

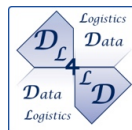
# Training AI models using Digital Data Marketplaces



Leon Gommans

Science Officer Air France KLM Group IT Technology Office - R&D


Guest Researcher, University of Amsterdam – Systems & Network Engineering Lab



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



# BUSINESS CONTEXT



Companies increasingly understand how to apply AI technologies to extract business value from data.

The more data the better: algorithm quality depends on data quantity and quality  
**Knowledge** how to translate such data into reliable algorithms is **competitive**

Companies are reluctant to share data when considering involved risk.

Emerging platform dominance: *“While creating real value for users, these companies are also capturing a **disproportionate and expanding share of the value**, and that ‘s shaping our collective economic future”.* \*

**Sharing data across companies increases the potential of creating business value no single organization can create on its own.**

# DATA IS INCREASINGLY CONSIDERED AN ASSET

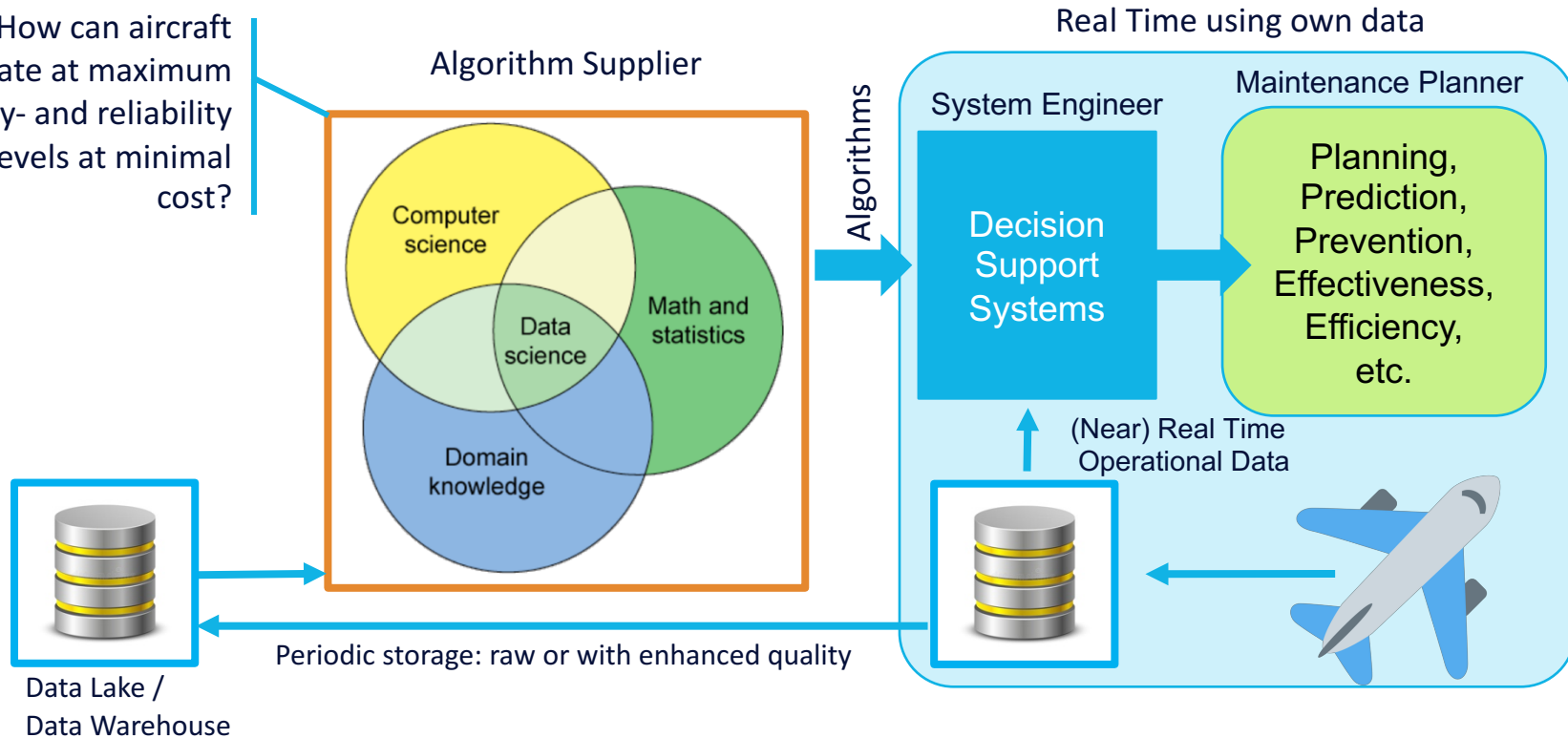
Considering value exchange and involved risk raises the main research question:

