EXCHANGEWELL GOVERNED DIGITAL DATA MARKETPLACE
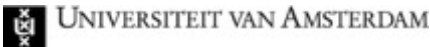
How to share data a trusted, fair and economic way supporting ML/AI development for PHM

Leon Gommans, Air France KLM Group & University of Amsterdam
George Saleh, Nokia Datapace

AIRFRANCE KLM

amsterdam economic **board**

AMdEX
THE DATA HYPERMARKET

CIMPL

ReMAP
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 · EQUINIX · DELL Technologies · NOKIA · ciena · EXCHANGEWELL A Program of SAE ITC

INTERNET HISTORY EUROPE

RESEARCH CONTEXT PROVIDED BY AMSTERDAM SCIENCE PARK

1988:
The Internet landed in Europe at CWI (CERN was first to connect to this place)

1994
The Amsterdam Internet Exchange was first housed by NIKHEF.
Grew into the largest Internet Exchange of the world interconnecting ISP's, clouds, web hosting,.. acting as 'market square'

University of Amsterdam
Faculty of Science
located right in the middle
of Amsterdam Science Park

Can we create Digital Data Marketplaces inside these 'market squares' ?

Digital Reality (3.0B$ >175 DCs)   Equinix (4.4 B$ >200 DCs)
Now > 60 datacenters emerged around the Amsterdam Internet Exchange, creating 12.500 jobs*

Data centers are neutral places housing equipment from multiple (cloud) providers in separate 'cages'

AIRFRANCE KLM

WHAT IS IT ABOUT?

A DIGITAL DATA MARKET PLACE:

- **Serves a common benefit no single organization can achieve on its own.**

- **Is created and governed by an industry consortium as a means to reduce risk, ensuring competition and fairness.**

- **Supply members advertise their assets, contracts arrange asset access and usage by other members.**

- **To prevent data asset exposure, members can use a consortium governed infrastructure to execute data science workflows**

- **Allows consortia to implement (digitally) enforceable contracts, whilst supporting dispute resolution by immutable logging.**

CONTEXT: AI DEVELOPMENT FOR AERONAUTICAL SYSTEMS

GENERATES MANY QUESTIONS: CREATING INITIATIVES TO ANSWER THEM

1950: "Can machines think?"
Alan Turing asked the question:
"Can a machine act as player in an imitation game?"
Alan Turing **"Computing Machinery and Intelligence",** Mind 49: 433-460, 1950.

Now "Can AI fly?"
Creg Hyslop, CTO Boeing, asked the question:
"How do we maintain the existing levels of safety with an AI-based system in the cockpit?"
Charlotte Jee, **"AI is set to change the aerospace industry - but won't be flying planes anytime soon",** MIT Technology Review, Sep 13th 2018.

Industry standards bodies are joining to consider the many questions around the role of AI in aeronautical systems:
- SAE International: G.34 Applied AI for Flight Critical Systems
- EUROCAE WG-114 Artificial Intelligence
- SAE ITC: ExchangeWell consortium.

