Improving AI models using Digital Data Marketplaces

Leon Gommans, PhD

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 innocvation 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.

* M. Iansiti, K.R. Lakhani, *Managing our hub economy,* Harvard Business Review, pg. 85-92, Sep/Oct 2017

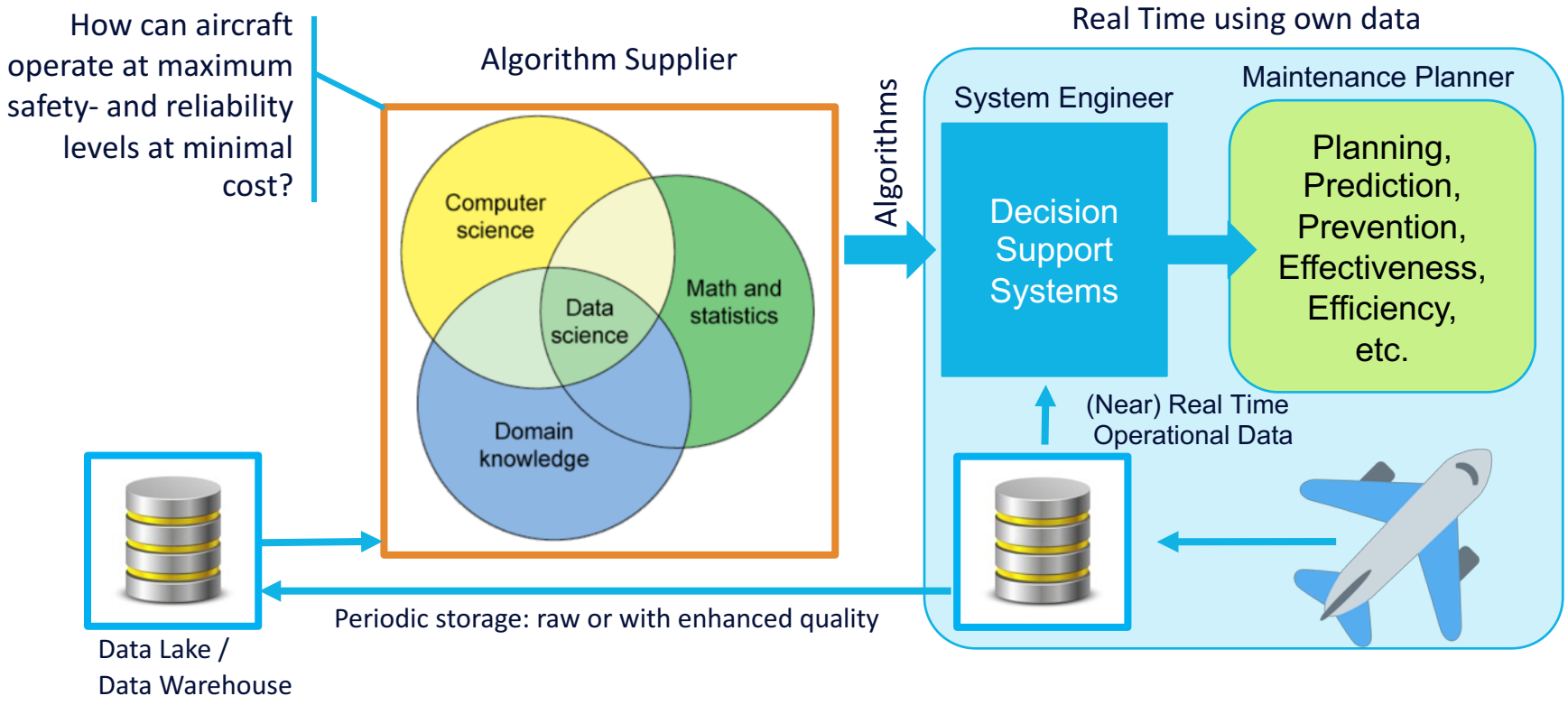
DATA IS INCREASINGLY CONSIDERED AN ASSET