*How can (big) data assets be shared between data suppliers and algorithms developers in*

- 1) A fair and economic way,*
- 2) whilst providing adequate means to reduce risk?*

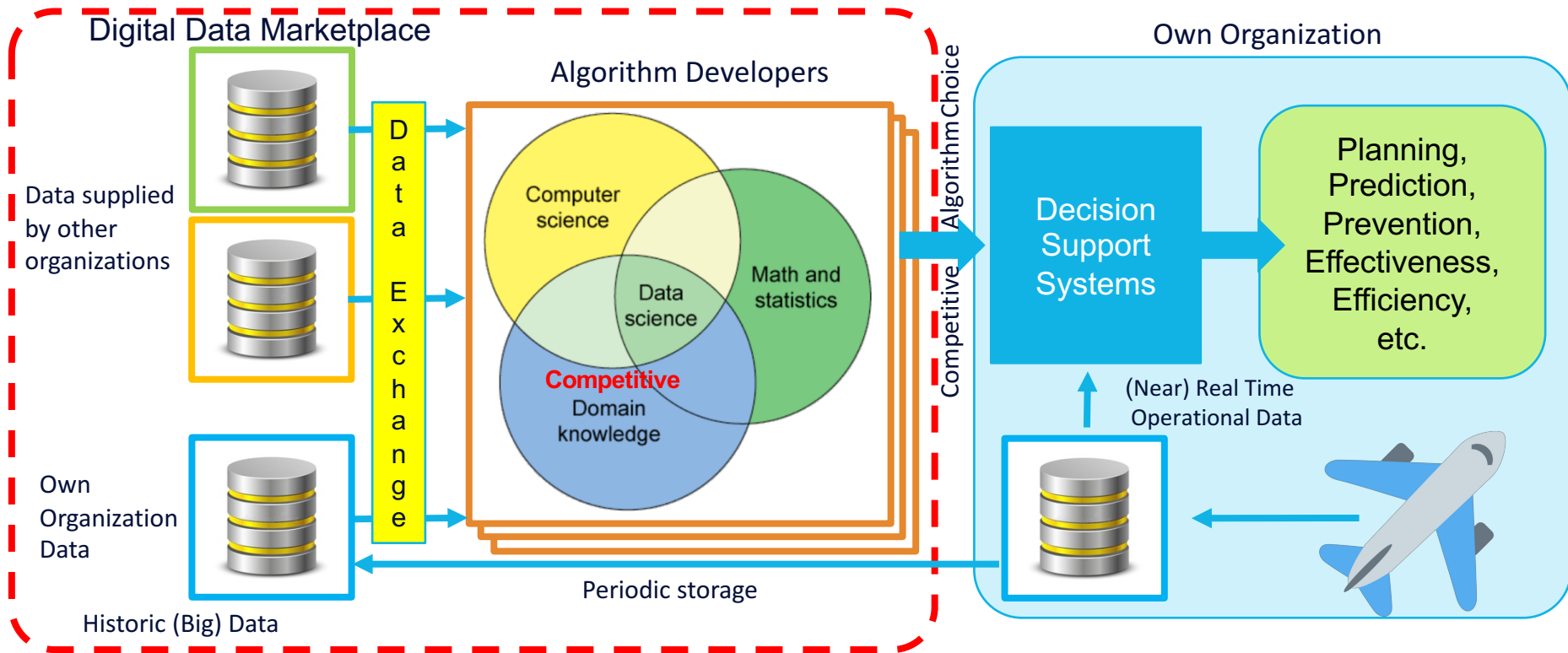
# CURRENT ALGORITHM DEVELOPMENT CONTEXT

How can aircraft operate at maximum safety- and reliability levels at minimal cost?



