

CHALLENGES ON DATA SHARING

AND HOW RESEARCH HELPS ADDRESSING THEM

Researching consortium driven Data Exchanges for AI/ML development

Leon Gommans, PhD
Science Officer

AIRFRANCEKLM
GROUP

Professor Data Exchange Systems,
University of Amsterdam, SNE



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769288



UNIVERSITEIT VAN AMSTERDAM

SURF



ESnet
ENERGY SCIENCES NETWORK



Industry

Research

CONTENT

What is Trust?

Context: two sided markets for AI development

What are the challenges to overcome?

- Platform archetypes in the light of trust
- Consortium creating challenges

How can trust be organized?

How to implement trust ?

Examples of solving some challenges.

WHAT IS TRUST ?

COMPLEX CONCEPT, MANY THEORIES

Dimensions (Bachman):

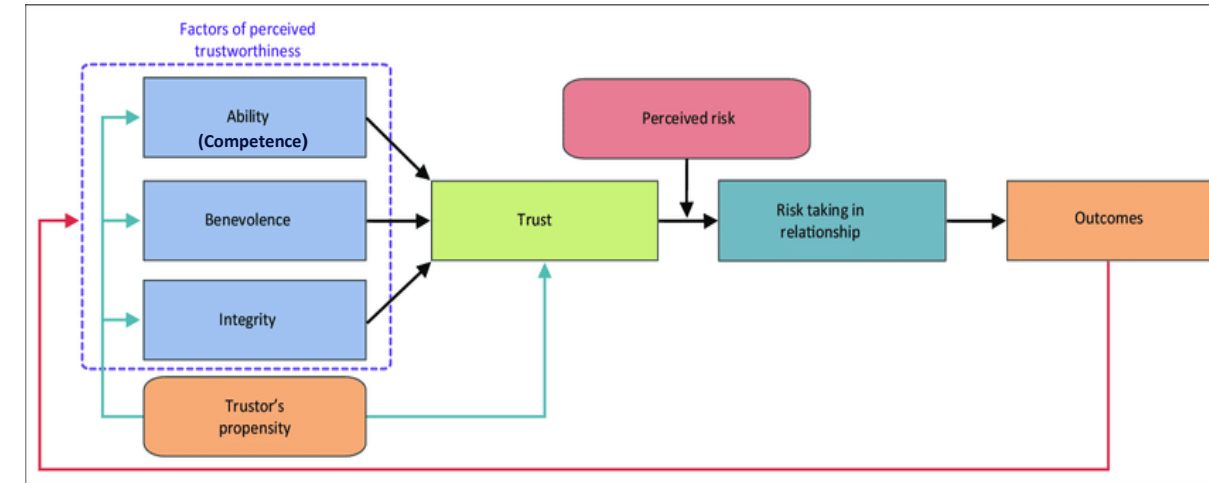
(Inter-) personal trust: Formed by the interaction between persons and growth with experience.

Rooted in the tacit understanding of personal trust he also recognizes **impersonal trust**, with sub-categories:

System trust: e.g. safe operation of a plane as a system with oversight from aviation authorities

Institutionalized trust: Organizations interacting based on rules, standards, code of conduct established by trade organizations, industry forums, standards bodies, or a dominant player.

A trust model (Mayer)

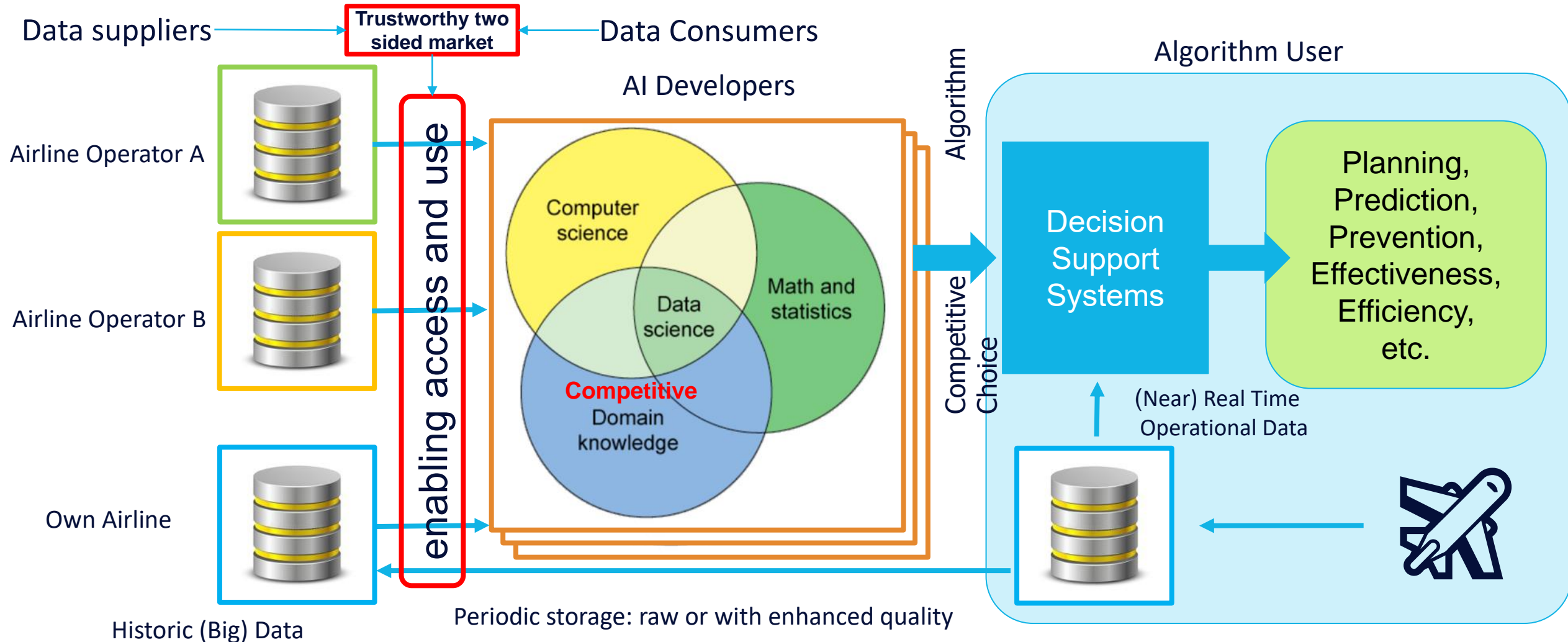


Source: Mayer, R.C., Davis, J.H., & Schoorman, F.D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20, 715.

*Trust is the **willingness** of a trustor to be vulnerable to the actions of a trustee based on the expectation that the trustee will perform a particular action important to the trustor,...*