Parties need to collaborate: OAMs, OEMs, MRO's, Operators, Regulatory bodies,..
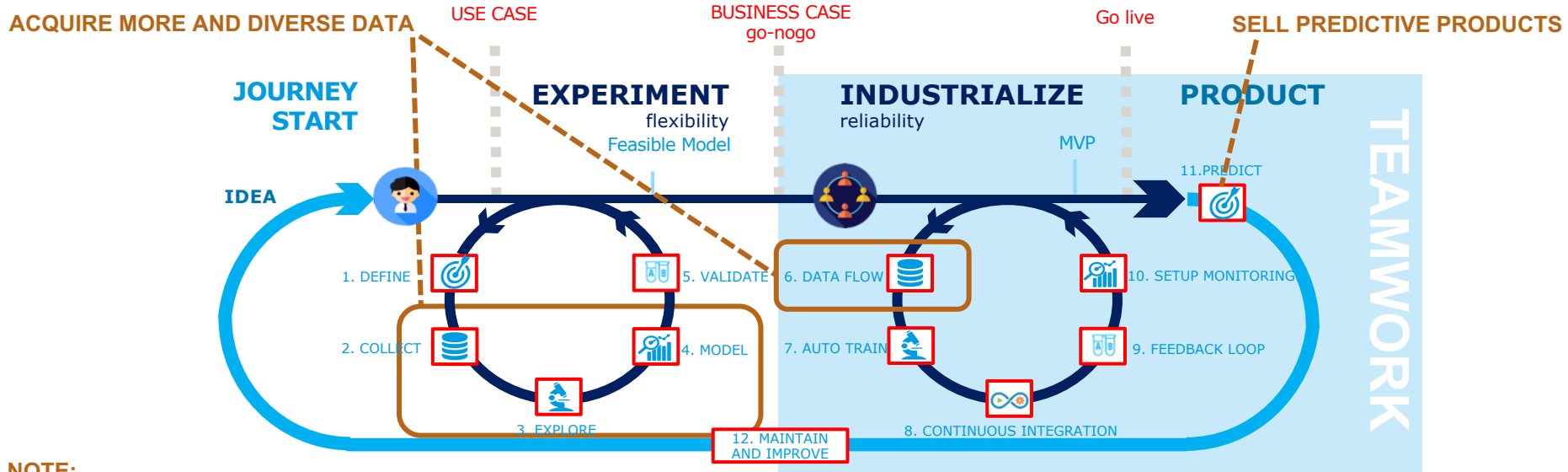**All have parts of the puzzle.**

AIRFRANCE KLM

RESEACHING DATA SHARING SOLUTIONS FOR AI DEVELOPMENT

A DIGITAL DATA MARKETPLACE GOVERNED BY A MEMBERSHIP CONSORTIUM
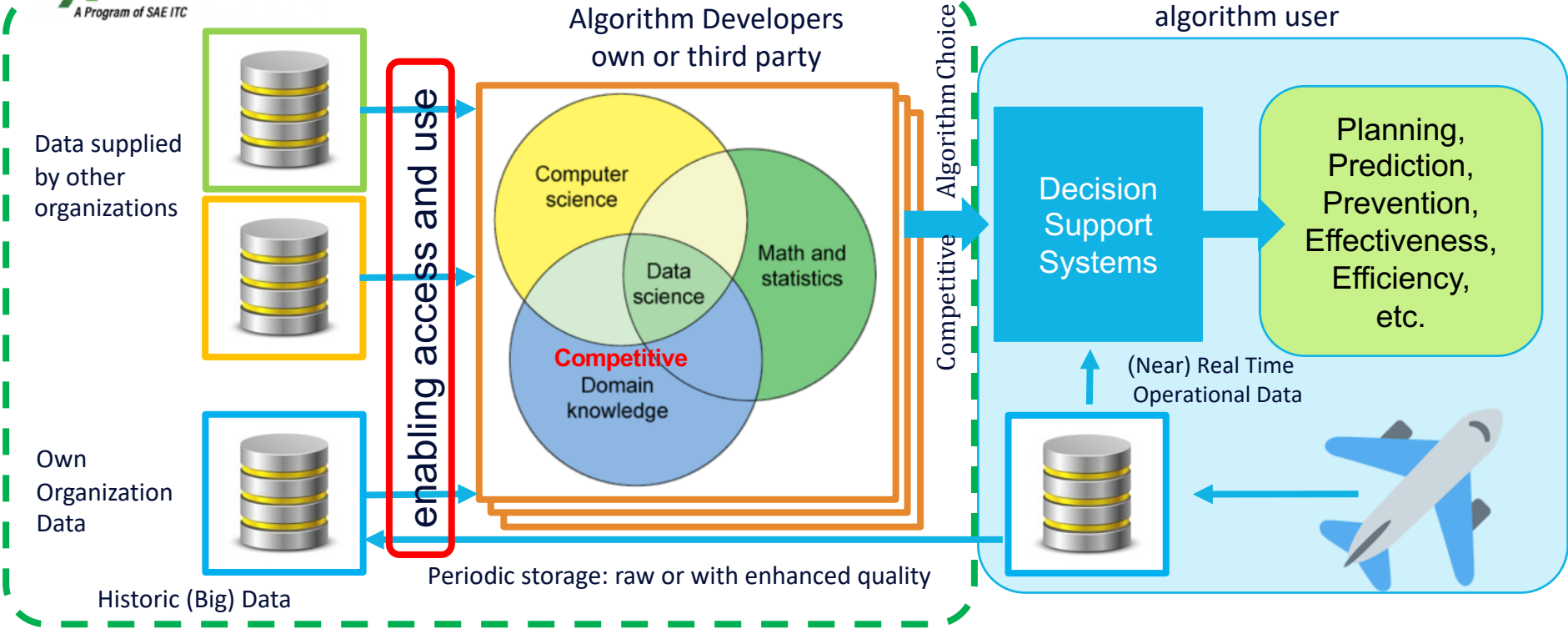
JOURNEY OF THE DATA SCIENTIST / ENGINEER