# RESEARCH CONTEXT

ARRANGE ADDITIONAL DATA TO IMPROVE ALGORITHM QUALITY & INNOVATION



# B2B DATA SHARING APPROACHES

AN EU STUDY (everis, JAN 2018)

Case studies

Approaches to B2B data sharing



## Five different approaches to B2B data sharing

### 1 DATA MONETISATION



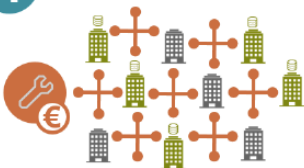
### 2 DATA MARKETPLACES



### 3 INDUSTRIAL DATA PLATFORMS



### 4 TECHNICAL ENABLERS



### 5 OPEN DATA



## Open vs Closed

### DATA MARKETPLACES



- ✓ Trusted intermediary between data suppliers and data users
- ✓ Data suppliers sell their data to interested data users
- ✓ Revenue is generated from each data transaction



### INDUSTRIAL DATA PLATFORMS



- ✓ Strategic and collaborative partnerships
- ✓ Mutual benefits for all parties
- ✓ Data shared (for free) in a closed, exclusive and secure environment
- ✓ Develop new or improved products and/or services
- ✓ Enhance internal performance

AIRBUS



Considering our research context:

difference with 2

Governance by a **membership organization**

difference with 3:

Data is stored **outside** data platforms to **allow multiple platforms to access & use same data**

# DATA SHARING CHALLENGES

WHEN TRAINING MODELS WITH AS MUCH DATA AS POSSIBLE

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

Many implications need to be considered:

## 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



# OVERCOMMING CHALLENGES

## ELEMENTS TO ORGANIZE TRUST AS MEANS TO REDUCE RISK



### COMMON BENEFIT

Define and agree common benefit no single organization can achieve on its own.



### GROUP RULES

Define consortium rules considering data use, access and benefit sharing



### ORGANIZE TRUST

Organize power and trust **as a means to reduce risk** for participating members



### IMPLEMENT INFRASTRUCTURE

Research operationalization of **Digital Data Marketplace & Data Exchange** concepts



# DEFINE AND AGREE COMMON BENEFIT



**Example: enable data sharing to improve quality of AI/ML innovations**

- Understand need: the more data the better
- Expect: capability that will help transform the MRO business in the digital era.

**Innovations that will improve air safety, passenger experience and additional cost reductions by:**

- avoiding unplanned maintenance
- increasing maintenance planning flexibility
- moving from fixed interval planning to maintenance when indicated
- less network disruptions by avoiding 'Aircraft On Ground' situations

