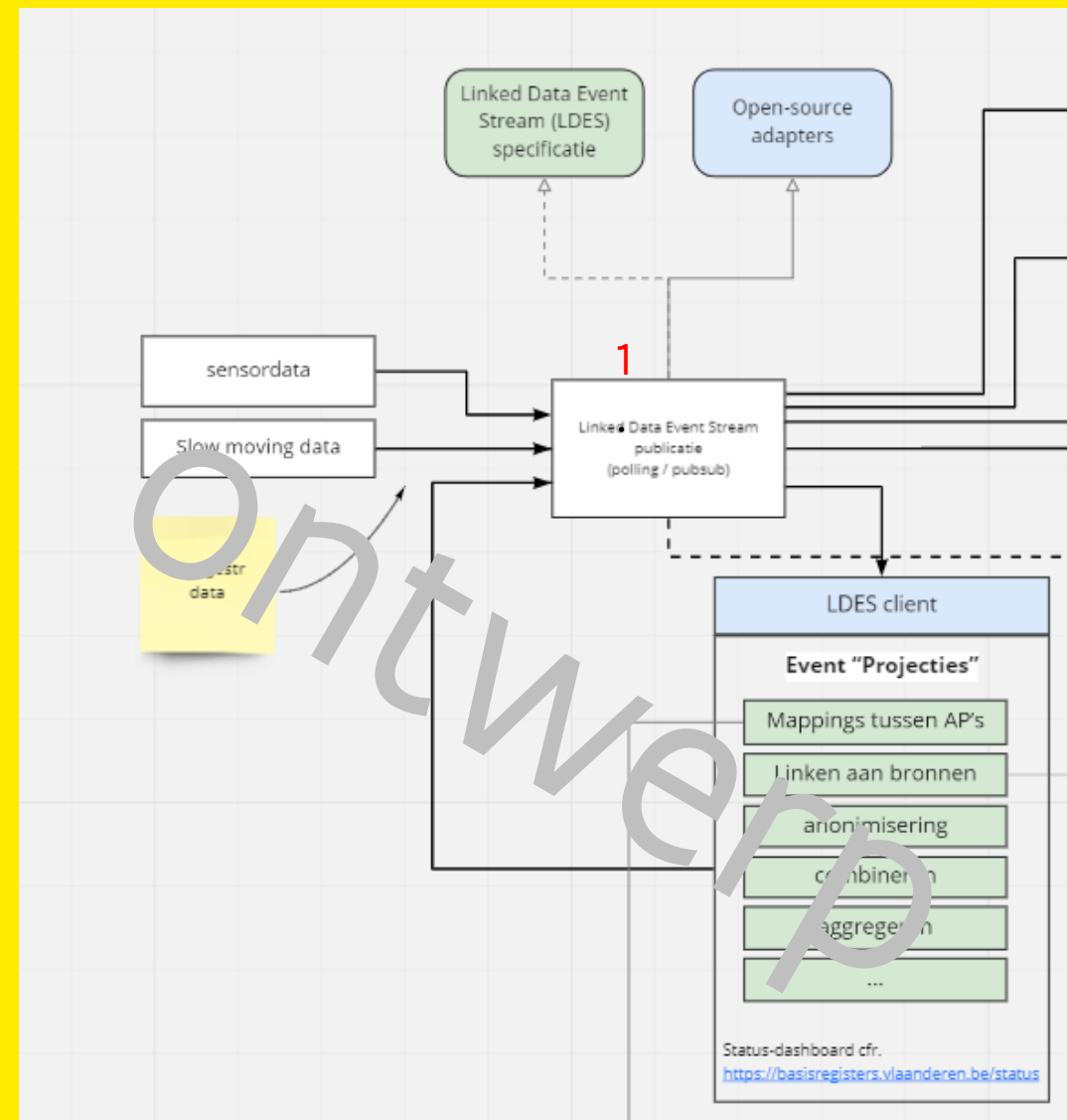
Functionele architectuur (Ontwerp)