## ROLE OF THE DIGITAL DATA MARKETPLACE: MVP FOCUS IS ON INDUSTRIALIZE PHASE

ACQUIRE MORE AND DIVERSE DATA

USE CASE

BUSINESS CASE go-nogo

Go live

SELL PREDICTIVE PRODUCTS

JOURNEY START

EXPERIMENT
flexibility
Feasible Model

INDUSTRIALIZE
reliability

MVP

PRODUCT

TEAMWORK

IDEA

1. DEFINE
2. COLLECT
3. EXPLORE
4. MODEL
5. VALIDATE
6. DATA FLOW
7. AUTO TRAIN
8. CONTINUOUS INTEGRATION
9. FEEDBACK LOOP
10. SETUP MONITORING
11. PREDICT
12. MAINTAIN AND IMPROVE

NOTE:
KNOWLEDGE SHARING IN OTHER PHASES (4,7,9,10) MAY ALSO BE A GOALS OF COLLABORATION IN A MARKETPLACE COMMUNITY.

DEFINE (1-3)	EXPERIMENT (4-5)	INDUSTRIALIZE (6-10)	IMPROVE (11-12)	SHARE
What is the end purpose? How will it add value?	Collect and explore data, research data science model	Deliver the solution in production environment	Feedback analysis, A/B testing, and performance monitoring	
Use case, data scope and stakeholders identification	Prototype, go/nogo to production	Product GO live	New product releases, business insights	

AIRFRANCE KLM

USE-CASE – THE 747 BLEED AIR SYSTEM

DATA FROM KLM E&M SPLIT ACROSS THREE PLACES (IN STEAD OF ONE)

Imagine if data scientist can use historic data from 747 aircraft operated by multiple airlines.
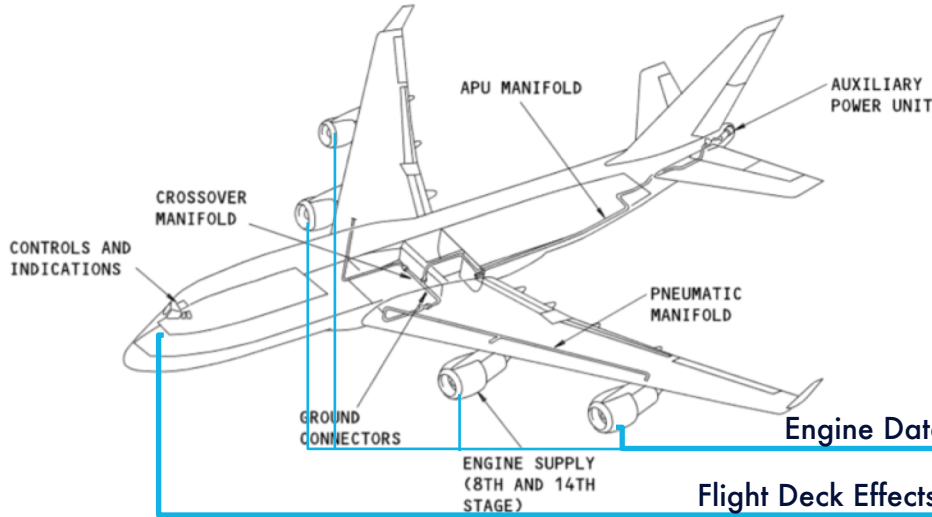
The Bleed Air System regulates pressure and temperature of air from a turbine engine needed by other aircraft systems taking care of:

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

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

CONTROLS AND INDICATIONS

CROSSOVER MANIFOLD

APU MANIFOLD

AUXILIARY POWER UNIT

PNEUMATIC MANIFOLD

GROUND CONNECTORS

ENGINE SUPPLY (8TH AND 14TH STAGE)

Near Real Time Data

Engine Data

Flight Deck Effects

Decision Support System

Will a Flight Deck Event occur within the next 10 flights?