# CONSORTIUM MEMBERSHIP RULES: WHAT KIND OF RULES DO WE NEED?



**Trust is considered as a means to reduce risk**

Defining consortium membership rules is a starting point

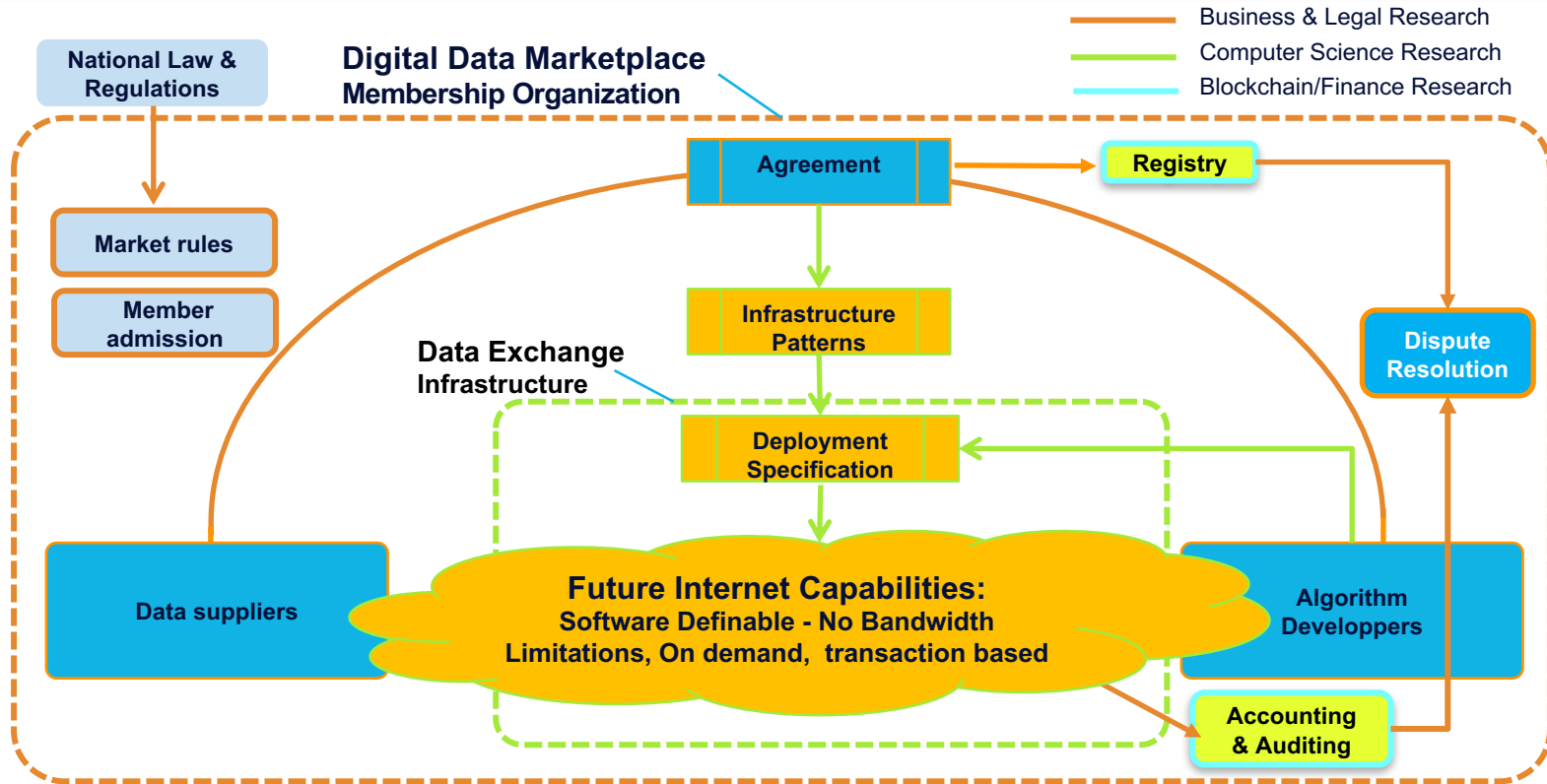
**Legal research topic's for discussion:**

- Data asset ownership
- Data access & usage
- Liability of owner & user
- Non-compliant behavior
- Market rules
- Purpose binding



