Federated learning:

a hype or a trend?

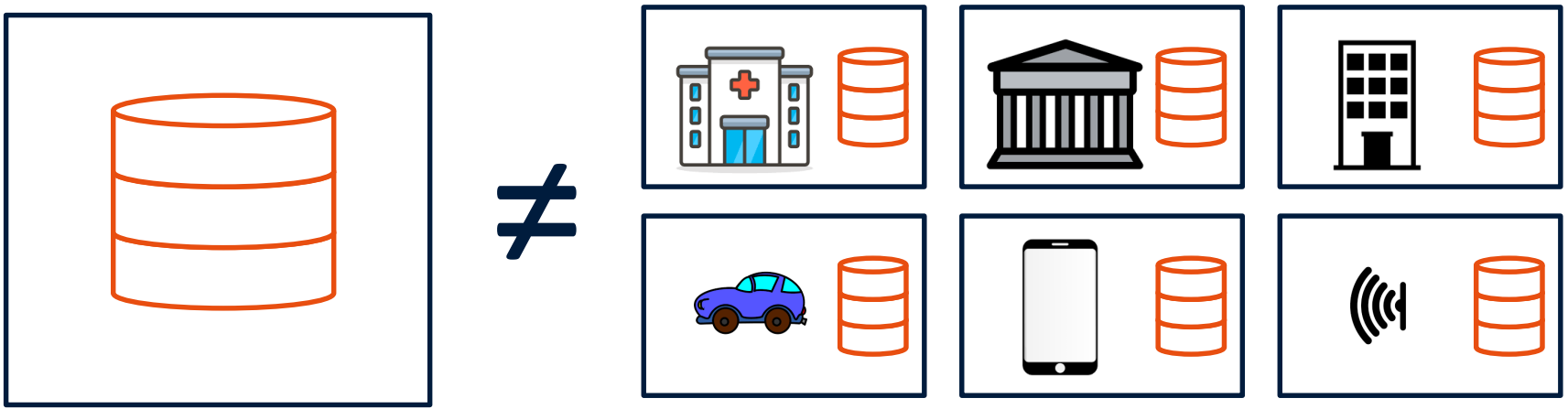
Anna Wilbik

Prof. Data Fusion & Intelligent Interaction

Agenda

- What is federated learning?
- **S**trengths & open challenges
- **W**eakness & **T**hreat
- **O**pportunity

From centralized to decentralized data



Federated learning

“We advocate an alternative that leaves the training data distributed on the mobile devices, and learns a shared model by aggregating locally-computed updates. We term this decentralized approach Federated Learning.”

McMahan et al. , Communication-Efficient Learning of Deep Networks from Decentralized Data, 2016.

Federated learning

“Federated learning is a machine learning setting where multiple entities (clients) collaborate in solving a machine learning problem, under the coordination of a central server or service provider. Each client’s raw data is stored locally and not exchanged or transferred; instead focused updates intended for immediate aggregation are used to achieve the learning objective.”

Kairouz et al., Advances and open problems in federated learning, 2019.

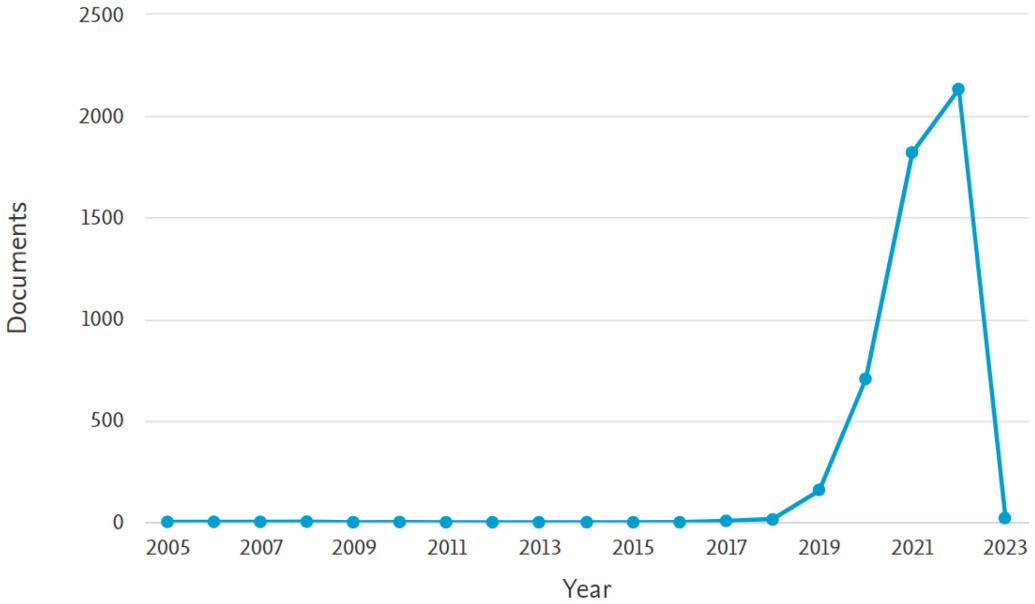
Federated learning

“collaborative learning without exchanging users’ original data”

Li et al., A survey on federated learning systems: vision, hype and reality for data privacy and protection, 2019.