Data stored in 3 places: KLM, UvA, EQX-SV

Data Science

Computer science

Math and statistics

Data science

Domain knowledge

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

AIRFRANCE KLM
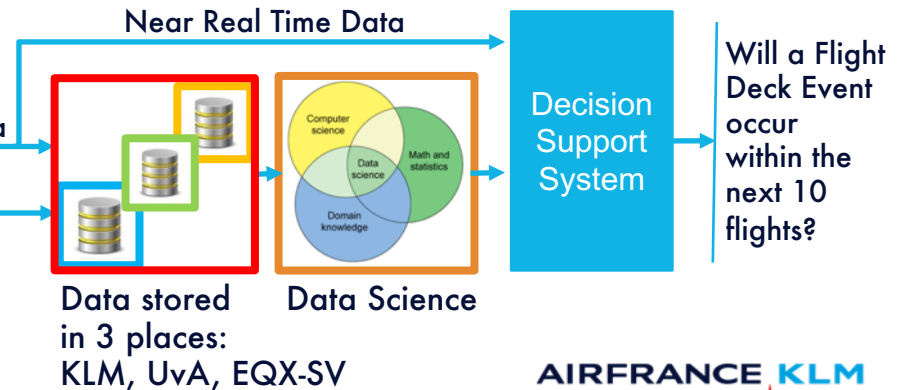
KEY CHALLENGE FOR DATA PROCESSING

WILL FEDERATED ANALYTICS PERFORM SIMILAR TO CENTRALIZED ANALYTICS?

Centralized

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

Centralized Repository

Data-zone Fabric
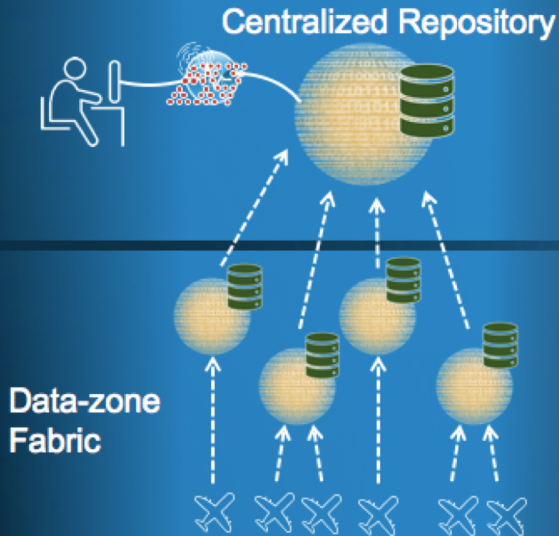
Federated

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

Virtual Computing Cluster

Data-zone Fabric

DELL Technologies
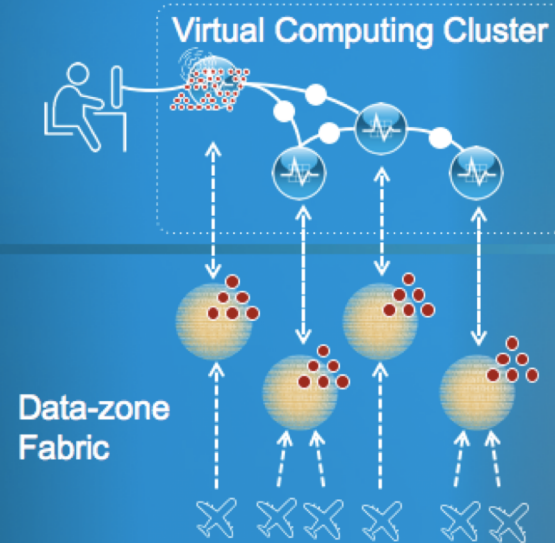
Predictive Performance and Data Transfer Comparisons