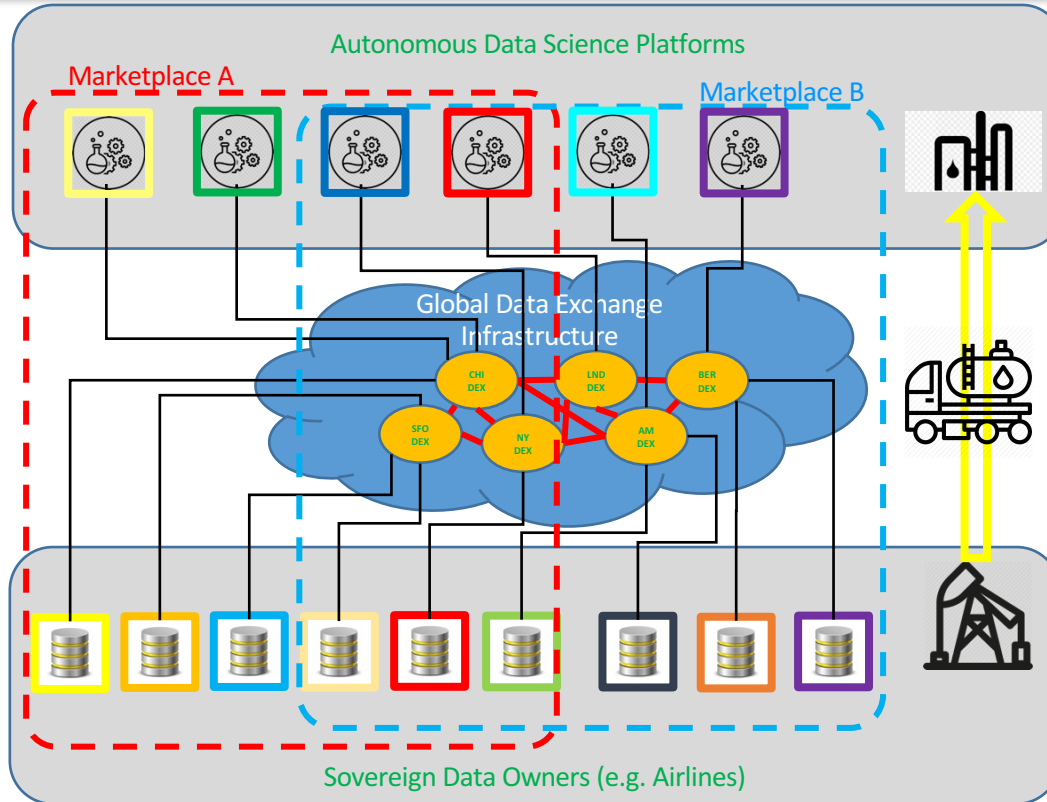
# DIGITAL DATA MARKETPLACE CONCEPT:

## COMBINED BUSINESS, LEGAL AND COMPUTER SCIENCE RESEARCH



# DATA EXCHANGE CONCEPT

## ENVISAGED GLOBAL EXCHANGE INFRASTRUCTURE



amsterdam  
economic  
board

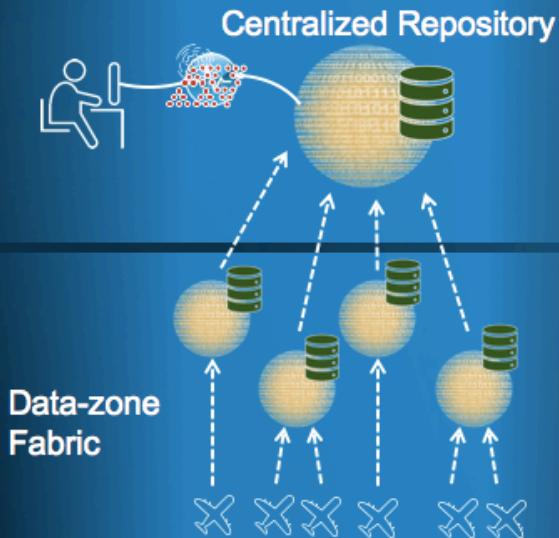


# PROCESSING & STORAGE: TRAINING STRATEGIES

## CENTRALIZED VS FEDERATED ANALYTICS

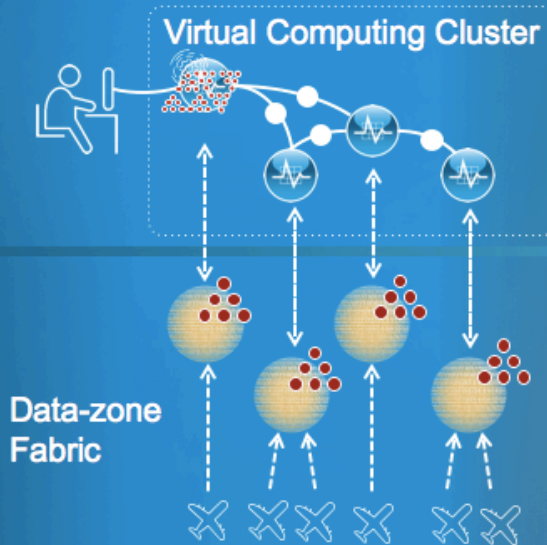
### Centralized

Raw data transferred from dispersed data zones to a central repository for analysis