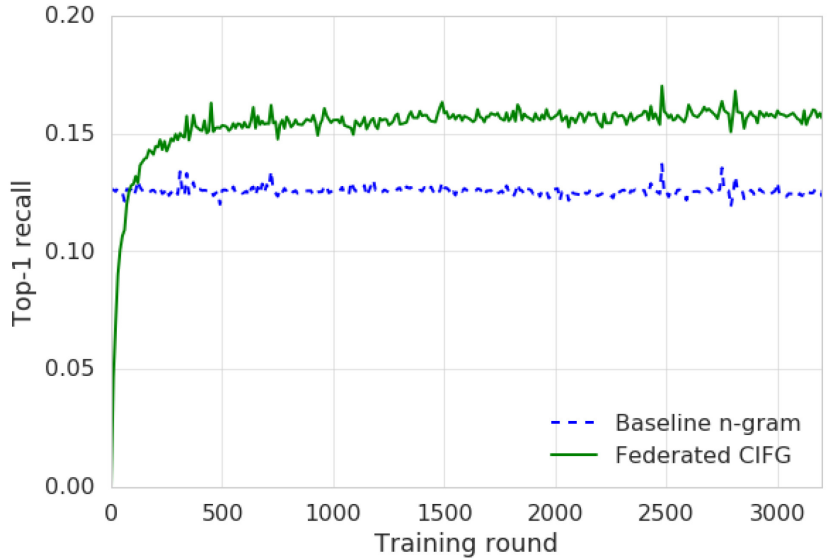
FL – area under development

Scopus



Gboard: next-word prediction

- Federated RNN (compared to prior n-gram model):
- Better next-word prediction accuracy: +24%
- More useful prediction strip: +10% more clicks



Hard et al. Federated Learning for Mobile Keyboard Prediction, arXiv:1811.03604

MIT Technology Review

Subscribe

ARTIFICIAL INTELLIGENCE

How Apple personalizes Siri without hoovering up your data

The tech giant is using privacy-preserving machine learning to improve its voice assistant while keeping your data on your phone.

By Karen Hao

December 11, 2019

https://www.technologyreview.com/2019/12/11/131629/apple-ai-personalizes-siri-federated-learning/

https://medcitynews.com/2020/05/upenn-intel-partner-to-use-federated-learning-ai-for-early-brain-tumor-detection/

ARTIFICIAL INTELLIGENCE, DIAGNOSTICS

UPenn, Intel partner to use federated learning AI for early brain tumor detection

The project will bring in 29 institutions from North America, Europe and India and will use privacy-preserved data to train AI models. Federated learning has been described as being born at the intersection of AI, blockchain, edge computing and the Internet of Things.

By ALARIC DEARMENT

Medical Institutions Collaborate to Improve Mammogram Assessment AI with NVIDIA Clara Federated Learning

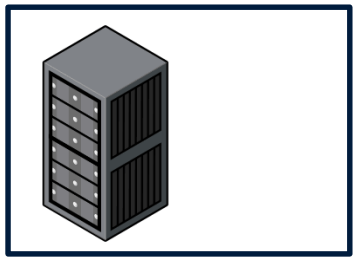
In a federated learning collaboration, the American College of Radiology, Diagnosticos da America, Partners HealthCare, Ohio State University and Stanford Medicine developed better predictive models to assess breast tissue density.

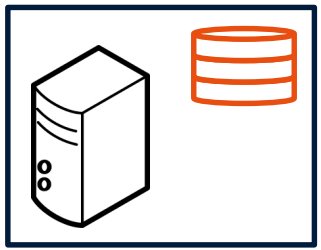
April 15, 2020 by MONA FLORES

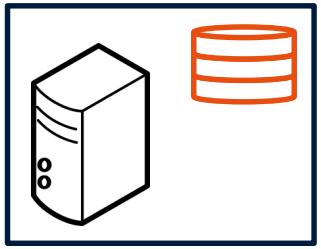
https://blogs.nvidia.com/blog/2020/04/15/federated-learning-mammogram-assessment/

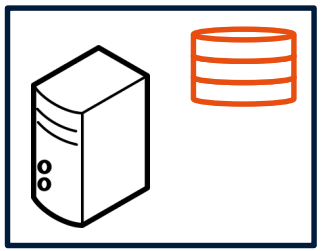
Scenario 1 – horizontal FL

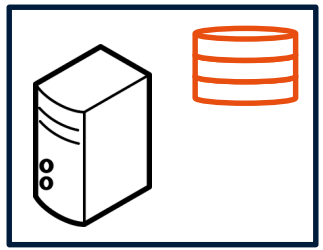
How does it work?



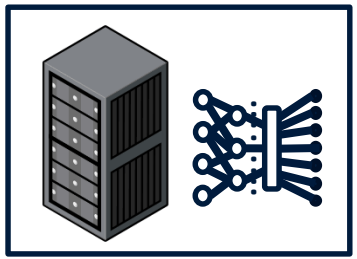


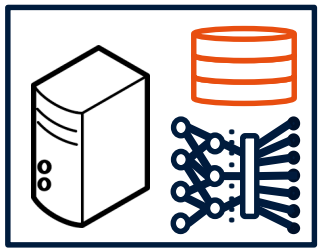


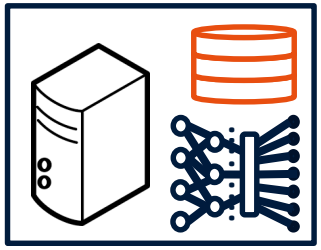


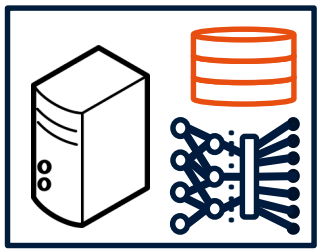


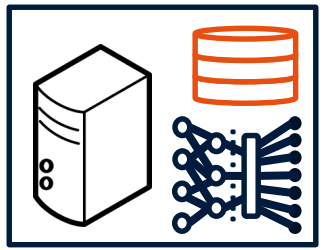
How does it work?