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

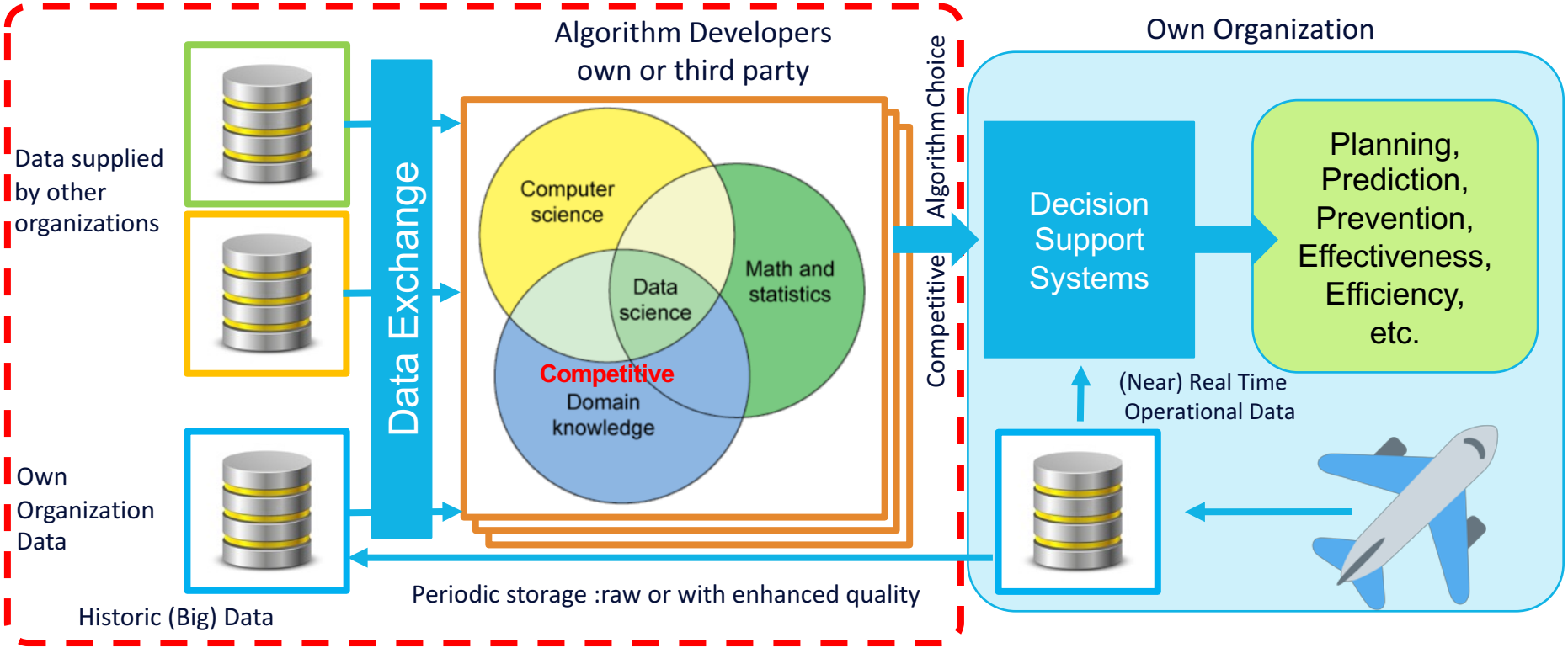
1) ***A fair and economic way,***
2) ***whilst providing adequate means to reduce risk?***

CURRENT ALGORITHM DEVELOPMENT CONTEXT



RESEARCH CONTEXT

ARRANGE ADDITIONAL DATA TO IMPROVE ALGORITHM QUALITY & INNOVATION



B2B DATA SHARING APPROACHES

AN EU STUDY BY EVERIS JAN 2018

Case studies | Approaches to B2B data sharing

European Commission

everis an NTT DATA Company

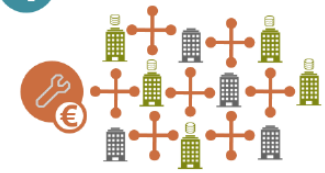
Five different approaches to B2B data sharing

1 DATA MONETISATION

2 DATA MARKETPLACES

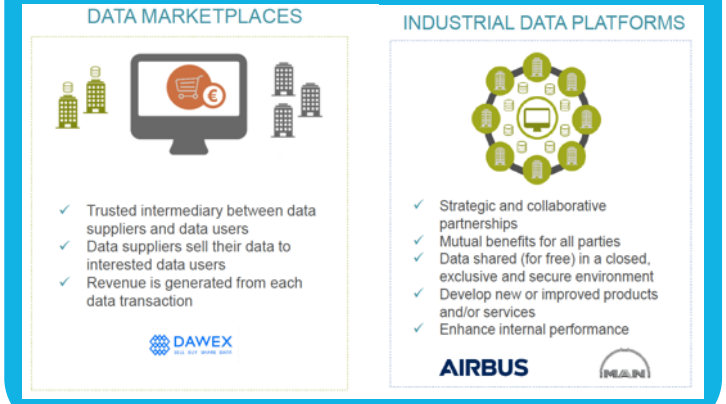
3 INDUSTRIAL DATA PLATFORMS

4 TECHNICAL ENABLERS



5 OPEN DATA

Open vs Closed



Difference with Data Marketplaces:

Governance by a **membership organization**

Difference with Industrial Data Platforms:

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

Contracts determine access / use

Market rules arrange pre-contractual elements

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	Legal level	Data level
Value	Ownership	Processing
Cost	Access	Storage
Benefits	Usage	Management
Agreements	Compliancy	Transport
Exchange	Liability	Transform
Trade	Market Rules	Security

Worldwide Scale

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

EXCHANGEWELL™
A Program of SAE ITC

INTRODUCTION

- Organized by SAE ITC, **ExchangeWell** brings data owners and algorithm developers together in a digital data marketplace that provides the required trust for mutual engagement.
- It enables members to share their data assets in a **fair and economic way** whilst providing an adequate means to **reduce risk**.
- Sharing data enables **digital transformation of the industry** and **business value creation.**

Objective: Help answer key question:

- Will ExchangeWell as proposed provide value to our industry?

COMMON BENEFIT

GROUP RULES

ORGANIZE TRUST

IMPLEMENT INFRASTRUCTURE

EXCHANGEWELL™

A Program of SAE ITC

A consortium program to provide the means for industry leaders to access industry experts, develop practical experience from pilots, collaborate on pre-competitive research and establish a strategic path forward to effectively implement data management strategies which positively impact and benefit industry.