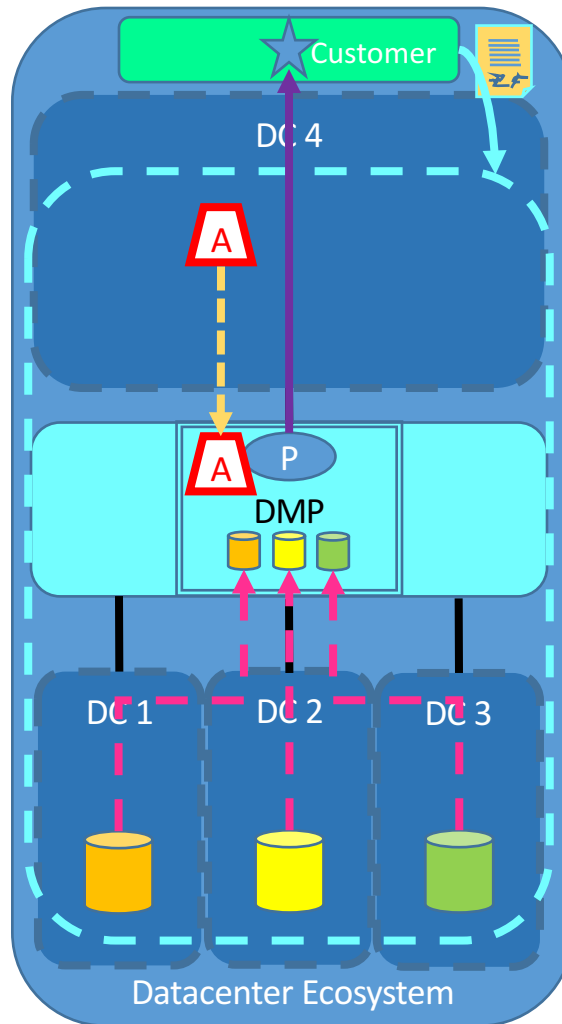













### Federated

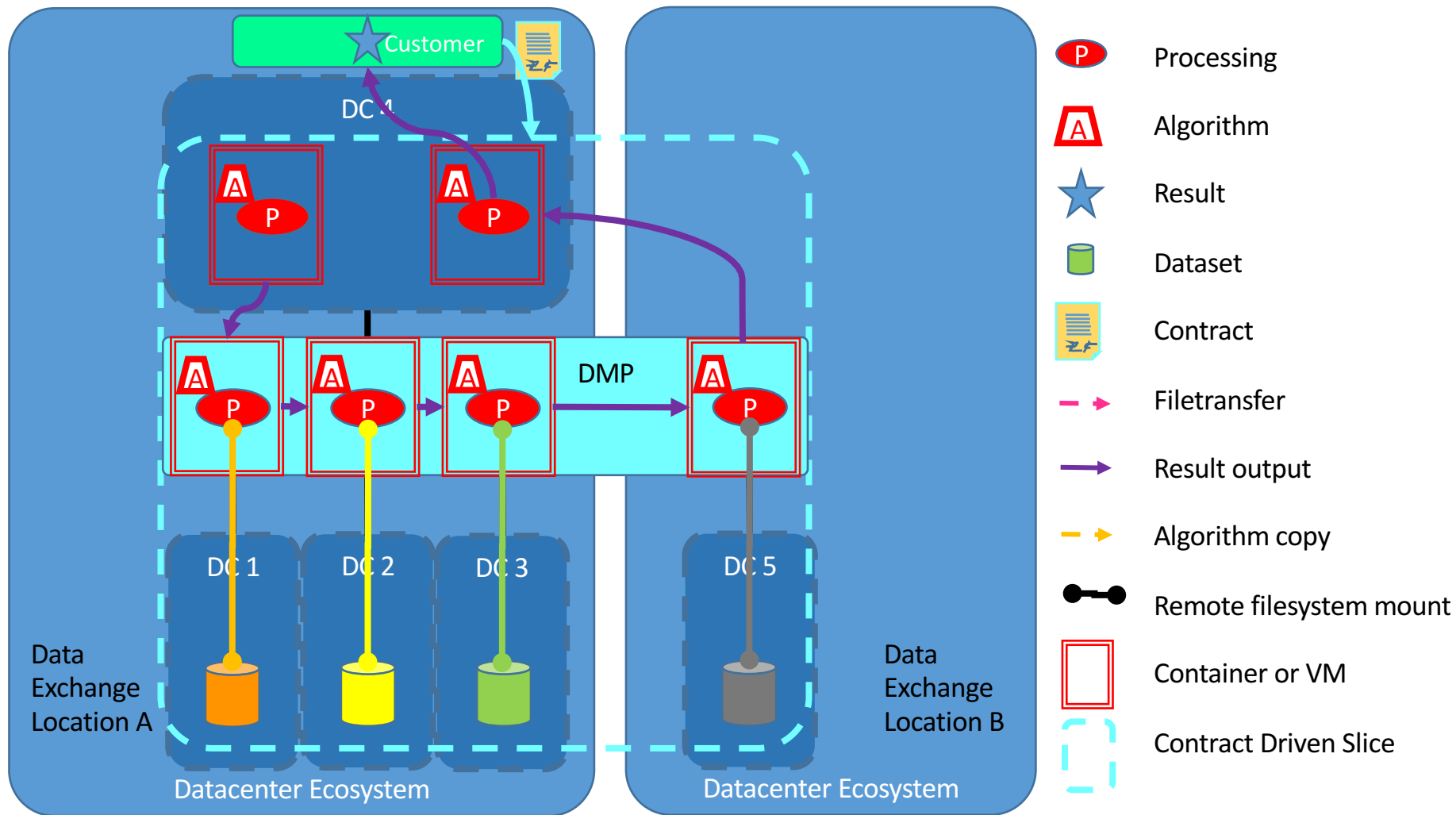
Raw data stays in place. Model trained through orchestration of local (at each data zone) and global computations



DMP provides neutral processing capabilities, dissolving after execution.



-  Processing
-  Algorithm
-  Result
-  Dataset
-  Contract
-  Filetransfer
-  Result output
-  Algorithm copy
-  Remote filesystem mount
-  Container or VM
-  Contract Driven Slice



- P** Processing
- A** Algorithm
- ★ Result
- Cylinder Dataset
- Document Contract
- - -> Filetransfer
- > Result output
- Algorithm copy
- Remote filesystem mount
- Red box Container or VM
- Dashed blue box Contract Driven Slice

# SUMMARY



Enterprises join a membership organization to achieve a common goal *no single enterprise can achieve on its own*



Membership rules are defined by rulemaking & standards processes, subsequently execution, enforcement and judgement is organized by membership organization as *a means to reduce risk.*



Members arrange data sharing and processing via *agreements deployed in an infrastructure*, provided by a secure digital market place owned by its members.



Members *achieve common benefits in a transparent way.* Members trust its operation based on use of accounting & auditing mechanisms, relying on market dispute resolution mechanisms.