CONTEXT: DATA SHARING FOR AI DEVELOPMENT

A MODEL STUDIED



PLATFORM ARCHETYPES FOR SHARING DATA

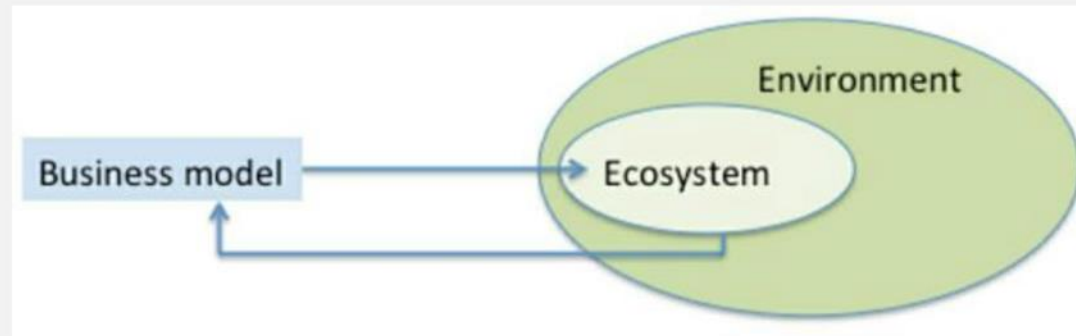
DRIVEN BY DIFFERENT WAYS TO ORGANIZE PLATFORM GOVERNANCE

Driver	Self interest		Common interest	
Trust	Trust a Single Party		Organize Trust	
Driven by	Existing enterprise	Investors in new enterprise	Alliance / Consortium (with specific aim)	Federation (with holistic aim)
Archetype	Internal platform offered externally	Centralized Platform	Distributed Platform	Federated Business Ecosystem
Goal	Be the best in your environment	Create shareholder value (typically at the expense of the existing environment)	Extend <i>reach</i> for suppliers and/or <i>offering</i> to a client using a common environment	Arbitrary collaboration with environment often for the common benefit of the environment
Role IT	Efficiency: Digitizing as means	Support disruption with Agility: Move fast / break things	Support standards to integrate with a common <i>alliance / consortium governed infrastructure</i>	Support creation of new business models by integrating own services with standardized, neutral federation services & Infra
Example	GE Predix	Uber	SkyTeam	Dataspaces (GAIA-X)

BUSINESS MODEL THINKING & ECOSYSTEMS

FUTURE: TOWARDS FEDERATIVE COLLABORATION VIA PLATFORMS (GAIA-X)

THE INTERACTIONS BETWEEN THE BUSINESS MODEL OF THE ORGANIZATION, THE ECOSYSTEM AND THE ENVIRONMENT



Source: “**Business model thinking**”, **business ecosystems and platforms: The new perspective on the environment of the organization**

Benoît Demil, Xavier Lecocq, Vanessa Warnier
Dans M@n@gement 2018/4 (Vol. 21),
pages 1213 à 1228

Through the design and implementation choices, encapsulated in a business model, a company chooses its stakeholders and its importance (i.e. its bargaining power) in the ecosystem.

Enabling sovereignty of choices is a key: Requires neutrality of exchanges.

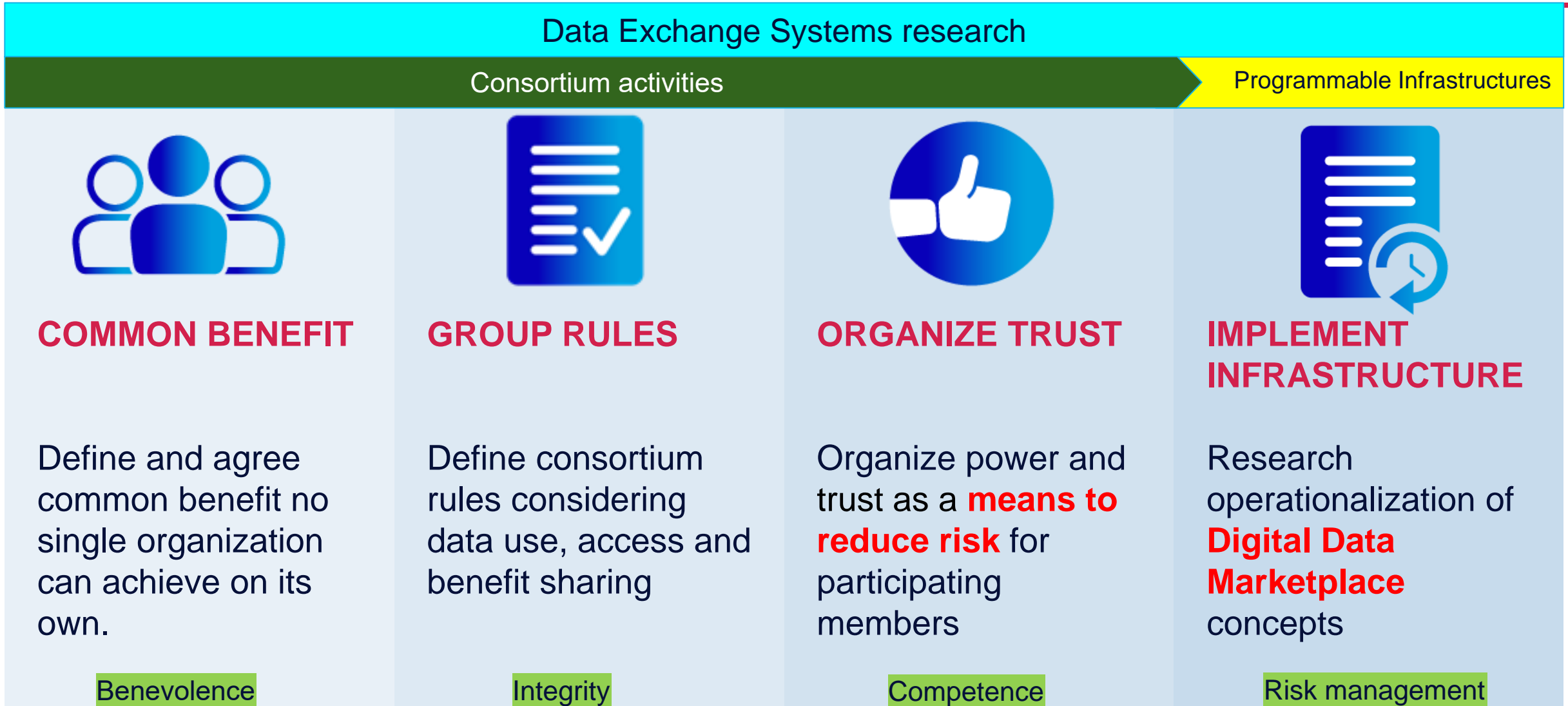
Competition is formed and defined between business ecosystems

Business ecosystems are believed to be capable of better explaining **how multi-sided businesses evolve** (demand/supply side)

Data can have a demand and supply side.

DIGITAL DATA MARKETPLACE APPROACH

ORGANIZING TRUST USING TRUSTWORTHY INFRASTRUCTURE FOR SPECIFIC GOAL



CHALLENGES TO BE CONSIDERED

BY A CONSORTIUM

Many organizations want to keep their historical data in their sovereign data zones.

Many implications need to be considered, to name a few:

Business level

Value
Cost
Benefits
Agreements
Exchange
Trade

Legal level

'Ownership'
Access
Usage
Compliance
Liability
Market Rules

Data level

Processing
Storage
Management
Transport
Transform
Security



DEMONSTRATED USE-CASE – THE 747 BLEED AIR SYSTEM

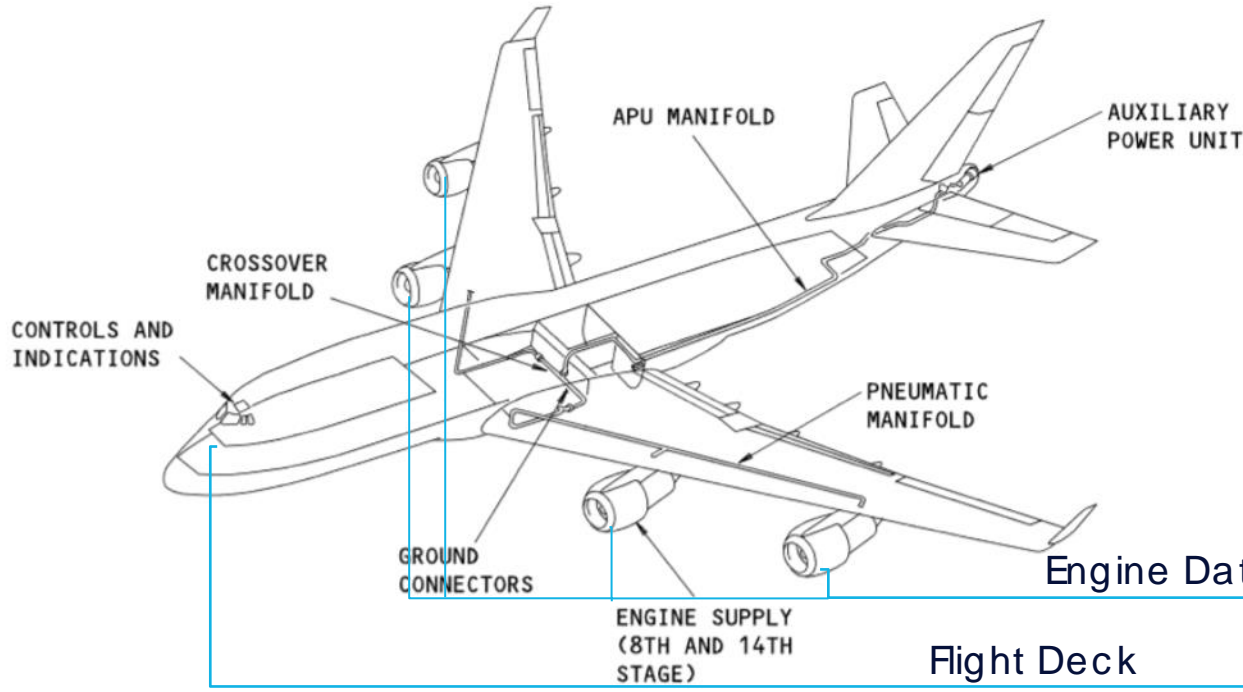
DATA FROM EXISTING USE-CASE NOW SPLIT ACROSS THREE PLACES

Imagine if data scientist can use historic data from specific aircraft types operated by multiple airlines

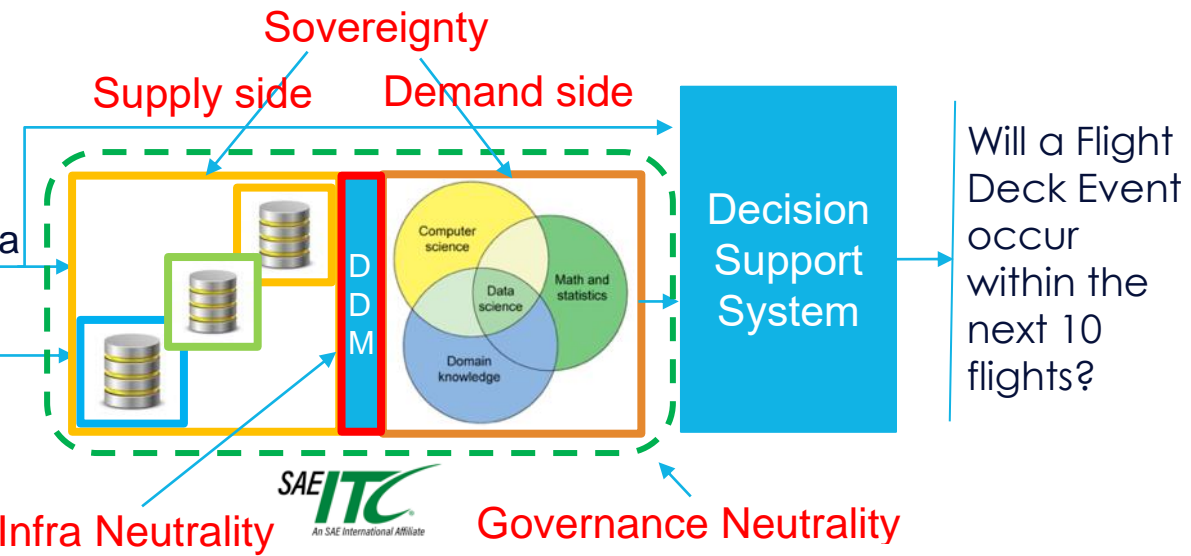
The Bleed Air System regulates pressure and temperature of air from a turbine engine needed by other aircraft systems taking care of:

- cabin pressure
- de-icing
- water pressure
- and more..

Flight Deck Effects indicate system functionality decreases and may trigger maintenance actions



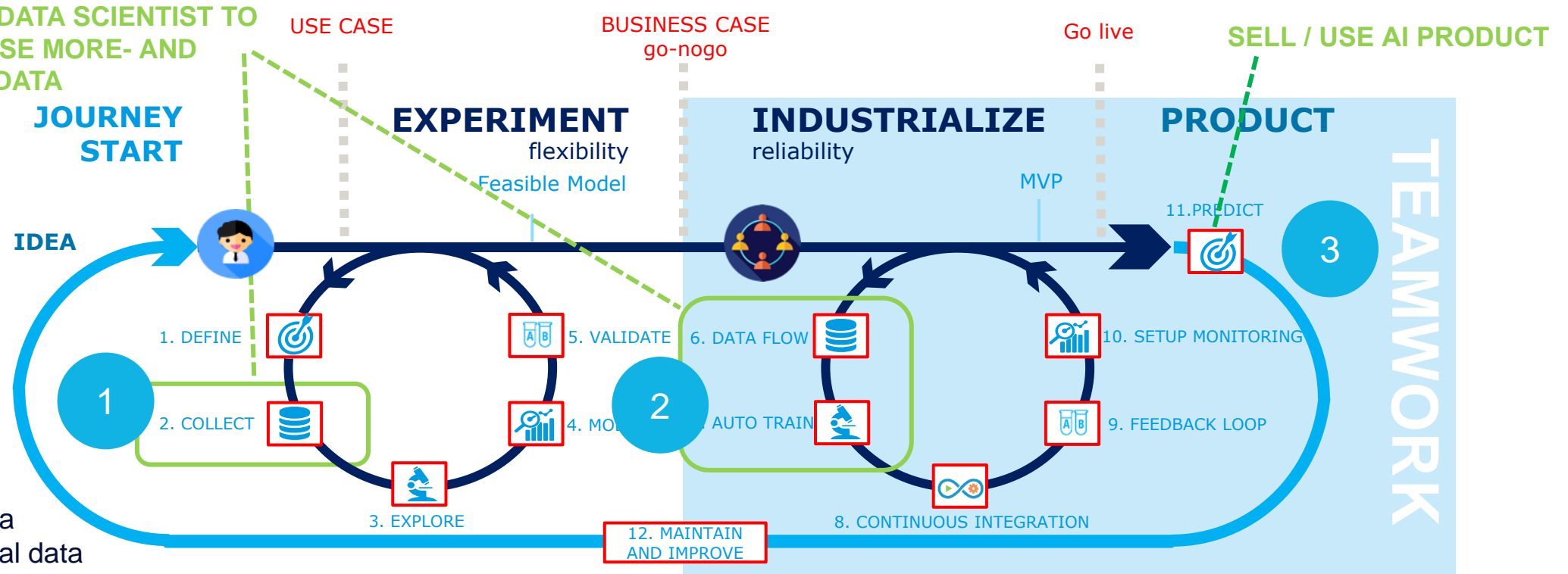
The more Flight Deck Effect occurrences are available, the more likely that a prognostic relation can be learnt.