1. Je ruwe datastroom publiceren als interoperabele linked data stroom

REAL TIME RUWE DATASTROOM

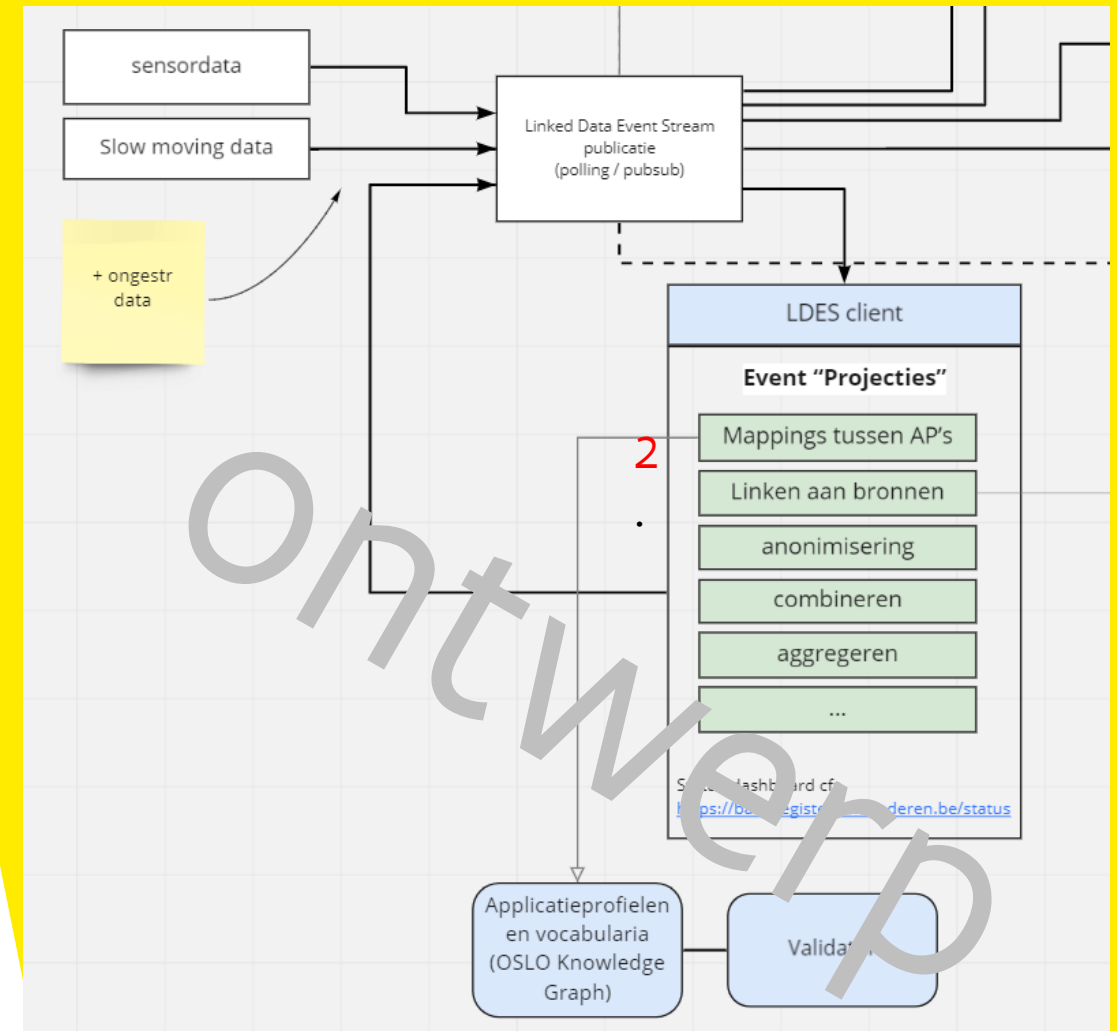
- Data kan in eender welke API aangeboden worden
- Omzetting naar een LDES (Linked data even stream publicatie): we beschouwen elke datastroom als een update reeks van **events**.
- Support – LDES specificaties
- Support – **open source adapters** (cfr een **reisstekker**) die je toelaat om een specifiek inkomend formaat te mappen op bestaande OSLO standaarden. (mappen dan delen)