SAE ITC®

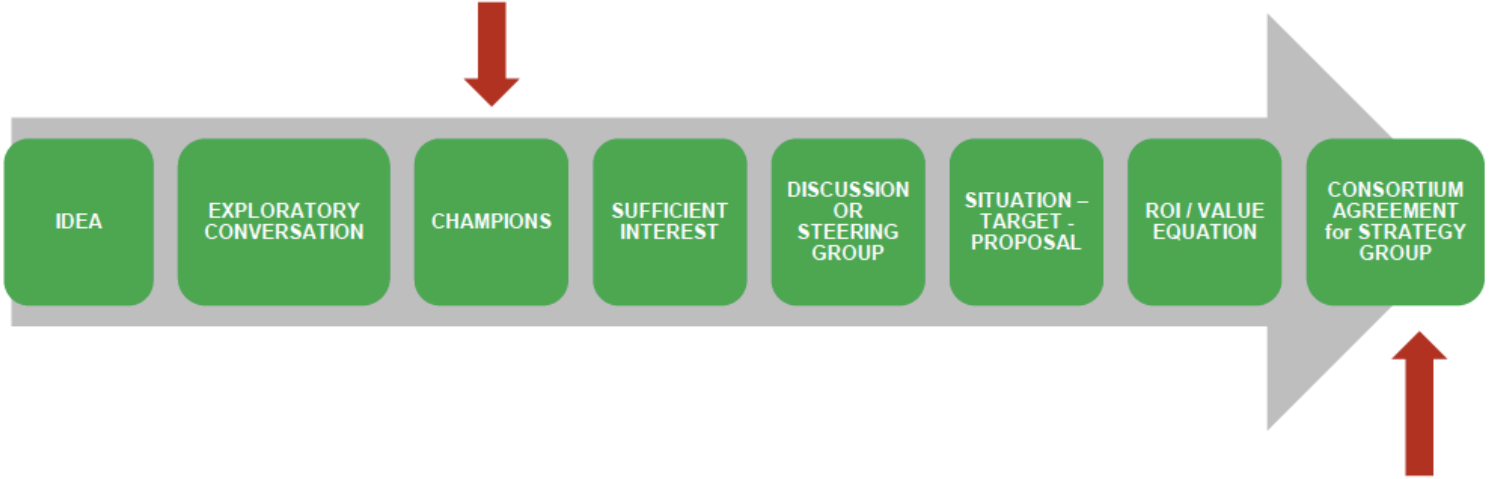
An SAE International Affiliate

Collaborative Innovation.
Trusted Implementation.