Predictive Performance	Model 1.0 Random Forest: Top 5 Features*			Model 2.0 Random Forest: Top 20 Features*			Baseline from previous study
	Centralized	Federated	% Diff.	Centralized	Federated	% Diff.	
Average Precision	0.19	0.19	0.0%	0.21	0.20	-4.8%	0.18
Area Under the ROC	0.59	0.60	1.7%	0.62	0.60	-3.2%	0.63
Bandwidth Use	~7.5GB	~2.8MB	-99.96%	~7.5GB	~2.9MB	-99.96%	--

→ Federated metrics within 5% of Centralized metrics in both models

→ Volume of data transferred in Federated training over 99% smaller than in Centralized training

*In terms of the feature importance metric from Random Forests

D∠LL Technologies

WHAT IS EXCHANGEWELL?

eXCHANGEWELL™
A Program of SAE ITC

Neutral forum where organizations represent their specific needs and interests on topics involving data exchange, interoperability, and security within highly regulated environments in order to identify the tools and services needed at an industry level to support the implementation of guidance, best practices and requirements captured in industry open forum consensus standards.

Consortium Benefits

Content. Access to standards and registries to support data exchange prototyping, testing, and development of future standards.

Contact. Creates an environment for building trusted relationships with industry experts in mobility and key IT technology providers.

Capabilities. Serves a common benefit no single organization can achieve on its own with experience in standardization, registrations and qualifications, algorithm development and validation- initially with standards, future with predictive maintenance.

Context. Understands use cases and positioned to help industry while reducing risk being governed by an industry membership organization

SAE ITC
An SAE International Affiliate
Collaborative Innovation. Trusted Implementation.

CONFIDENTIAL: NOT FOR DISTRIBUTION

10

EXCHANGEWELL SERVES 3 AREA'S

1 Standards data dissemination. — Technical Standards Data Interoperability →
- Reduce transcription error
- Improve change management
- Reduce redundancy

2 Tracking of qualified entities. — Registries →
- Improved transparency
- Efficient auditing
- Connecting standard parts and suppliers

3 Sharing of data and/or algorithms without centralization. — Digital Data Marketplace →
- Increased statistical relevance
- Improved algorithm performance
- De-centralize data

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

EXCHANGEWELL™
A Program of SAE ITC

Design/ Development

Recycle/ Sunset

Manufacture/ Inspect

Maintenance/ Repair

Integrate

Operations
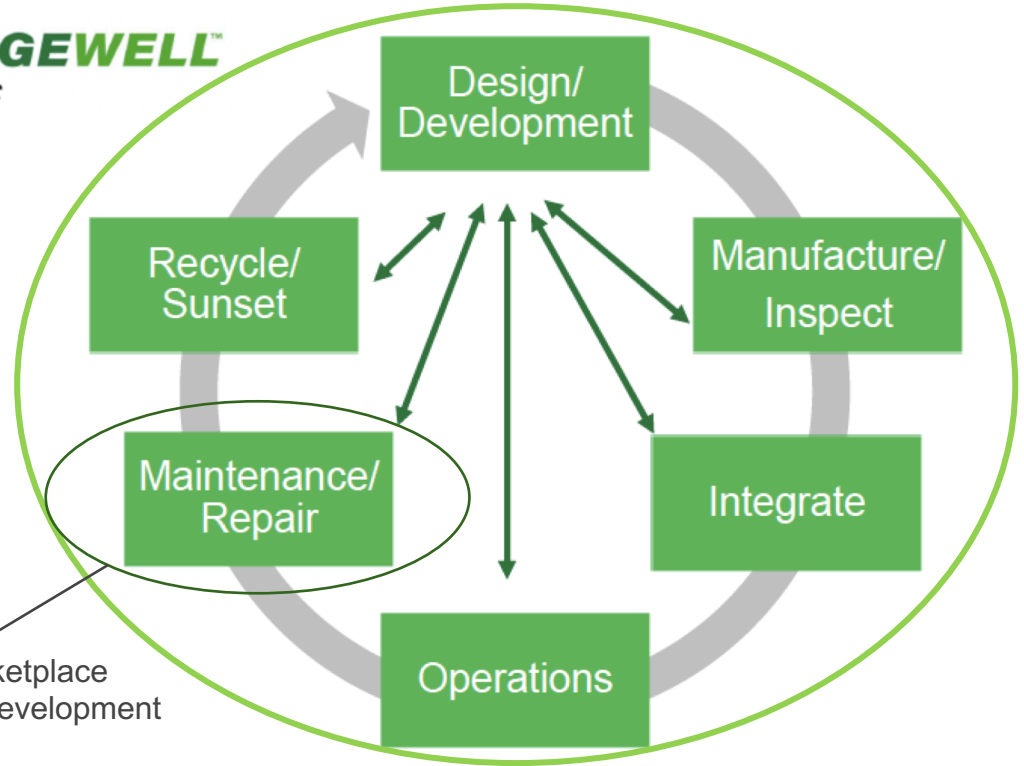
Digital Data Marketplace accelerating AI development