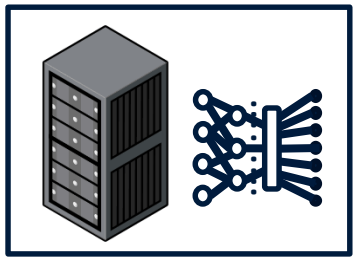


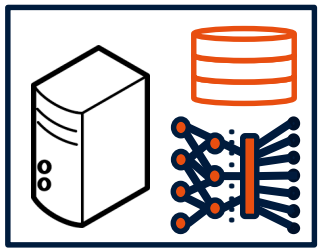


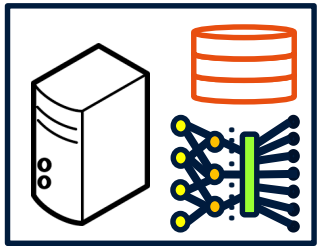


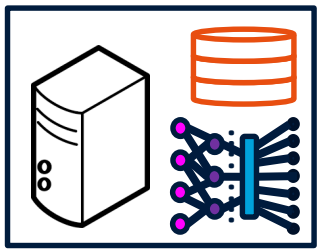


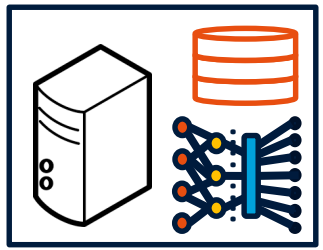
How does it work?