2. LDES clients voor verschillende doelen

CLIENT VOOR VERSCHILLENDE AFGELEIDE DATASTROMEN

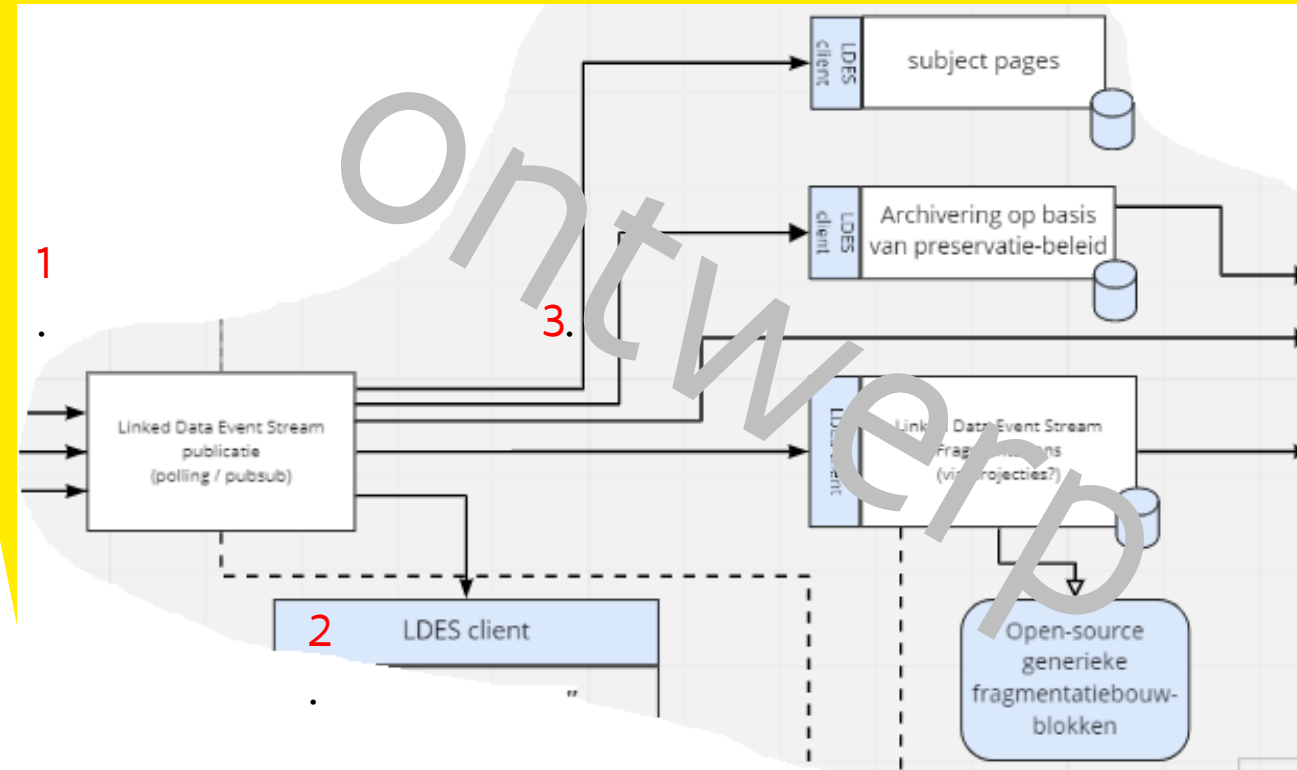
- Afgeleide datastromen creëren
 - via projecties (LDES client managed) en terug publiceren als LDES
- Daaronder specificaties die we gaan voorzien om afgeleide datastromen als event stream te genereren.
- Vb mappingcomponent applicatieprofiel
Vb linken aan bronnen, anonimisering, aggregatie etc...
- Valideren: OSLO Knowledge Graph en validator)



