DATA SHARING FOR AI

Towards an industrial lab

Leon Gommans, PhD – Science Officer
Professor Data Exchange Systems at University of Amsterdam

Air France KLM Group IT Technology Office
R&D department.

Presentation NMT Oct. 19th 2020
DATA SHARING SOLUTION
DIGITAL DATA MARKETPLACE GOVERNED BY A MEMBERSHIP CONSORTIUM

Algorithm Developers
own or third party

Computer science
Math and statistics
Data science

Competitive
Domain knowledge

Decision Support Systems

Planning, Prediction, Prevention, Effectiveness, Efficiency, etc.

(Near) Real Time Operational Data

Own Organization

Data supplied by other organizations

Data supplied by other organizations

Own Organization Data

Historic (Big) Data

Periodic storage: raw or with enhanced quality

enabling access and use

Competitive Algorithm Choice

Algorithm Developers
own or third party

Own Organization

Own Organization

Periodic storage: raw or with enhanced quality

enabling access and use

Own Organization

Own Organization

Own Organization

Own Organization
JOURNEY OF THE DATA SCIENTIST / ENGINEER
ROLE OF THE DIGITAL DATA MARKETPLACE: FOCUS IS ON INDUSTRIALIZE PHASE

ACQUIRE MORE AND DIVERSE DATA

JOURNEY START

IDEA

1. DEFINE

2. COLLECT

3. EXPLORE

EXPERIMENT

flexibility

Feasible Model

USE CASE

BUSINESS CASE

go-nogo

INDUSTRIALIZE

reliability

MVP

Go live

PRODUCT

SHELL PREDICTIVE PRODUCTS

10. SETUP MONITORING

9. FEEDBACK LOOP

11. PREDICT

8. CONTINUOUS INTEGRATION

7. AUTO TRAIN

6. DATA FLOW

5. VALIDATE

4. MODEL

DEFINITE

EXPLORE

FEATURE

MODEL

VALIDATE

FLEXIBILITY

FEASIBLE MODEL

RELIABILITY

MVP

GO LIVE

TEAMWORK

NOTE:

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

DEFINE (1-3)

What is the end purpose?
How will it add value?
Use case, data scope and stakeholders identification

EXPERIMENT (4-5)

Collect and explore data, research data science model
Prototype, go/nogo to production

INDUSTRIALIZE (6-10)

Deliver the solution in production environment
Product GO live

IMPROVE (11-12)

Feedback analysis, A/B testing, and performance monitoring
New product releases, business insights

SHARE
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

Data owners

Data owners

consolidate

Data owners
ESSENTIAL INFRASTRUCTURE ARCHETYPES
CONSORTIUM DRIVING SECURITY & INFRASTRUCTURE ARCHITECTURE

Centralized
Bring data to the algorithm

Distributed
Bring algorithm to the data

Federated
Using trusted infrastructure

Data owners
Developer

consolidate

Data owners

Data owners
DIGITAL DATA MARKETPLACE GOVERNANCE
A FOUR STEP APPROACH:

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.

SAE ITC ExchangeWell

Equinix/Dell/Nokia/Ciena
Innopay T.R.U.S.T# framework applied to organizing collaborative governance approaches

### PLATFORM BRAND ORGANIZATION

<table>
<thead>
<tr>
<th>TRADE</th>
<th>RULES</th>
<th>USE &amp; STANDARDS</th>
<th>TECHNOLOGY</th>
</tr>
</thead>
</table>
| • Strategic decisions about the essence (why) of the platform  
• Value creation, Earning & Business models  
• Branding and communication  
• .. | • Stakeholder roles and responsibilities  
• Influence and decision taking  
• Rules of engagements  
• Operations & Governance  
• Liability, usage terms & conditions, privacy, .. | • Customer journey & experience  
• Propositions to supply side  
• Risk perception & management  
• Messaging / API standards  
• Industry standards & frameworks  
• Data models & semantics  
• Payment standards,.. | • Platform Architecture  
• Technology stack  
• Software components  
• Scaling & performance  
• Security  
• Sourcing  
• Software definability,.. |

### 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

---

DIGITAL DATA MARKETPLACE ARCHITECTURE

RESEARCHING IMPLEMENTATION OF ESSENTIAL ELEMENTS

Digital Data Marketplace Membership Organization

- National Law & Regulations
- Market rules
- Member admission

Data Exchange Infrastructure

- Agreement
- Infrastructure Archetype
- Centralized Distributed Federated
- Data Science Transaction
- Dispute Resolution

Global Digital Data Market Infrastructure

- Data supplier(s)
- Algorithm Developer
- Accounting & Auditing

Research Testbed

- EQX
- UvA
- KLM
- SV
- SURF
- ESnet

Centralized Distributed Federated
MARKETPLACE JOURNEY

Setup
- Member Admission

Asset Registration
- Trade Agreements

Governance

Transactions
- Clearing & Settlement

Auditing & Security
MEMBERS ALLOW EACH OTHER VISIBILITY

AI flow Data/Algorithm

Note: Nokia originally designed Datapace to work for streaming data
NEXT STEPS – CREATING LAB ENVIRONMENTS

Research Track
Work on use-cases in research lab with AMdEX/UvA/Innopay
Governance, normative reasoning, secure container networking, data exchange archetypes

Industry Track
Work on use-cases in industrial lab with Equinix/Dell/Nokia/SAE-ITC
Sustainability, AI for Aviation Systems, data driven MRO, NLP log processing, multi-modal travel

Funding sources
EU
GAIA-X
National
DATA SHARING COALITION
Regional
amsix
amsix
SAE
INTERNATIONAL
DNC TECHNOLOGY
XMNIA
ORTHOC
GO
DATA
DRIVEN
AMDEX
AIRFRANCE
KLM