JOURNEY OF THE DATA SCIENTIST / ENGINEER

ROLE OF THE DIGITAL DATA MARKETPLACE (DDM):

DDM ENABLING DATA SCIENTIST TO
EXPLORE AND USE MORE- AND
MORE DIVERSE DATA



Role DDM:

- 1 Explore sample data
- 2 Industrialize with real data
- 3 Offer AI products

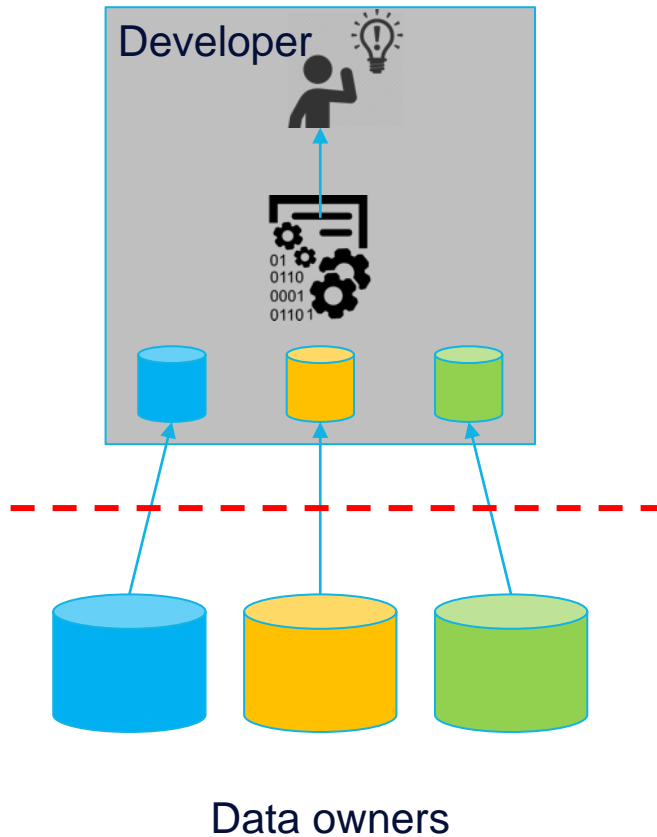
Offering seamless integration
With AI sandbox environment