SAE ITC
An SAE International Affiliate
Collaborative Innovation.
Trusted Implementation.

DIGITAL DATA MARKETPLACE GOVERNANCE

A FOUR STEP APPROACH

ExchangeWell

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** concepts

DEMO

AIRFRANCE KLM

DIGITAL DATA MARKETPLACE ARCHITECTURE

RESEARCHING IMPLEMENTATION OF ESSENTIAL ELEMENTS
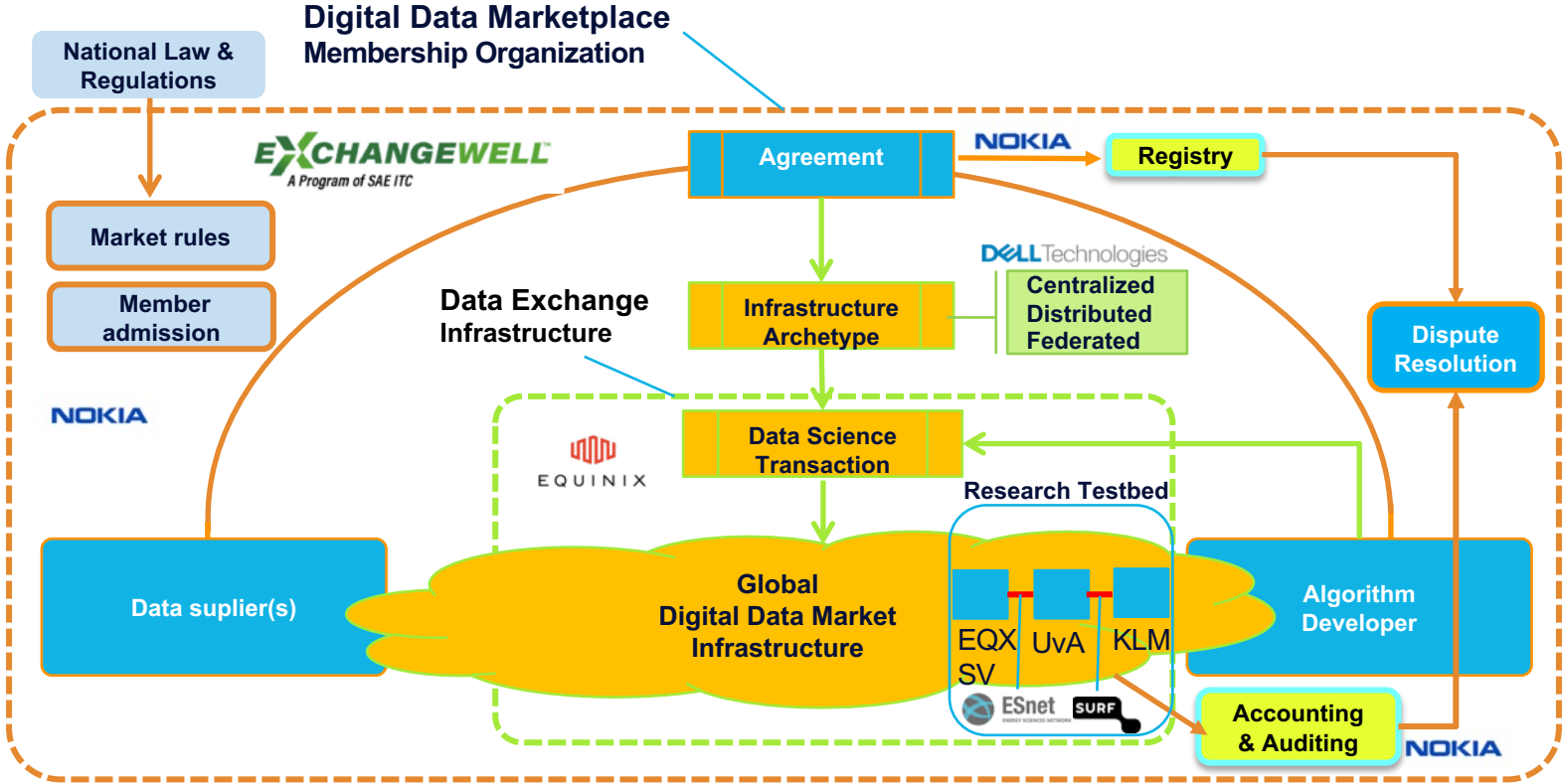
PROOF OF CONCEPT
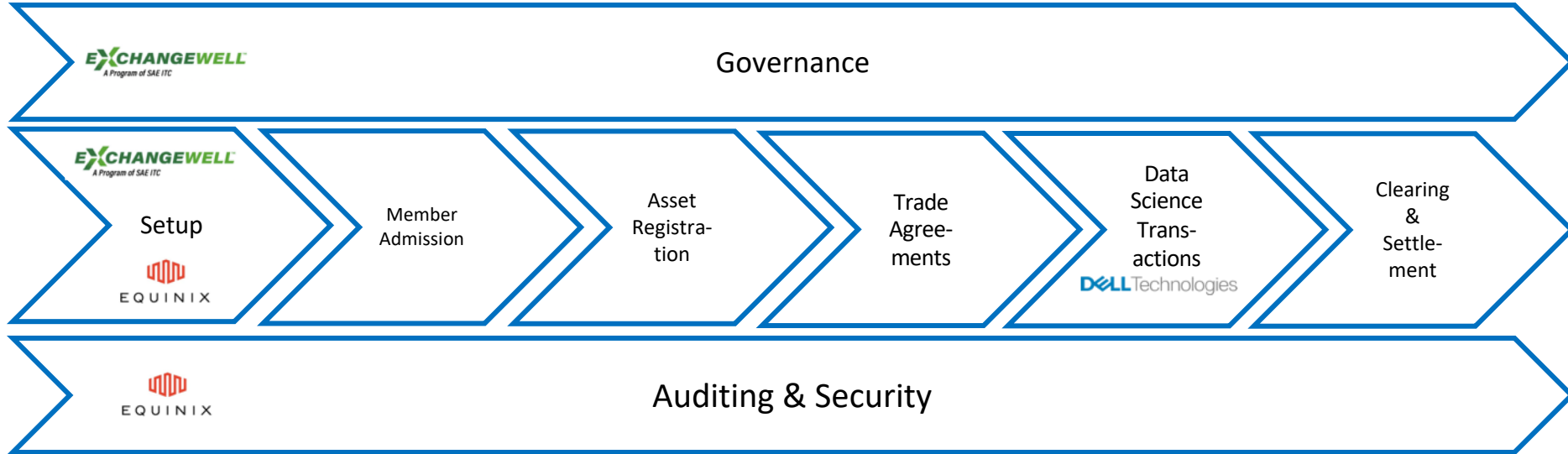## PARTNERS INVOLVED IN DEMO OF THE ARCHITECTURE

	Partner	Key Points
AIRLINE CONSORTIUM	ExCHANGEWELL (A Program of SAE ITC)	• Governance Layer
DATA MARKETPLACE EXCHANGE	NOKIA	• Orchestrator/Control Plane for Marketplace
FEDERATED ANALYTICS	DELL Technologies	• Distributed/Federated Analytics
EQUINIX PLATFORM 2.0 Computer/Container Platform Provider / Storage Provider	NetApp, EQUINIX, DELL Technologies, vmware	• Compute/Storage Hardware Platform Layer
Network Provider / Data Center Provider	EQUINIX	• Data Center & Network Fabric