SAE ITC

CONSORTIUM MATURATION PATH

We're here!
ExchangeWell



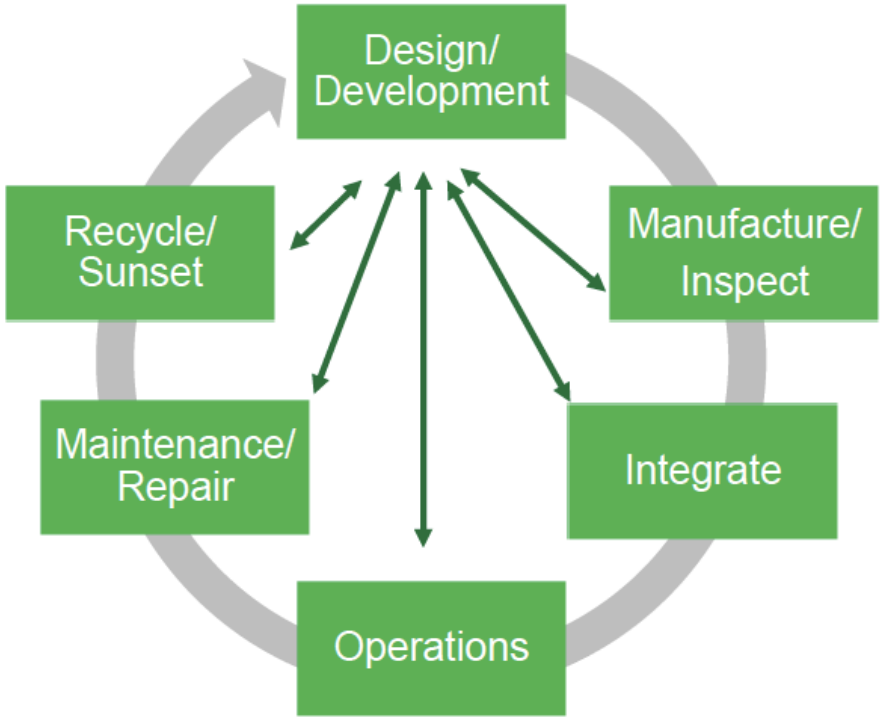
We're here!
ARINC IA

STAKEHOLDERS

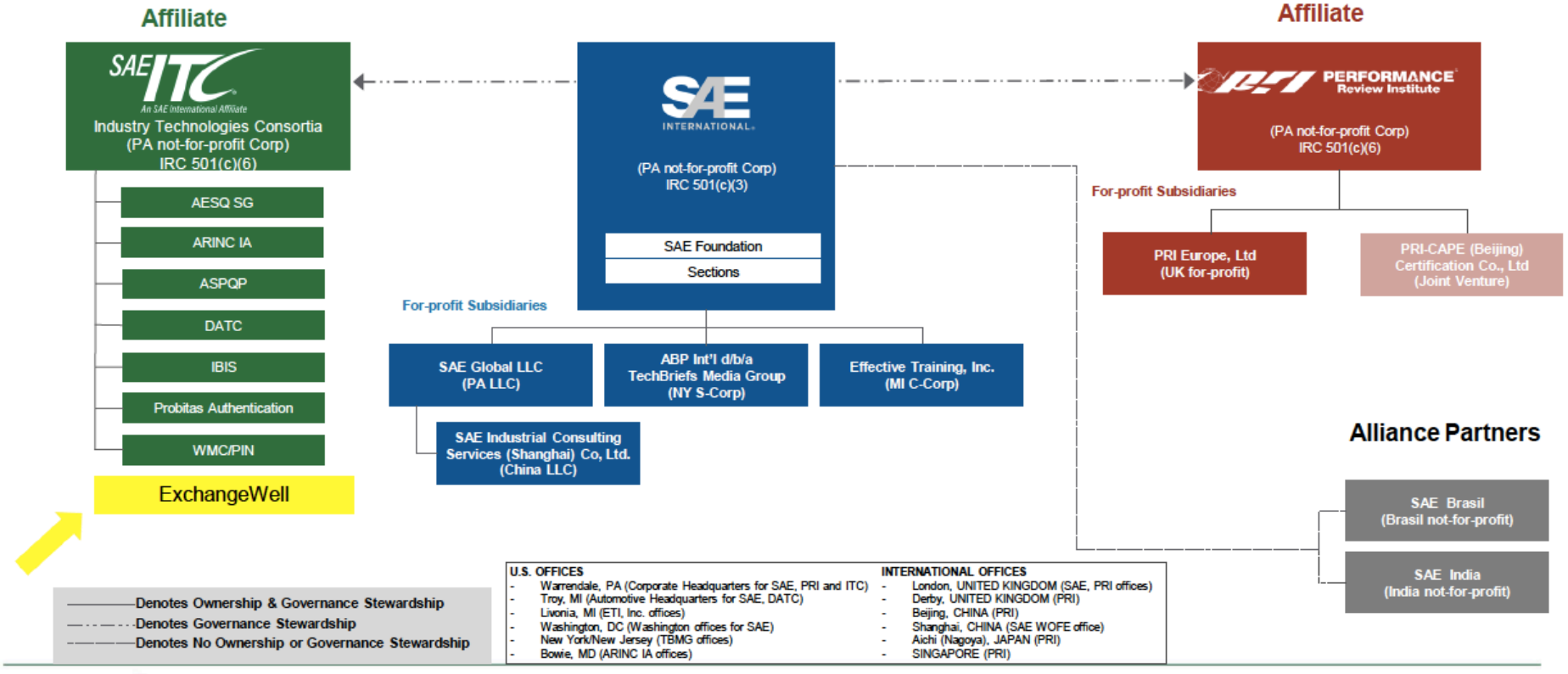
INTEGRATED PRODUCT CYCLE

Stakeholders

- Regulatory
- Airline/ Operator
- Airframer/ Integrator
- OEM
- Sub Assembly Manufacturer
- Distributor
- Component/ Part Manufacturer
- Standards Organization
- Industry Review Body
- Auditor/ Mandated Body
- SAE ITC
- Registrar
- Maintenance
- Training Provider
- IT System and Software Tools Provider
- Data Aggregators and Analyzers
- Insurers
- Legal
- Access Authorizing Agent
- Research/ Academics



SAE ORGANIZATION



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

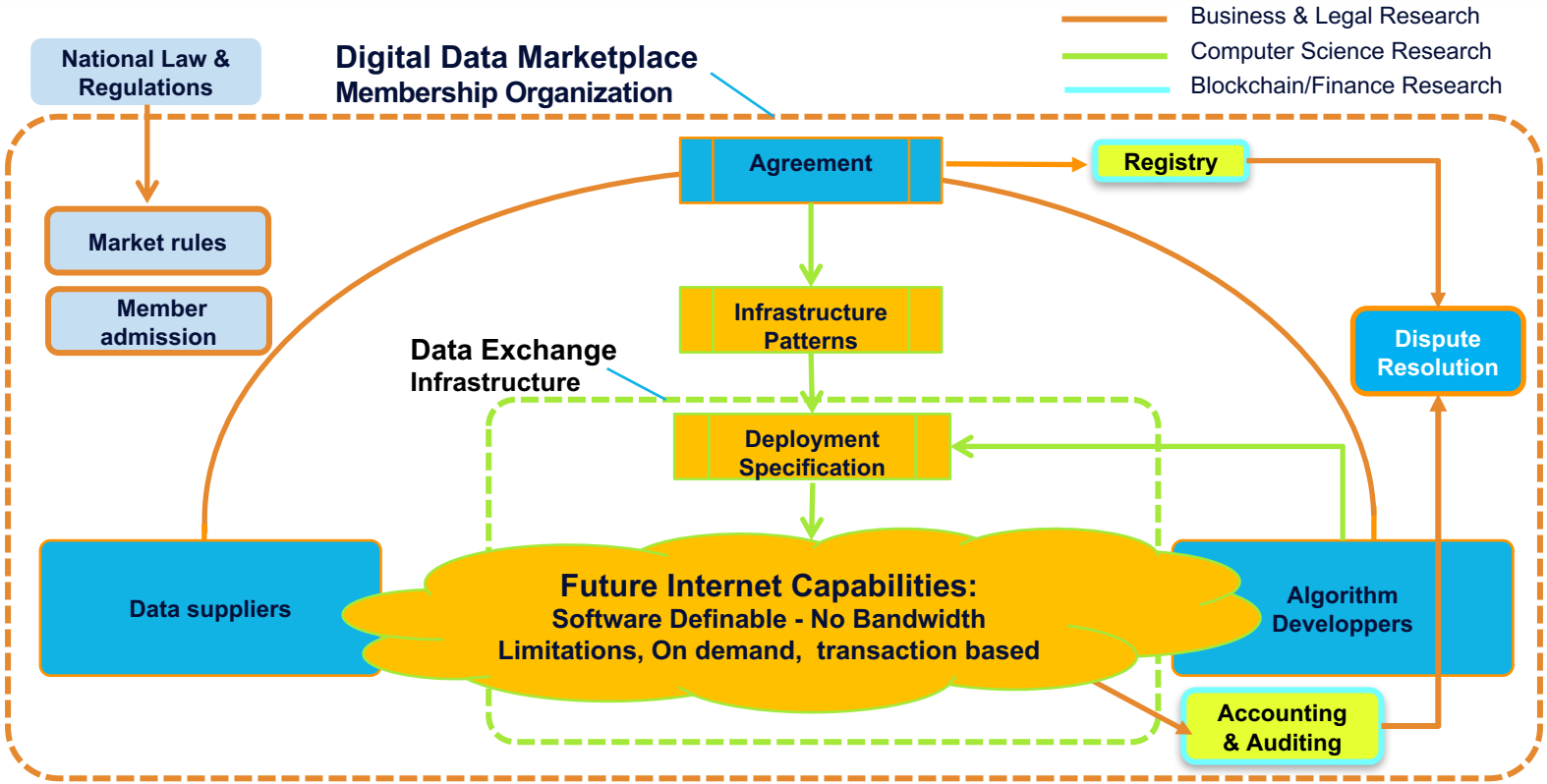
Legal research topic's for discussion:

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

EXCHANGEWELL™
A Program of SAE ITC

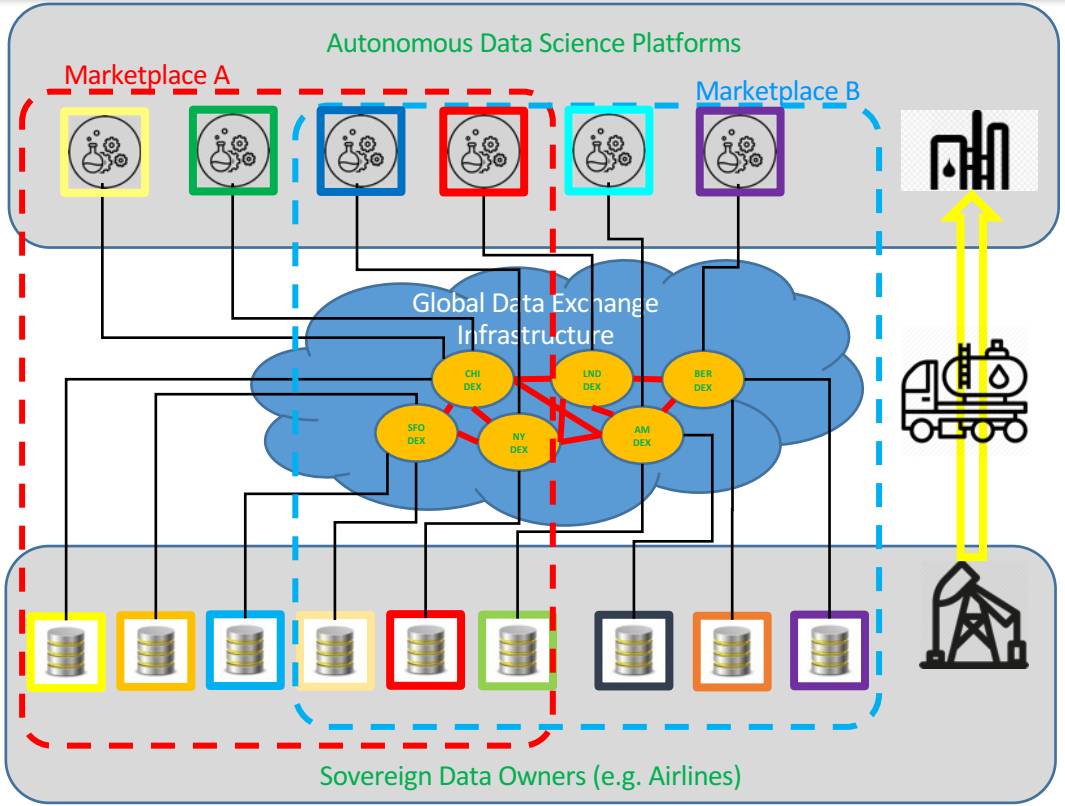
DIGITAL DATA MARKETPLACE CONCEPT:

COMBINED BUSINESS, LEGAL AND COMPUTER SCIENCE RESEARCH



DATA EXCHANGE CONCEPT

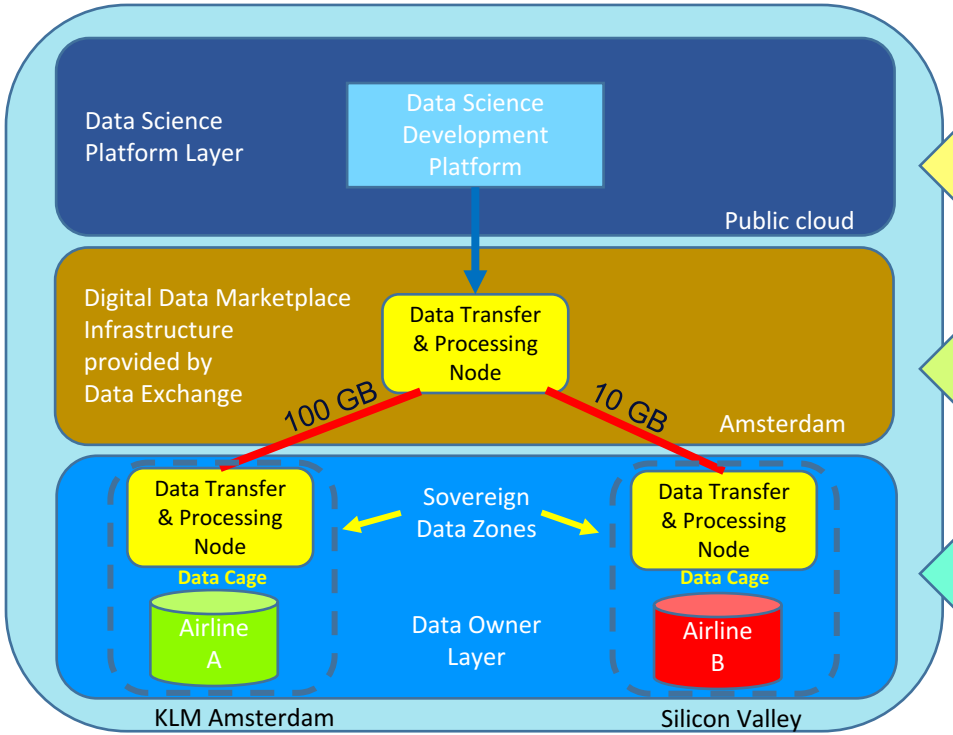
ENVISAGED GLOBAL EXCHANGE INFRASTRUCTURE



amsterdam economic board

AMdEX
THE DATA HYPERMARKET

RESEARCHING EXCHANGE ARCHITECTURES



Trust Modelling:

What is the optimal infrastructure archetype, describing storage and processing locations and their relationships, which best suit member requirements when considering risk?

See CIENA booth 2847 and demo

Processing Models:

What are the implications of distributing data processing across membership organization owned infrastructures in terms of achievable model accuracy and processing performance using federated/distributed models vs centralized models

Marketplace Reference Architecture:

What constitutes a marketplace? Researching needed functions, personas, flows, credentials, contracts & rules, conflict resolution, and much more …

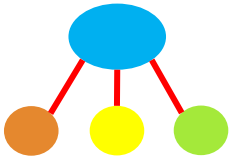
RESEARCHING PHYSICAL IMPLEMENTATION

INVOLVING BOTH RESEARCH AND IT INDUSTRY

GLOBAL RESEARCH INFRASTRUCTURES

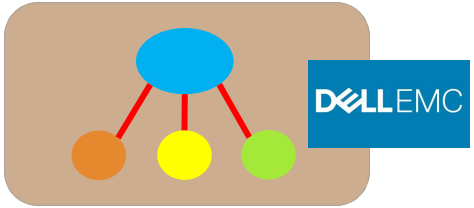
Data Sharing Infrastructure Model Research using Future Internet capabilites

UNIVERSITEIT VAN AMSTERDAM



Research collaboration

GLOBAL DATACENTER INFRASTRUCTURES

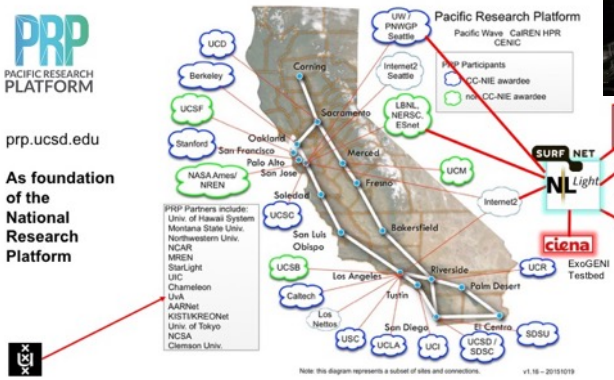


How to create a Global Digital Data Market Ecosystem via Data Exchanges

PRP PACIFIC RESEARCH PLATFORM

prp.ucsd.edu

As foundation of the National Research Platform



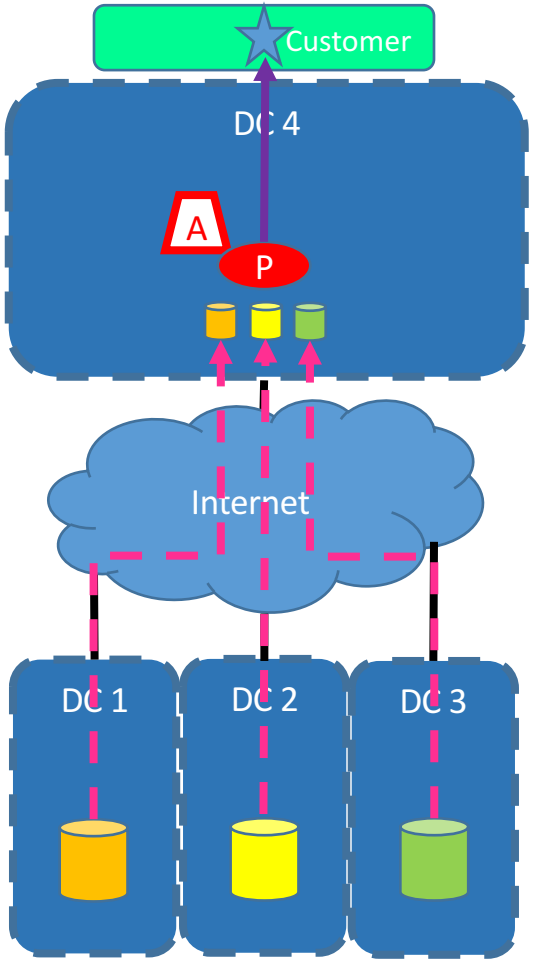
Data Transfer Node at KLM fieldlab with 100 gb/s link to enable SDMP research thanks to UvA, SURFnet and Ciena