ESSENTIAL INFRASTRUCTURE ARCHETYPES

MANY VARIANTS: FOCUS ON CONSORTIUM DRIVEN APPROACH TO ORGANIZE TRUST

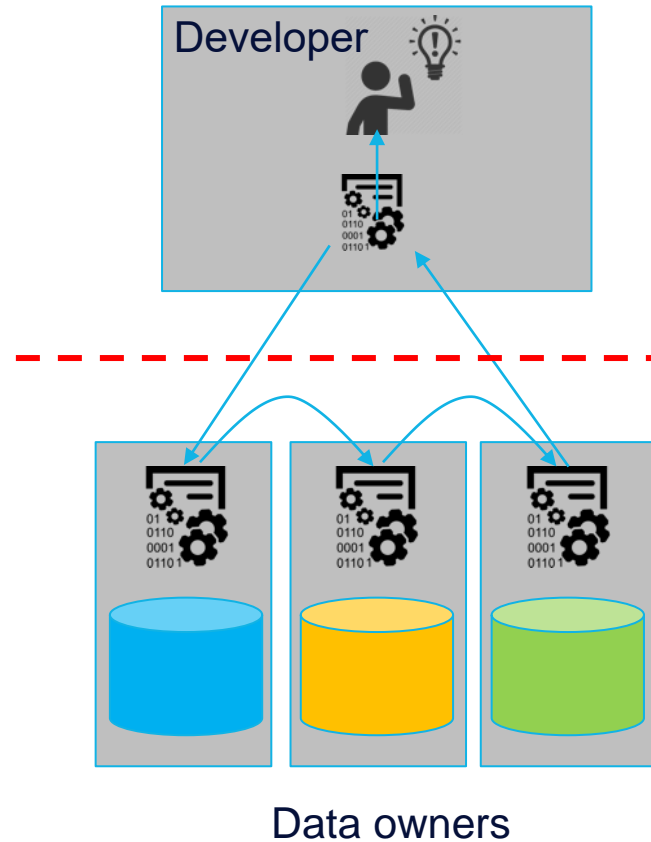
Centralized

Bring data to the algorithm



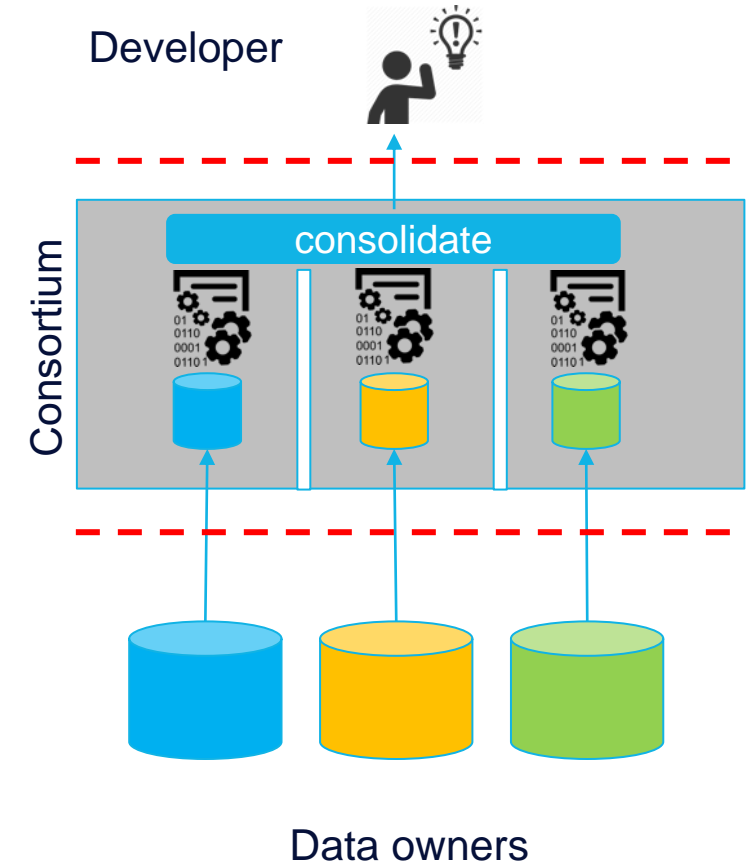
Distributed

Bring algorithm to the data



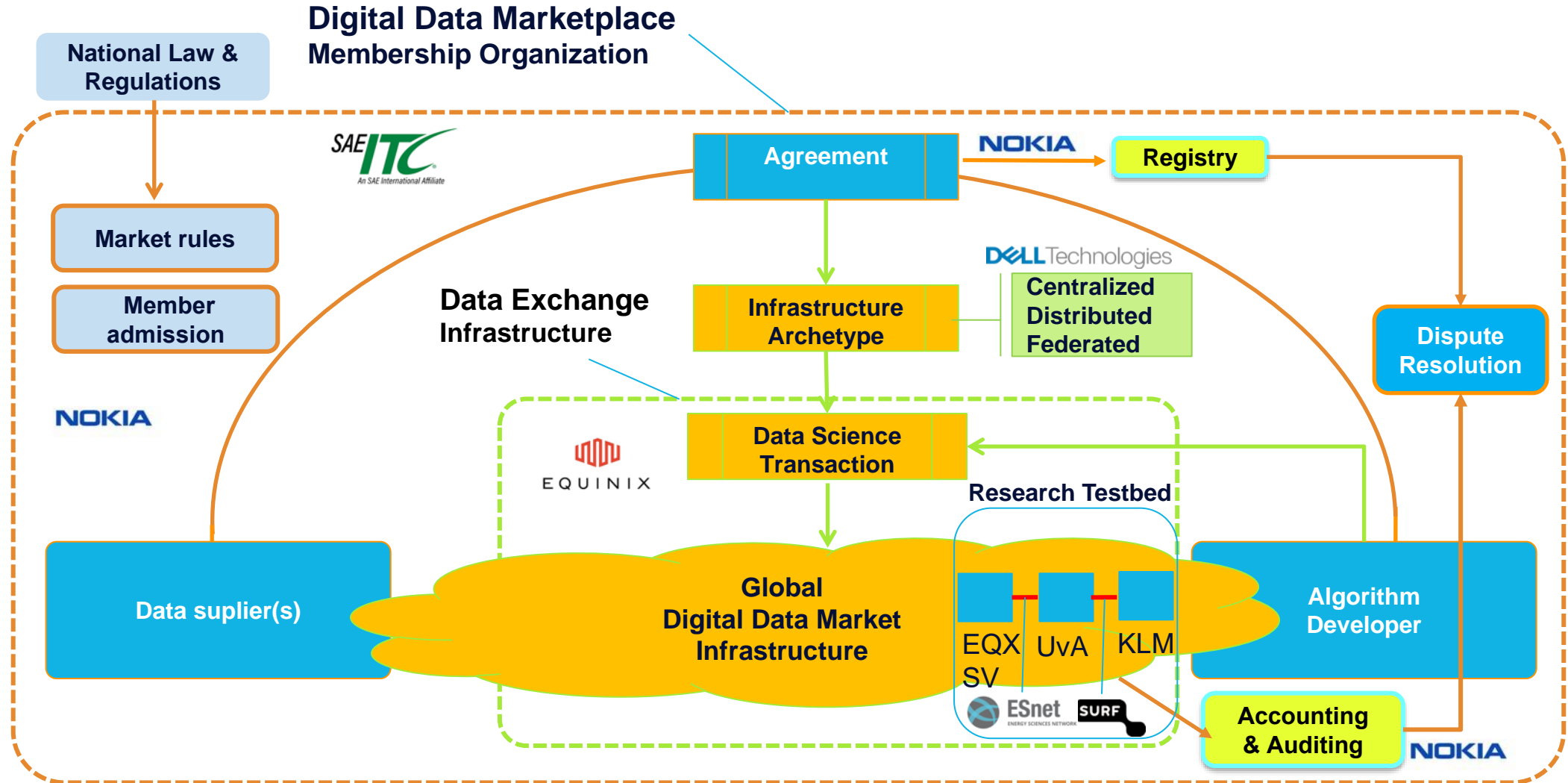
Federated

Using trusted infrastructure



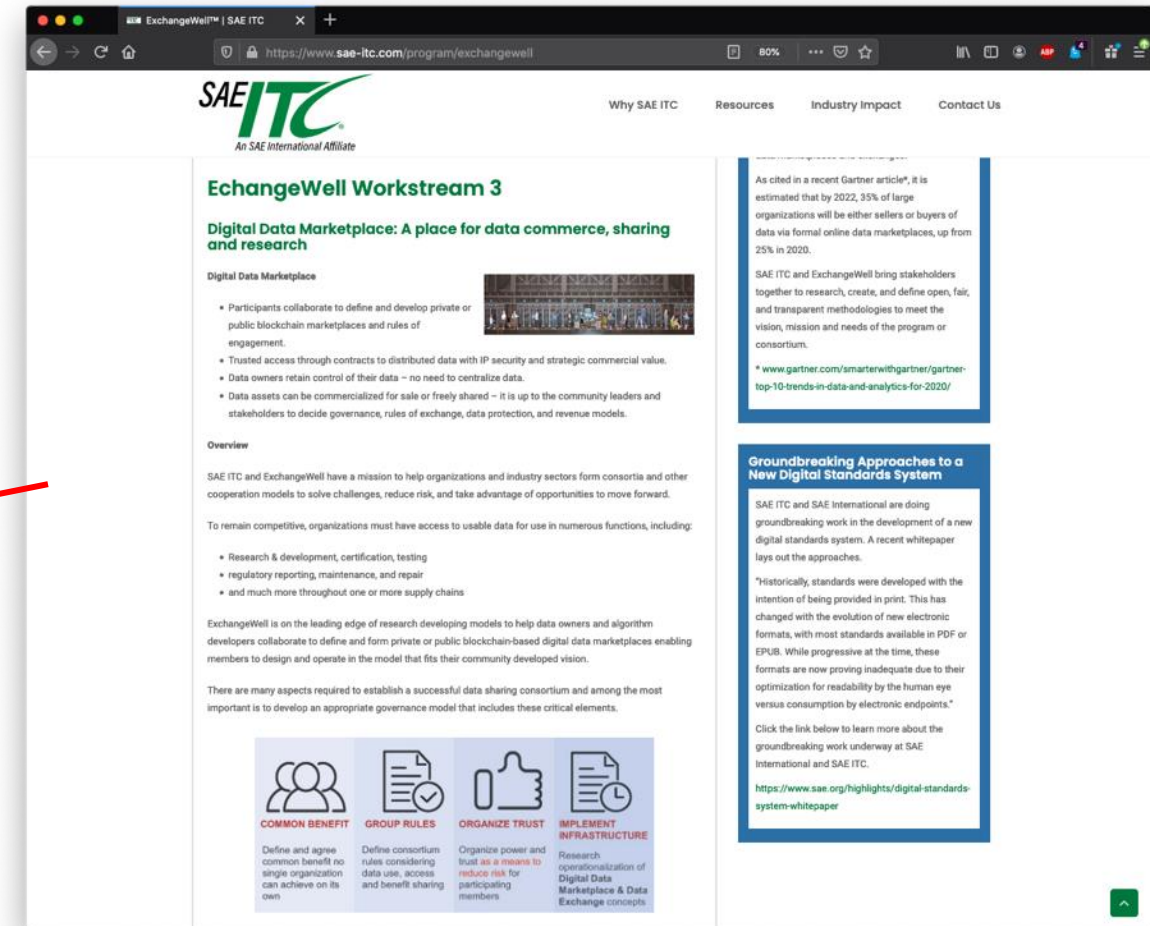
DIGITAL DATA MARKETPLACE ARCHITECTURE

RESEARCHING IMPLEMENTATION OF ESSENTIAL ELEMENTS WITH IT INDUSTRY



SAE ITC EXCHANGEWELL SUPPORTS DDM CONCEPT

<https://www.sae-itc.com/program/exchangewell>



SAE ITC
Is member of
and
Endorses



Is member of

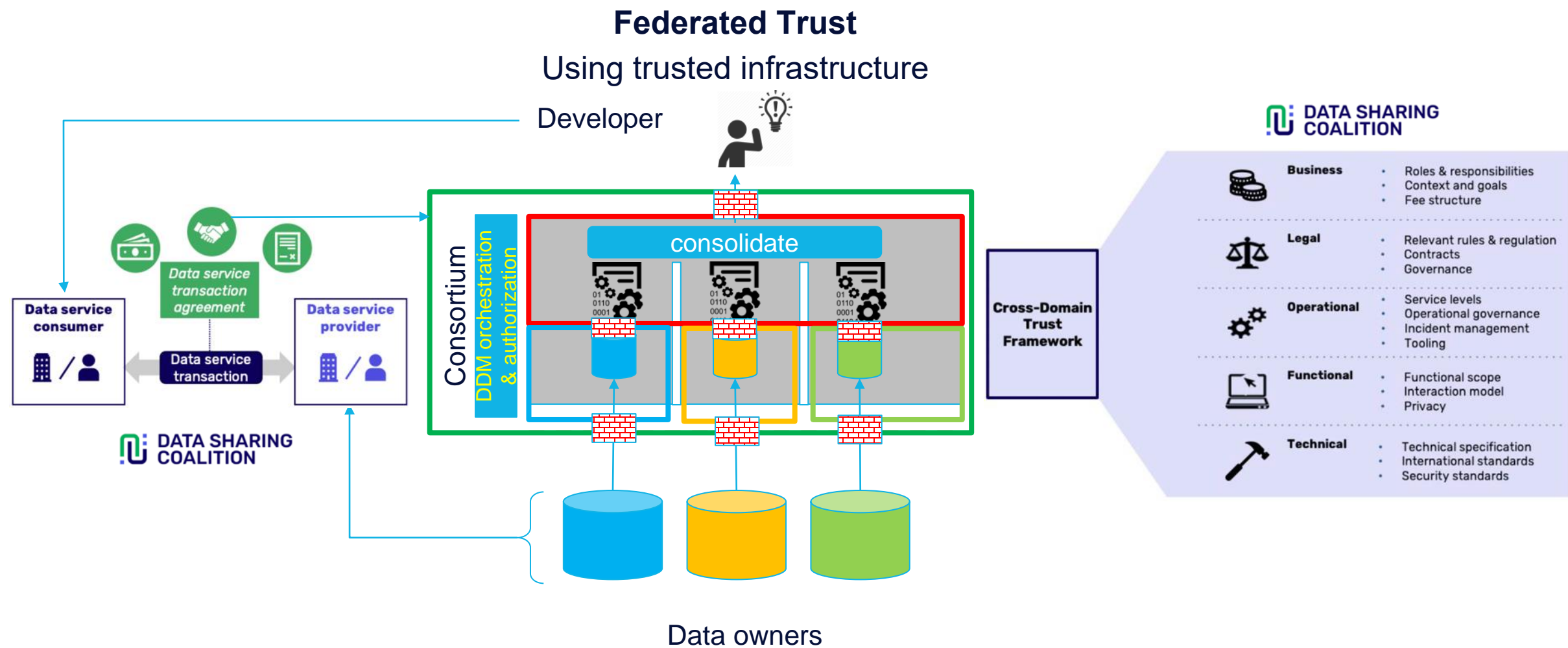


UNIVERSITEIT VAN AMSTERDAM



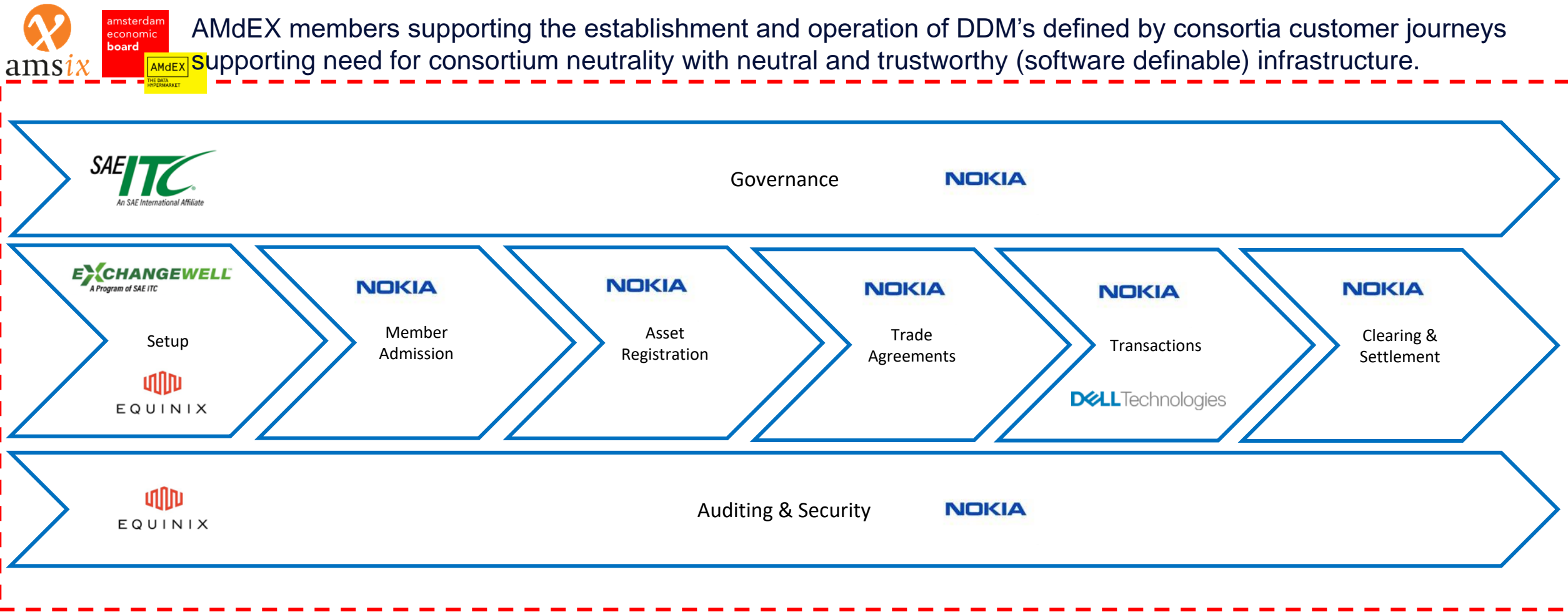