AIRFRANCE KLM

Solution Integration
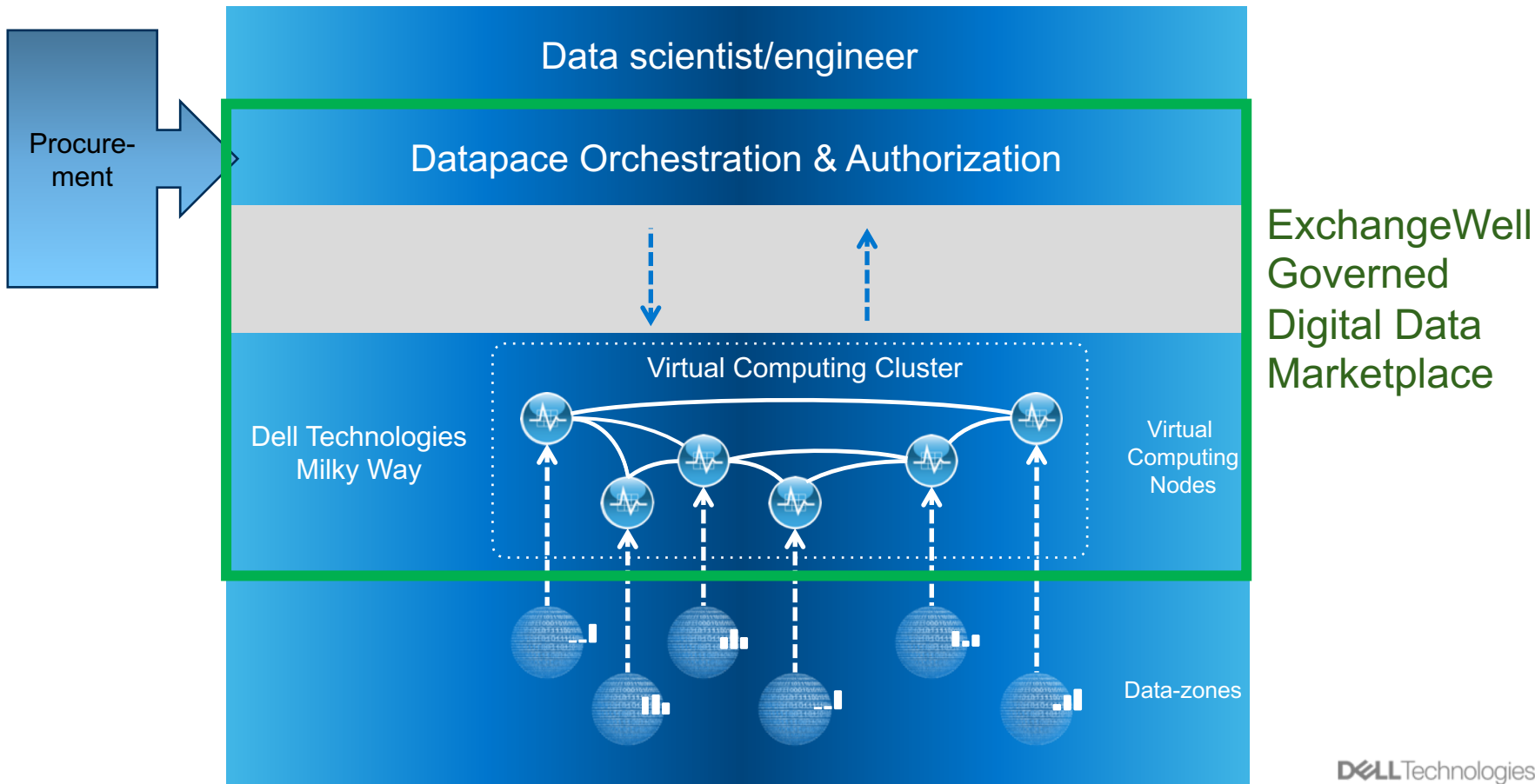
DEMO: DATAPACE IMPLEMENTING WORKFLOW GOVERNANCE & AUDITING USING BLOCKCHAIN & SMART CONTRACTS

QUESTIONS