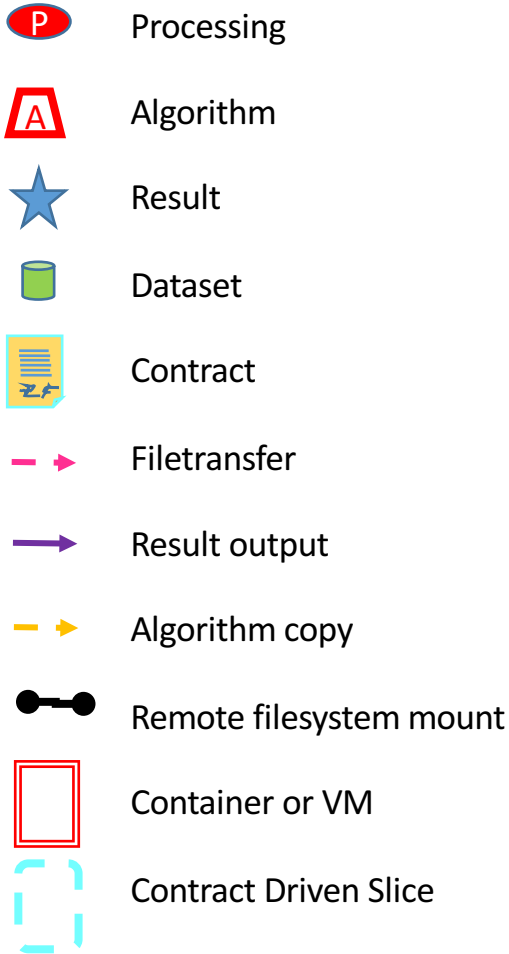
AM3 and AM4 Datacenters Amsterdam Science Park