FEDERATED ARCHETYPE IMPLEMENTATION

CONSORTIUM BUILDING USING DATA SHARING COALITION CANVAS & APPROACH



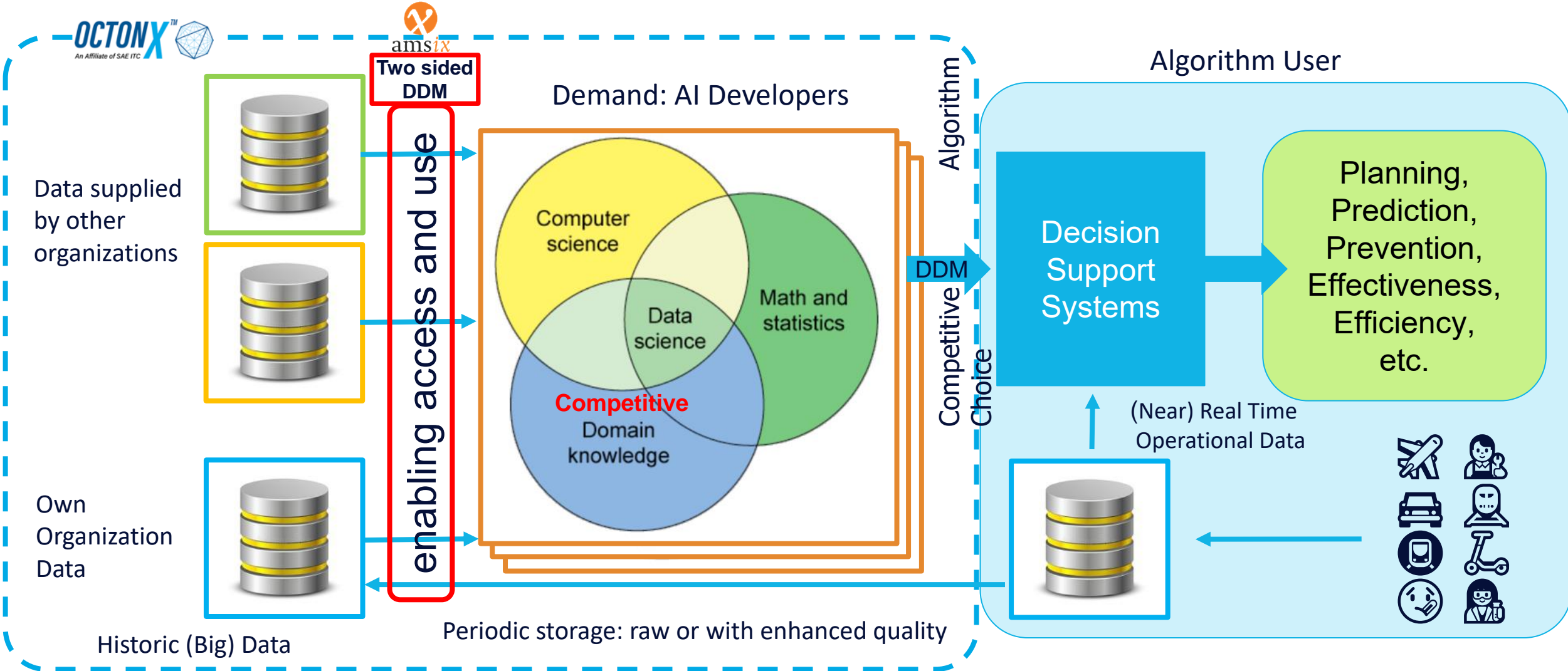
DEMONSTRATED DIGITAL DATA MARKETPLACE AS PROTOTYPE

AMS-IX PROJECT AMDEX AS FACILITATOR TO ENABLE DDM'S



USE-CASE: DATA SHARING FOR AI DEVELOPMENT

USING A DIGITAL DATA MARKETPLACE GOVERNED BY A MEMBERSHIP CONSORTIUM



DIGITAL DATA MARKETPLACE SOLUTION NOKIA/EQUINIX

FROM TRL5 TO TRL9

Research creating a TRL 5 prototype performed within DL4LD project lead to product
<https://www.nokia.com/networks/services/nokia-data-marketplace/>



Nokia Data Marketplace Solution at Equinix

Download

Whitepaper showcasing how Nokia Data Marketplace is complemented by Equinix infrastructure.