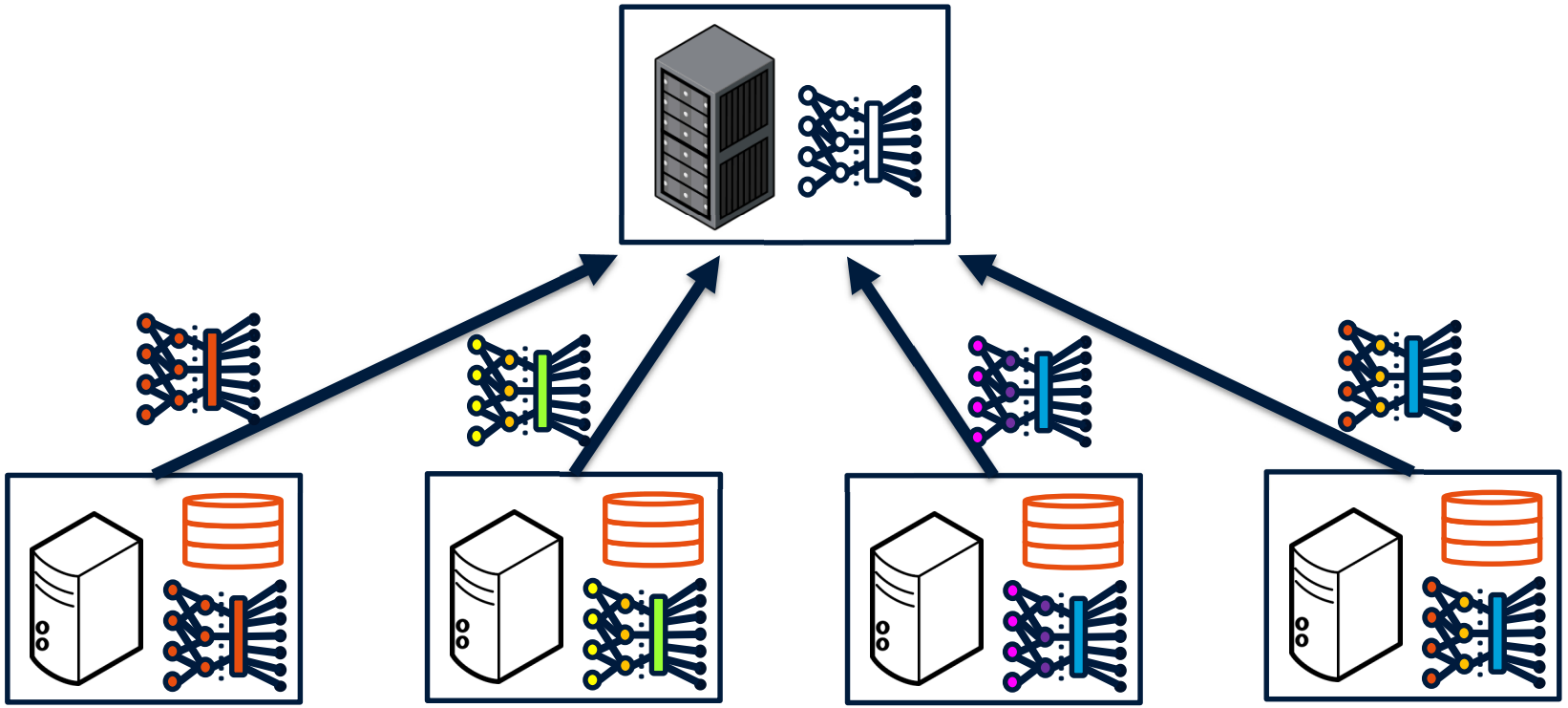




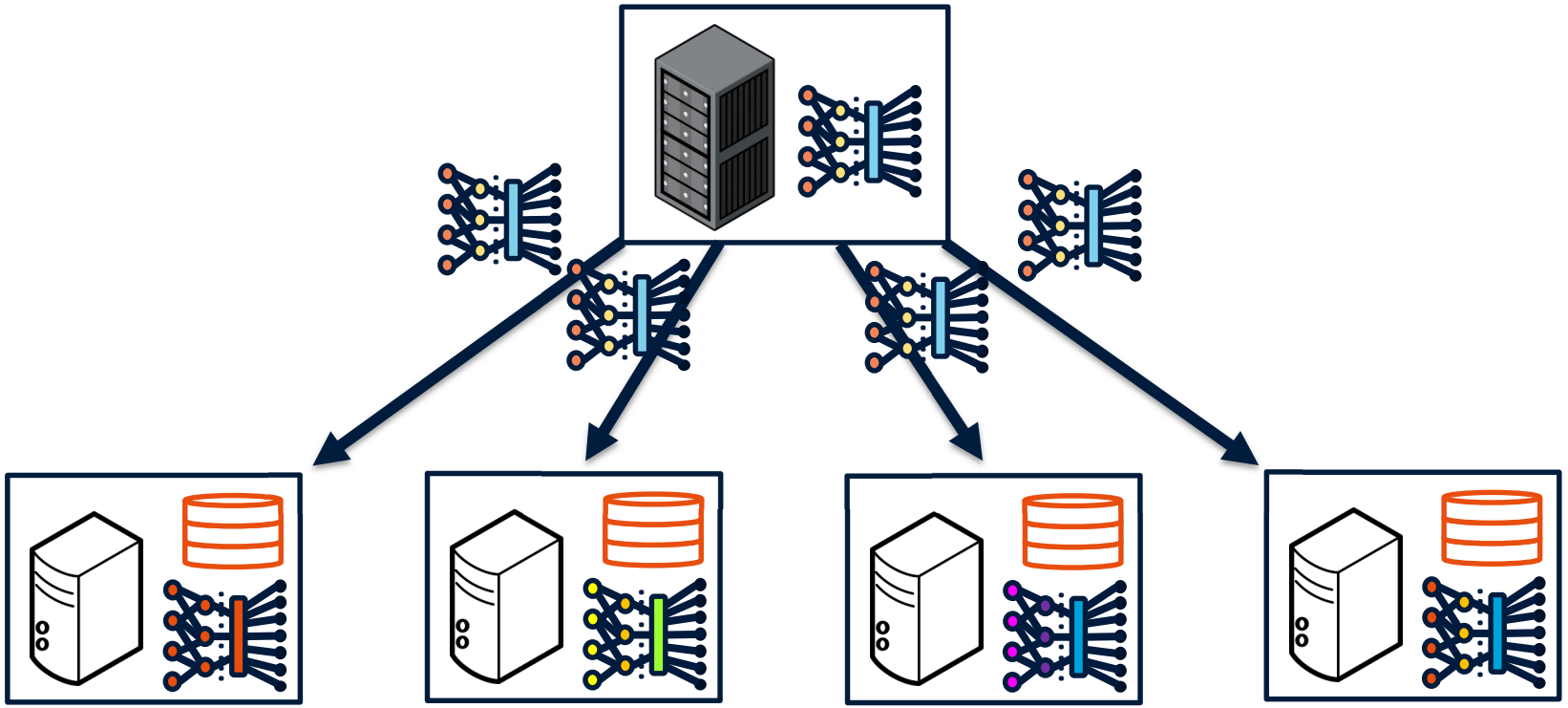




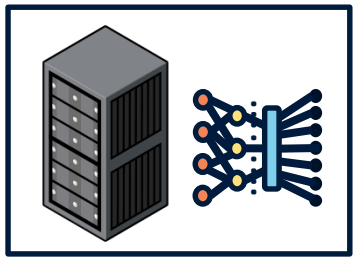
How does it work?

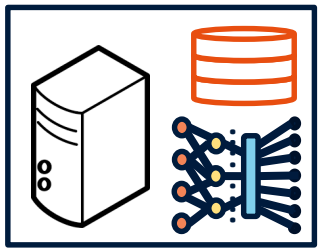


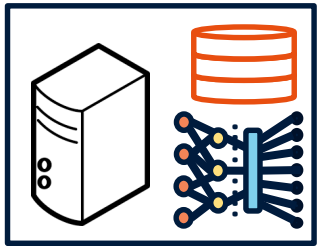
How does it work?

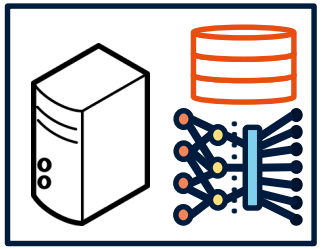


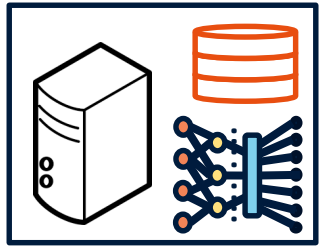
How does it work?