3. Verschillende LDES clients voor verschillende doelen

COMPONENTEN ONAFHANKELIJK – VERSCHILLENDE MANIER OM DE DATA TE PUBLICEREN

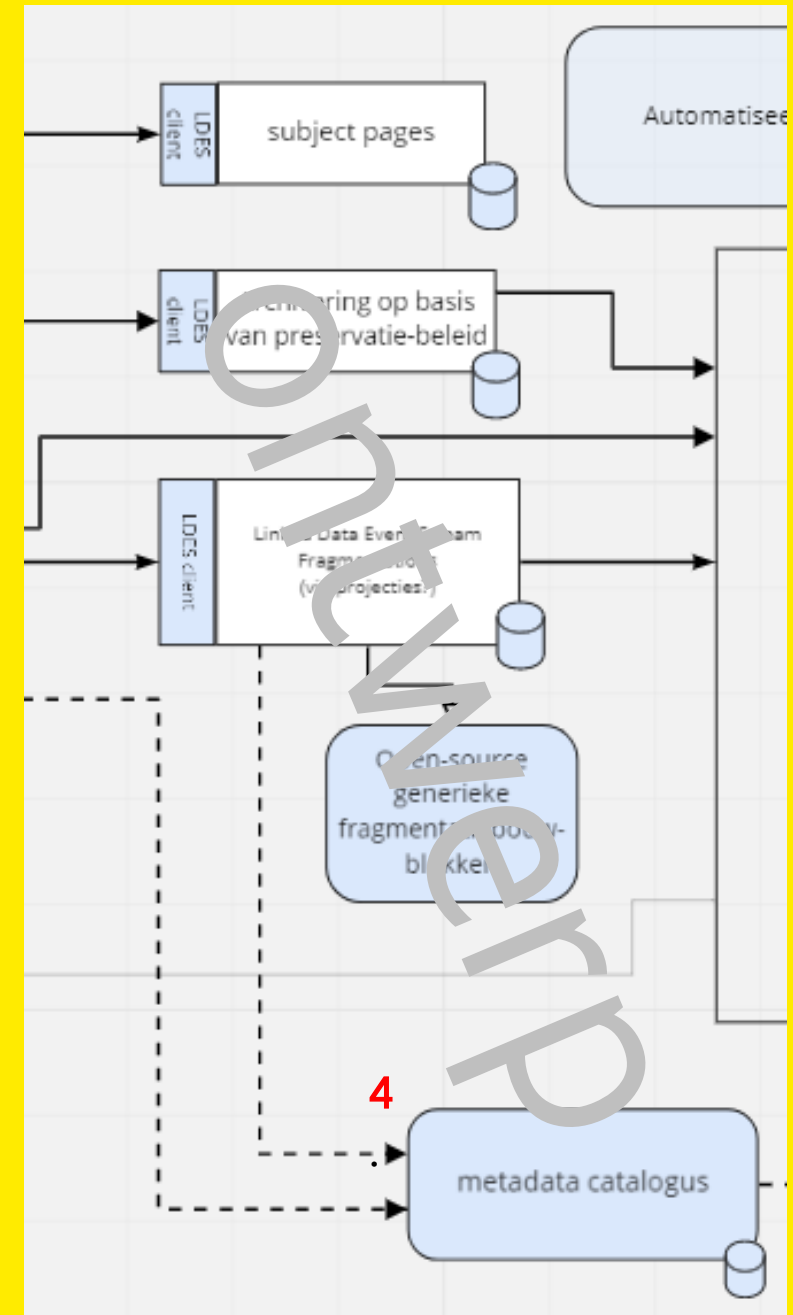
- LDES client – subjectpages
 - URI raadpleegbaar voor mens en machine
- LDES client – archivering/retentie
 - We voorzien een specificatie en bouwblok ifv een gestandaardiseerd retentiebeleid. (YTT scenario's)
- LDES client – Fragmentaties



4. vindbaarheid, herkomst van de data

- Metadata catalogus: vindbaar maken van de data/metadata

3.



5. Ontsluiting data vanuit fit-for-purpose storage

- Ontsluiting data ifv de vraag/behoefte van de afnemer (bijvoorbeeld WFS, NGSI, HTTP, ..)
- Afstemmen van opslagmethode op ontsluiting

Open Discussie – AS IS

Hoe kunnen bestaande
platformen / architecturen / initiatieven
hieraan bijdragen en/of hiermee aligneren?

KH Open voor
co-creatie

https://vloca-kennishub.vlaanderen.be/vloca-kennishub/Categorie:Data_Broker

AGENTSCHAP
**BINNENLANDS
BESTUUR**

Open Discussie – TO BE

Welke initiatieven zitten er aan te komen en hoe kunnen deze hiermee aligneren ?



https://vloca-kennishub.vlaanderen.be/vloca-kennishub/Categorie:Data_Broker

Meer informatie



vloca.vlaanderen.be
vloca.vlaanderen.be/trajecten/Mobiliteit-
mobipunten/



vloca@vlaanderen.be



AGENTSCHAP
BINNENLANDS
BESTUUR