2.2 Data Marketplace Overview

The high-level architecture of a data marketplace is depicted in Figure 1. It shows the data marketplace as an entity owned and operated by a membership organization through which data providers and consumers can interact and transact based on community rules and individual agreements.

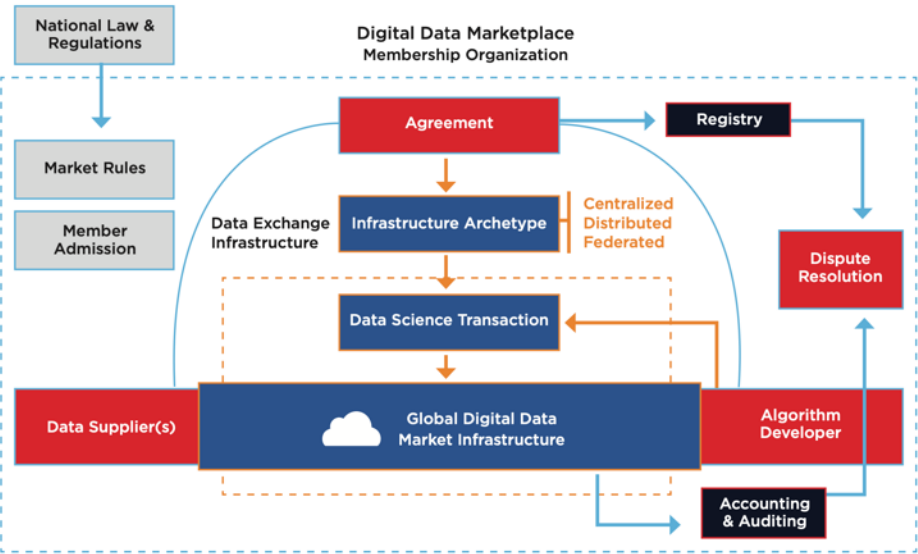


Fig. 1. Data marketplace architecture

A data marketplace must have a process for creating membership rules, and its process for admission must require a prospective member to agree to comply with all membership rules. A member might be a data supplier, algorithm developer or data services provider. After being admitted, members can decide which other members they want to interact with based on an established understanding or agreement. Member groups can compete with other member groups.

DDM ORCHESTRATION & AUTHORIZATION PROTOTYPE

AI flow Data/Algorithm

Datapace

Dashboard

Sell stream

Buy stream

Smart contract

Access control

AI

My wallet

manu@gmail.com
0 DPC

AI

+ Add Algorithm

+ Add Dataset

+ Execute

Algorithms

	Name	Type	Price	
<input type="checkbox"/>	MyMetaAlgo >	Algorithm	0.45 DPC	<div><div></div><div></div></div>

< Previous

1234

Next >

Datasets

	Name	Type	Price	
<input type="checkbox"/>	publicManu2Stream >	Dataset	0.67 DPC	
<input type="checkbox"/>	dsfsdfds >	Dataset	0.44 DPC	<div><div></div><div></div></div>
<input type="checkbox"/>	MyMetaData >	Dataset	0.44 DPC	<div><div></div><div></div></div>

Choose DataSet

Jobs Queue

ID	Mode	Algo	Data	State
5c780634afdade0001a0d3a5 >	federated	mySuperAlgo	pu-Dataset	executing
5c7a6f31afdade00013a4f31 >	centralized	locoo	proproprotetd	executing