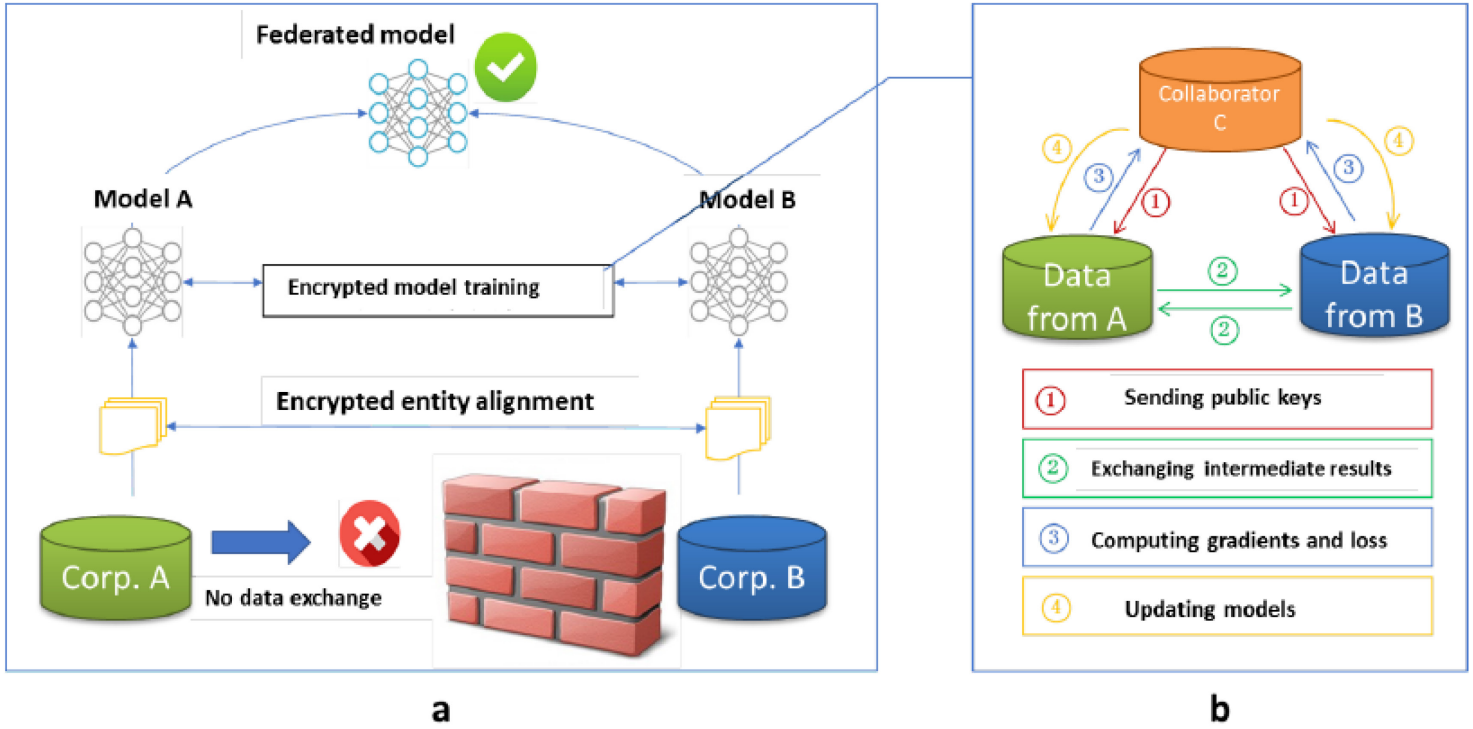






Scenario 2 – vertical FL

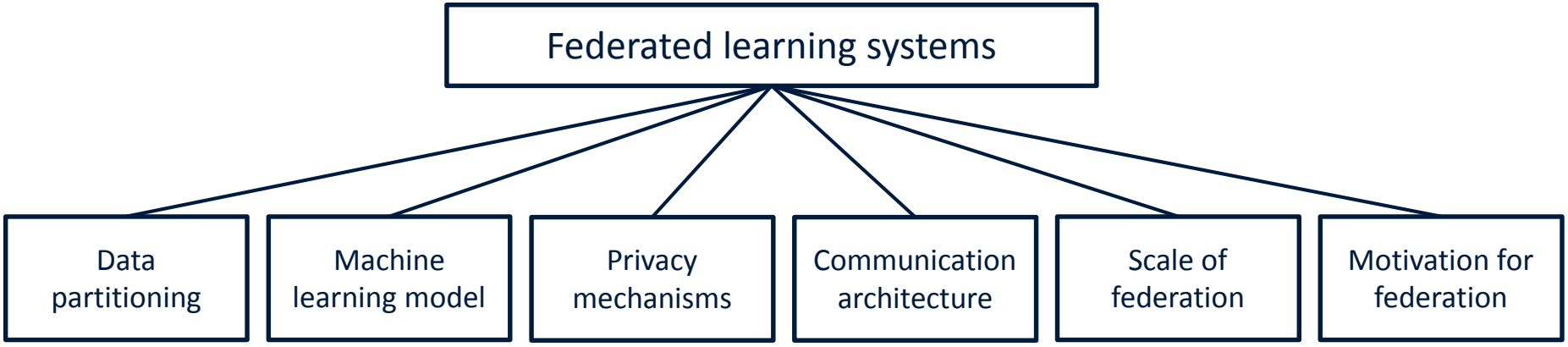
Vertical federated learning



Yang, et al., Federated Machine Learning: Concept and Applications

Scenario 3 – hybrid FL

Taxonomy of Federated Learning



Li et al., A survey on federated learning systems: vision, hype and reality for data privacy and protection, arXiv preprint arXiv:1907.09693, 2019.