SV10 Datacenter Silicon Valley

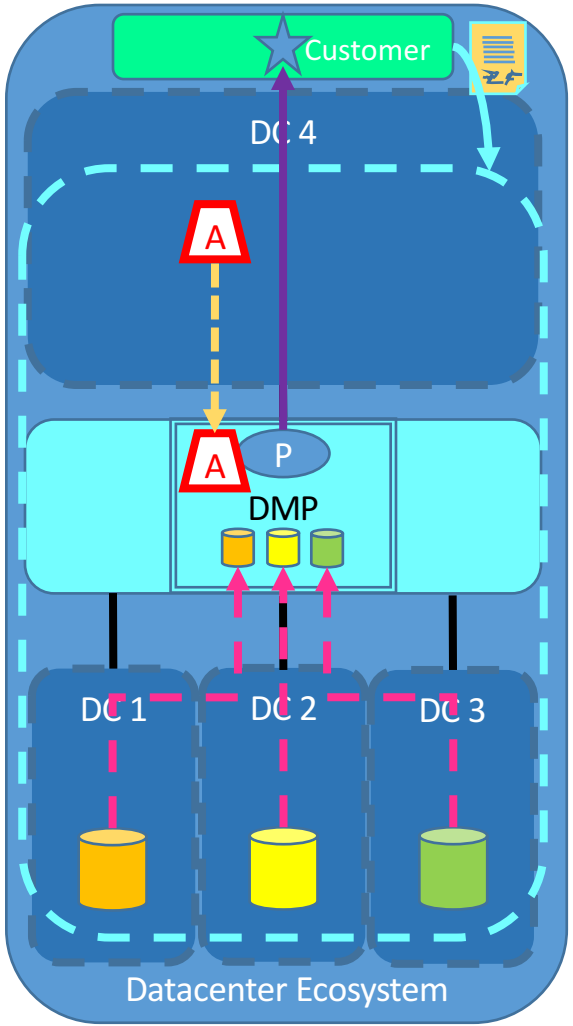
Traditional model

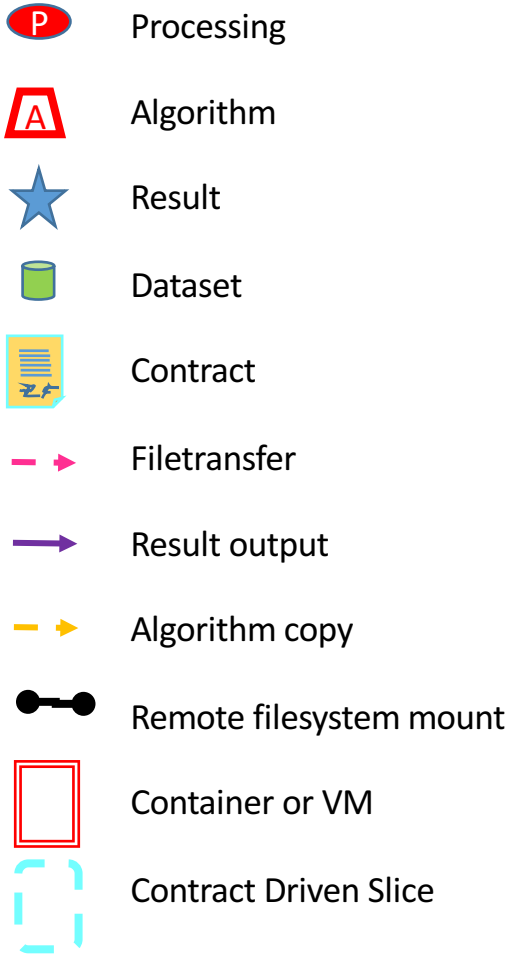
DC4 acts as platform:
1: creates potential competitive bottleneck / lock-in.
and
2: raises data owner concerns about risk

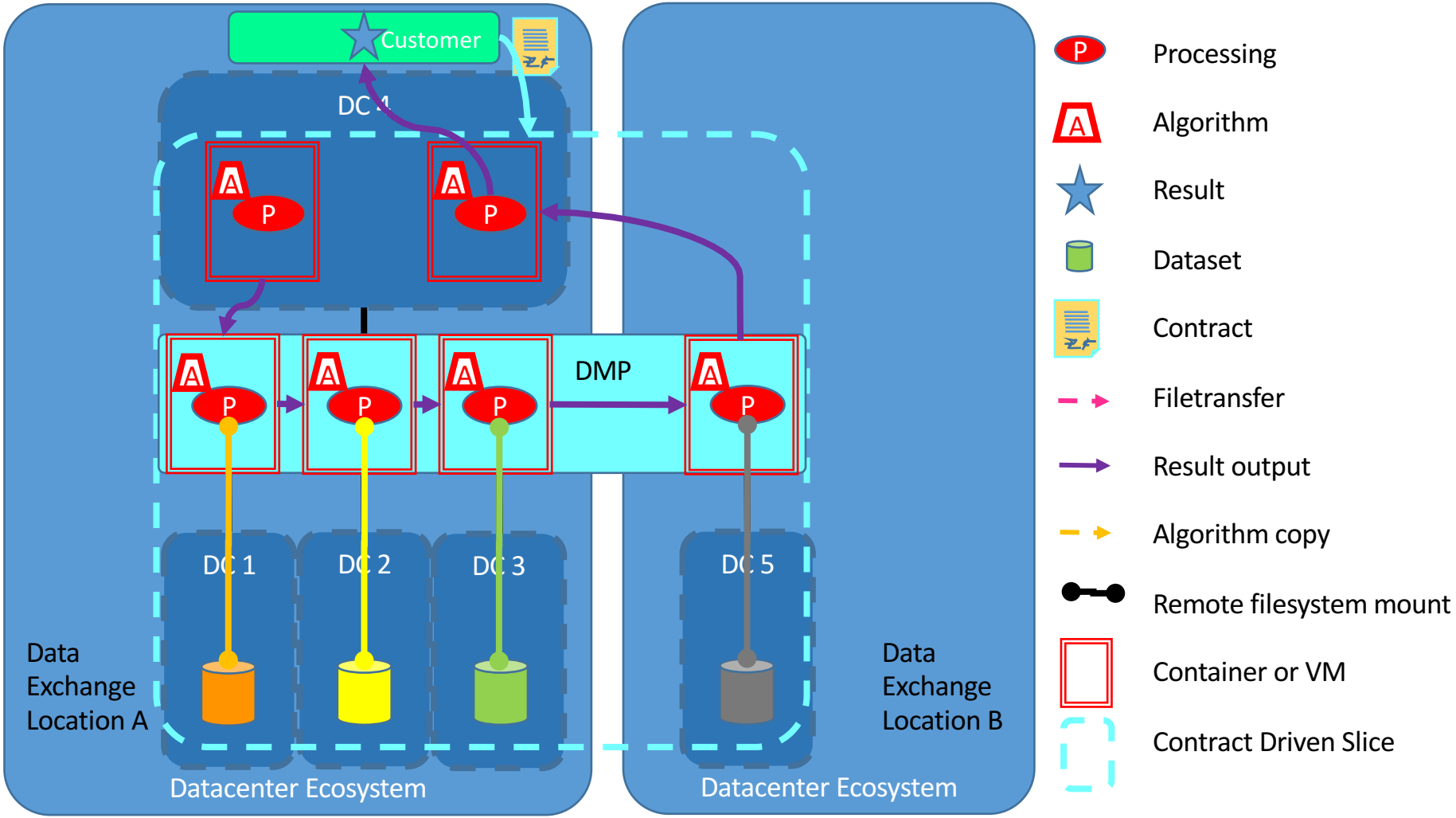




DMP provides neutral processing capabilities, which dissolves after Execution.






Customer
DC 4
A
P
A
P
A
P
A
P
A
P
DMP
DC 1
DC 2
DC 3
DC 5
Data
Exchange
Location A
Datacenter Ecosystem
Data
Exchange
Location B
Datacenter Ecosystem
Processing
Algorithm
Result
Dataset
Contract
Filetransfer
Result output
Algorithm copy
Remote filesystem mount
Container or VM
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.