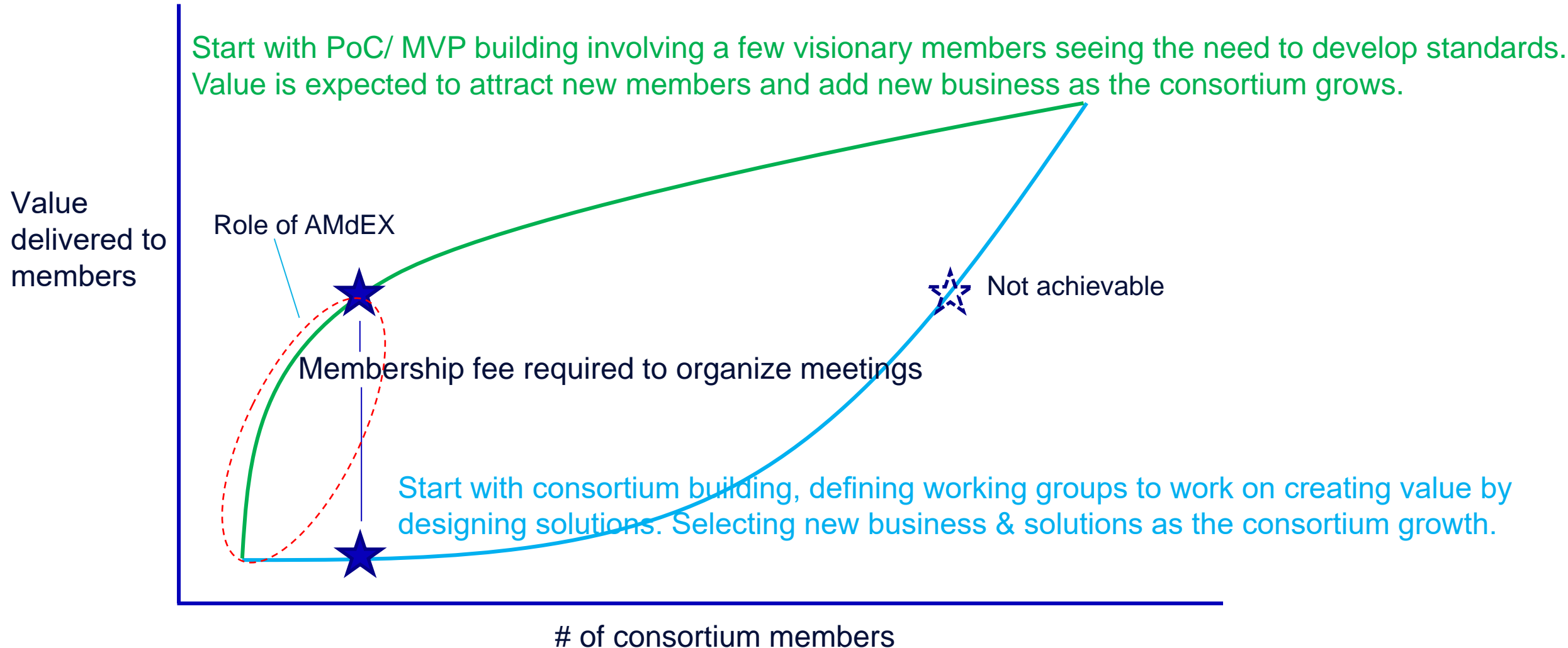
Choose Algorithm

Choose DataSet

Execute Training, Model creation, Prediction

Log tracking, execution tracking

CONSORTIUM SCALING – OVERCOMING THE CATCH-22

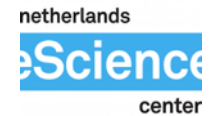
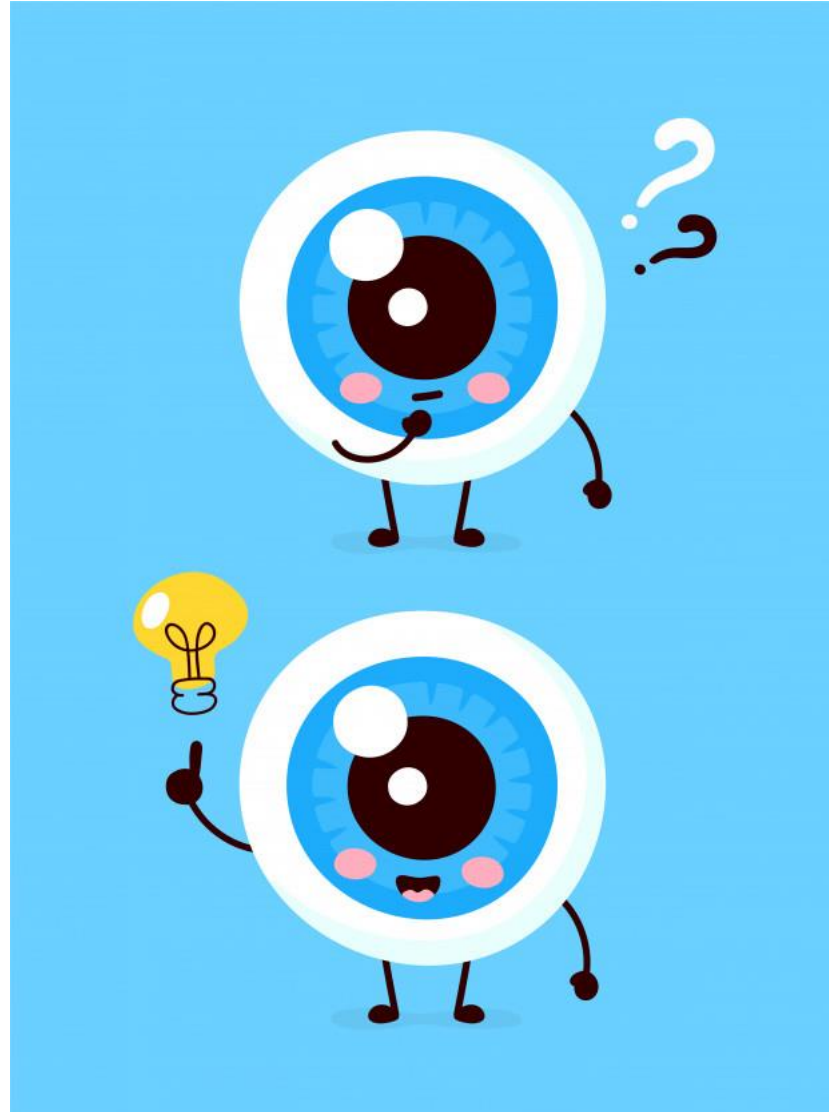


QUESTIONS



SSPDDP

SecConNet



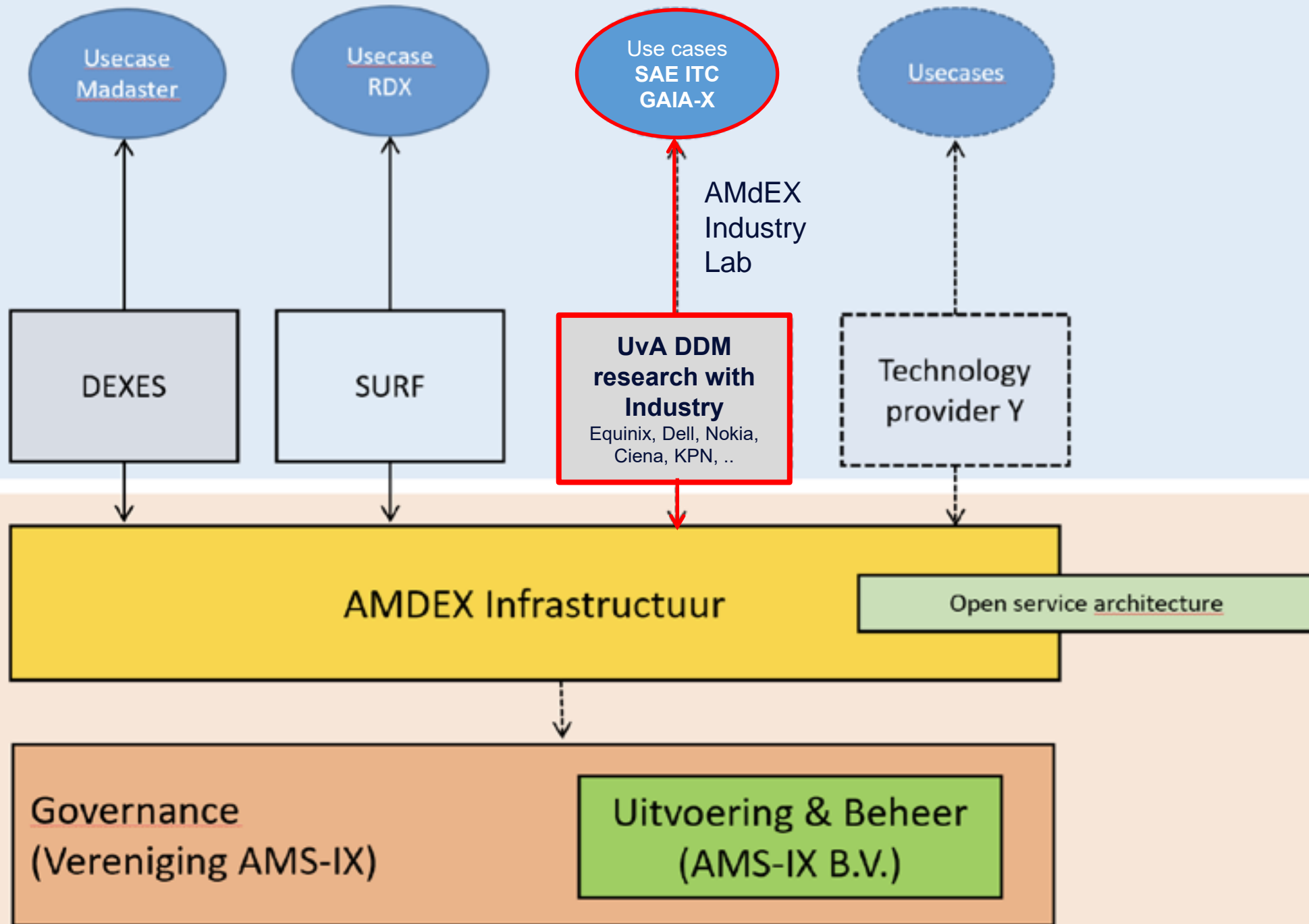
Marktplaatsen

AMDEX Fieldlab

Marktplaats services

Open exchange infrastructuur

Exchange operatie



UVA RESEARCH CONTRIBUTIONS

Research Area's



Consortia development

Identify 'why' (use-case benefits), rulemaking, organize trust, policy scheme's



Normative reasoning

Implement trust via automated compliance monitoring, request assessment, enforcement using legal concepts



'Data Space' Infrastructure Development

Data sharing architectures & archetypes for streaming, transactional, historic data, AI development (DDM)



Data Exchange Systems

Applied Industrial Research Lab using Future Internet concepts to create software definable global data exchanges