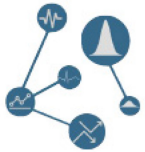
HFL: research

expensive communication

- massive, slow networks

privacy concerns

- user privacy constraints



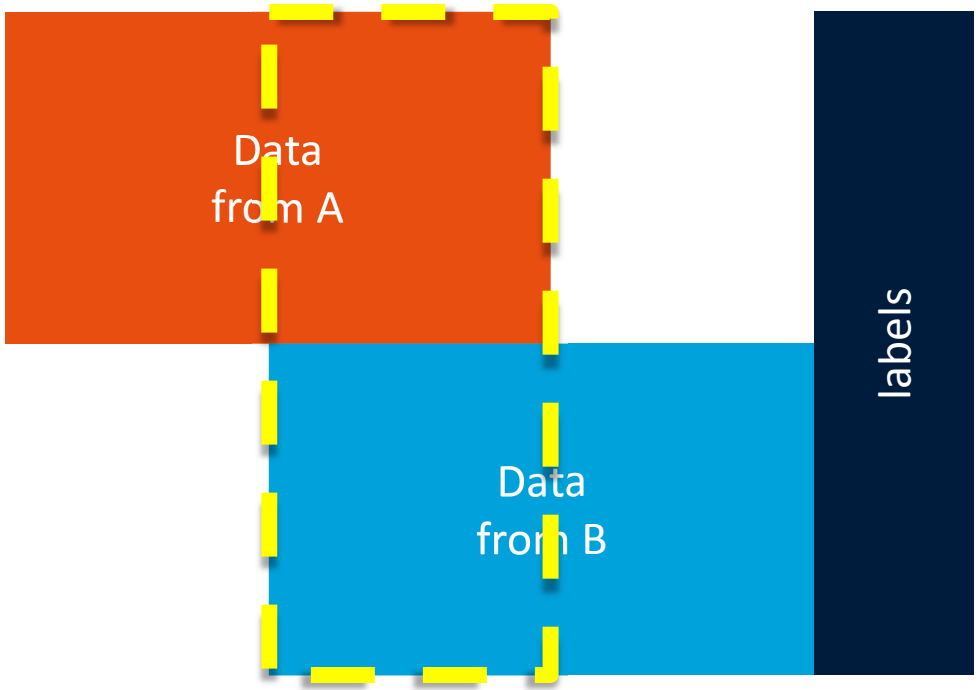
statistical heterogeneity

- unbalanced, non-IID data

systems heterogeneity

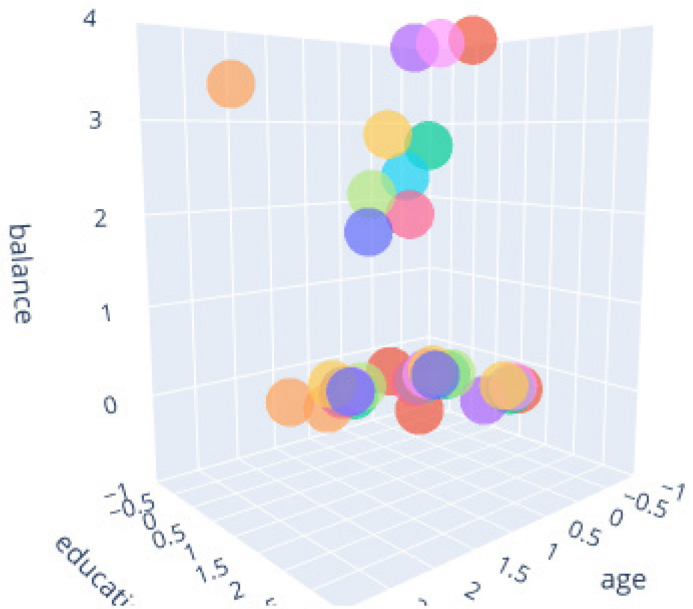
- variable hardware, connectivity, etc

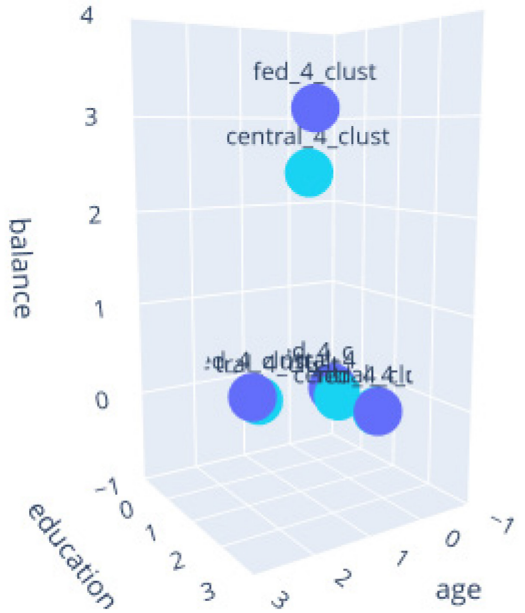
Not all parties collect same data



Federated learning with uncertainty on the example of a medical data K Dyczkowski, B Pękala, J Szkoła, A Wilbik
2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)

Federated clustering



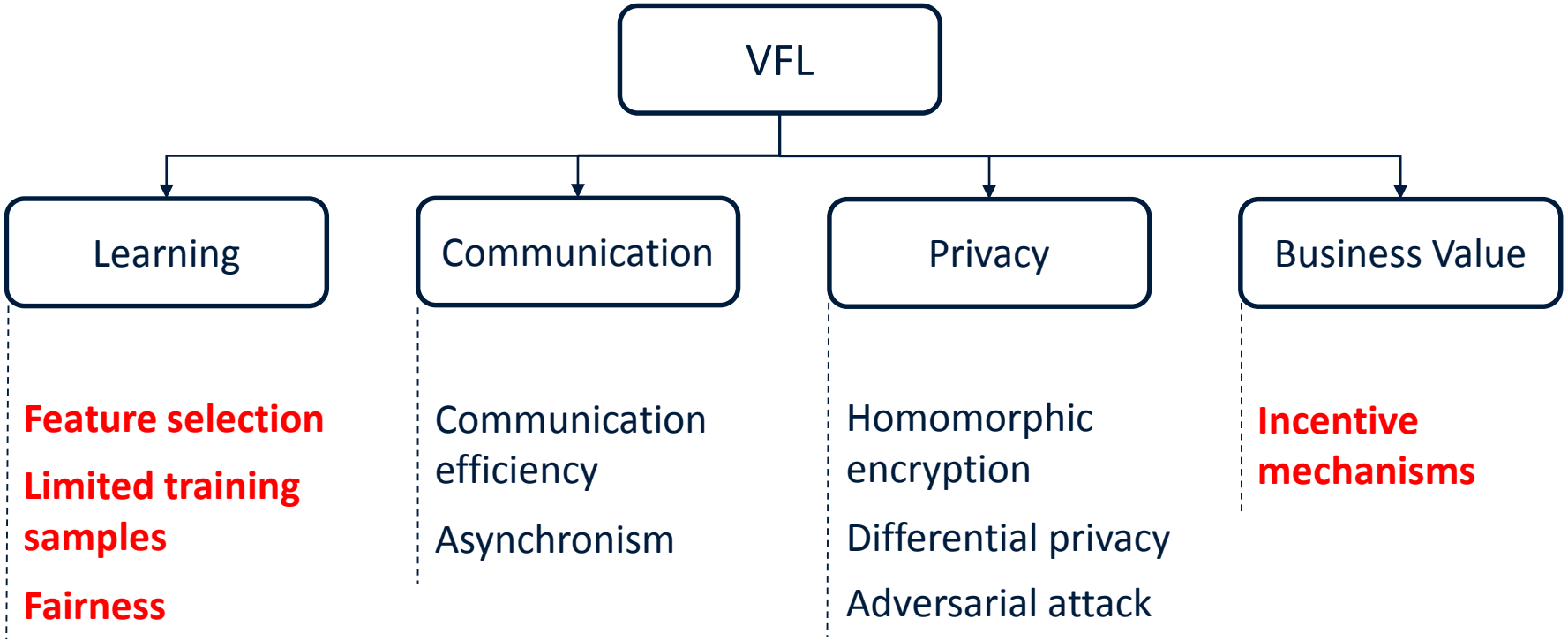


On a Framework for Federated Cluster Analysis
M. Stallmann, A. Wilbik, Applied Sciences, in press

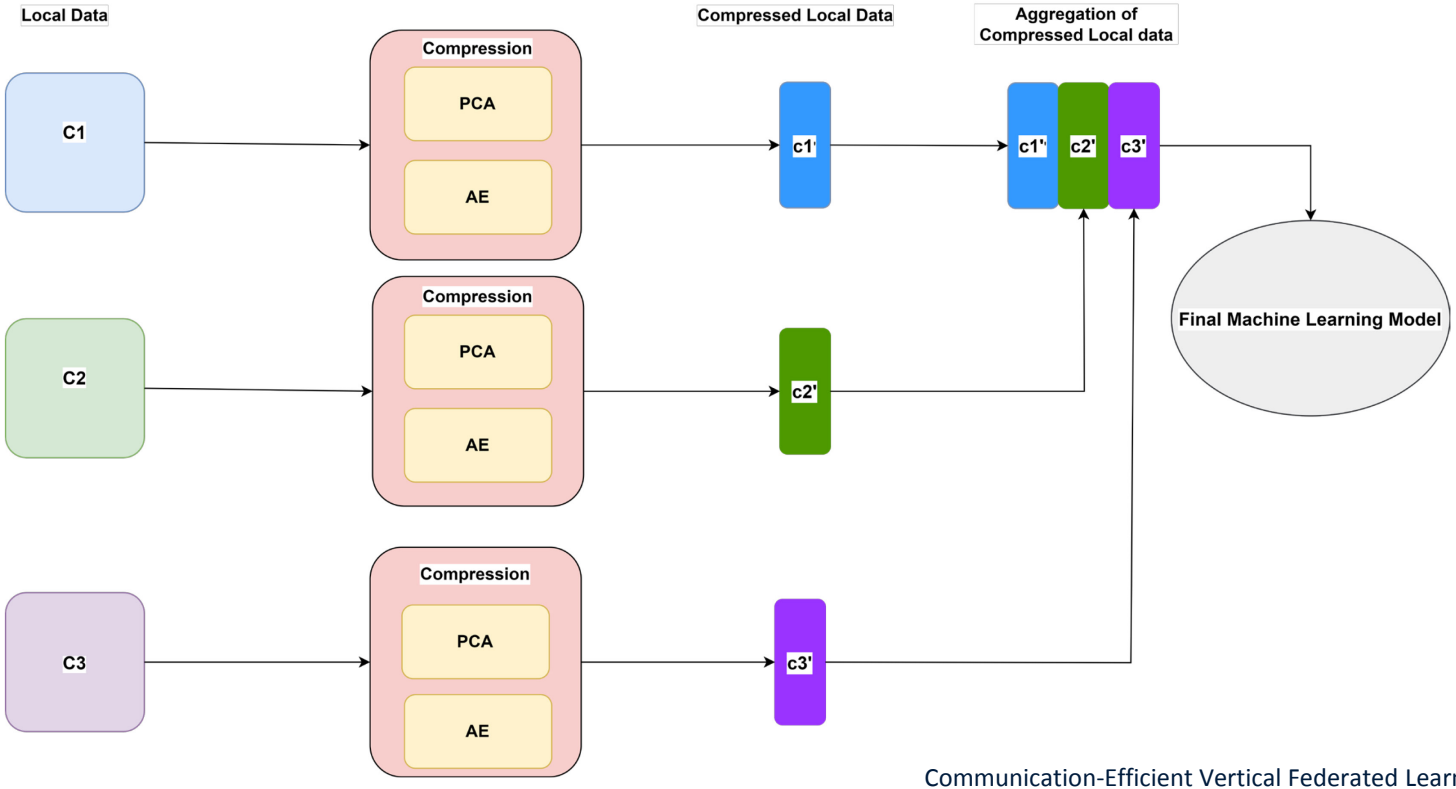
HFL: opportunities

- Going beyond empirical risk minimization formulations: tree-based methods, online learning, Bayesian learning...
- RL, unsupervised and semi-supervised, active learning
- Support ML workflows like hyperparameter searches
- Data alignment
- Make trained models smaller
- Fairness

VFL: research & opportunities



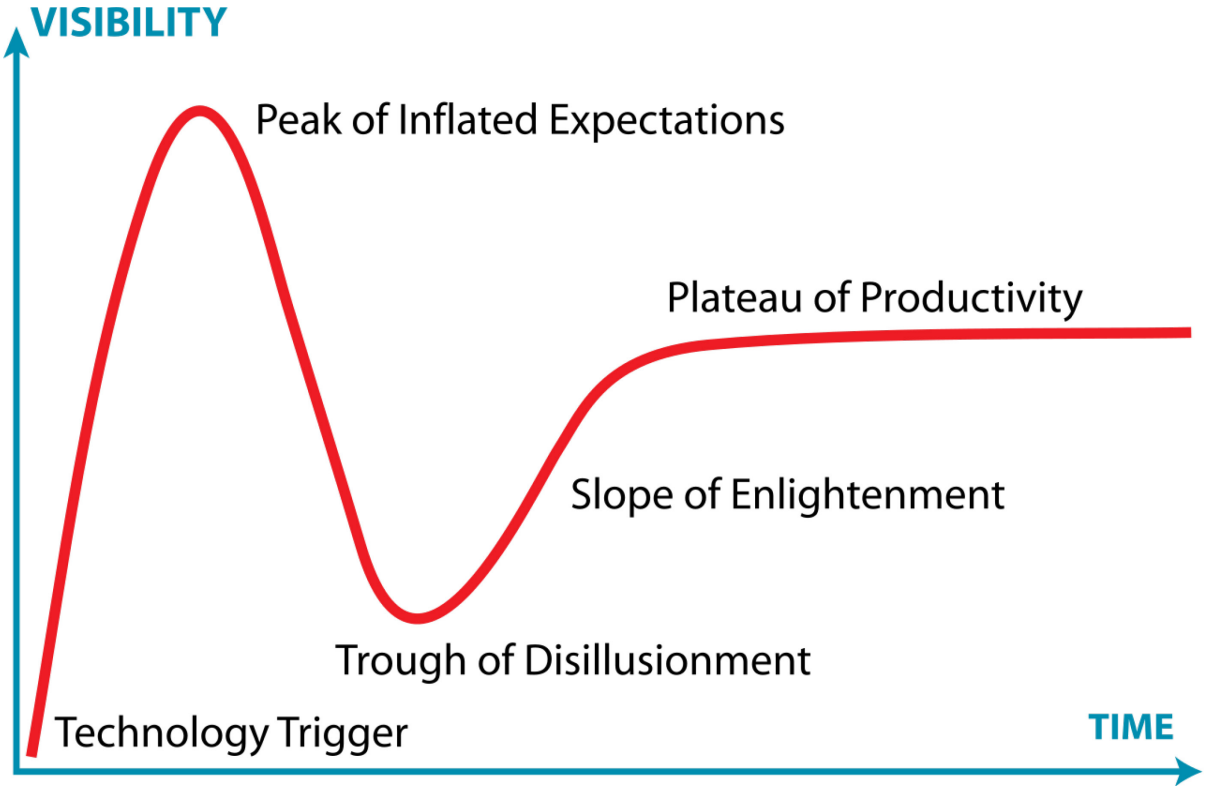
Communication efficiency

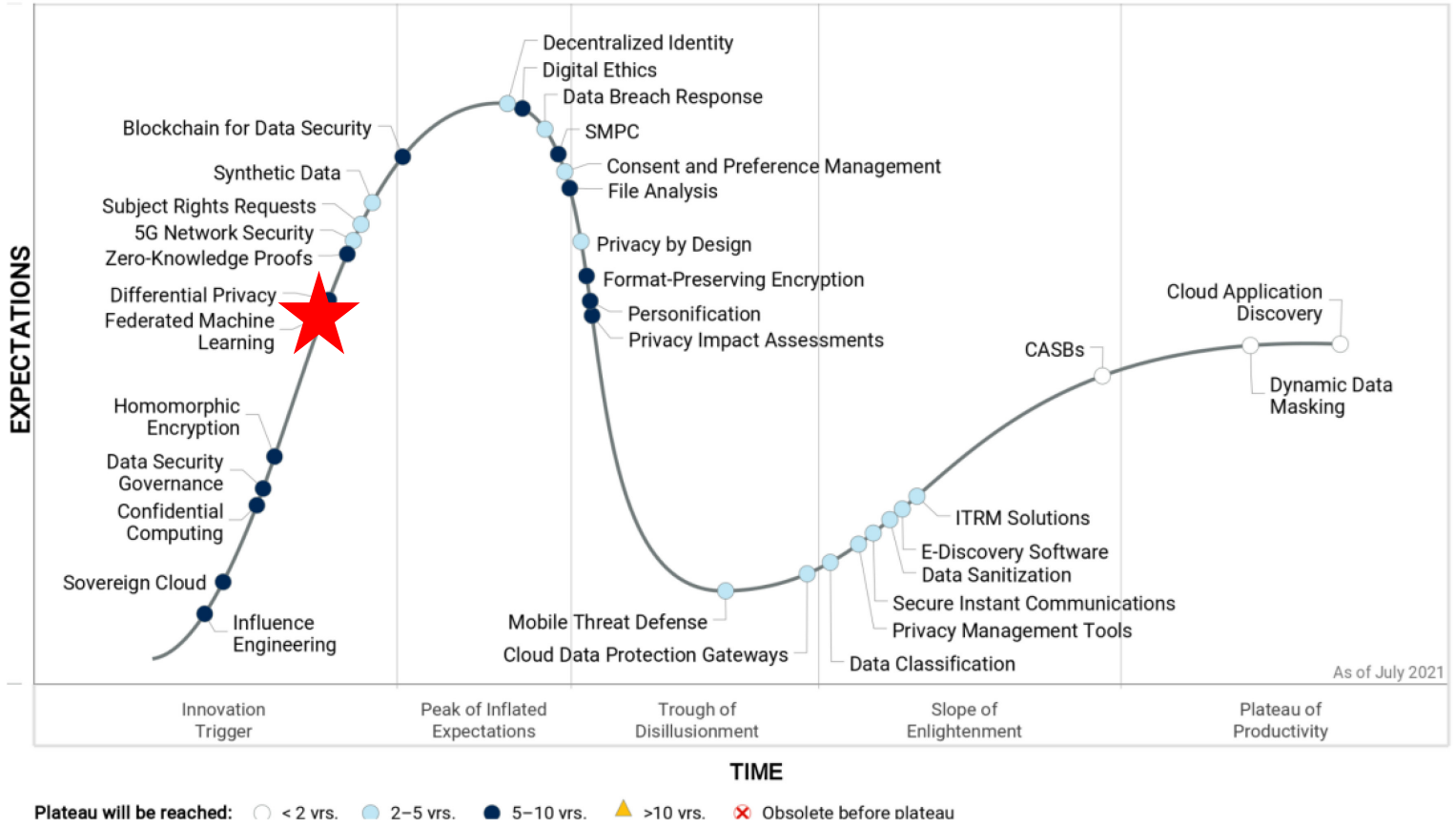


Communication-Efficient Vertical Federated Learning
A Khan, M ten Thij, A Wilbik, Algorithms 15 (8), 273

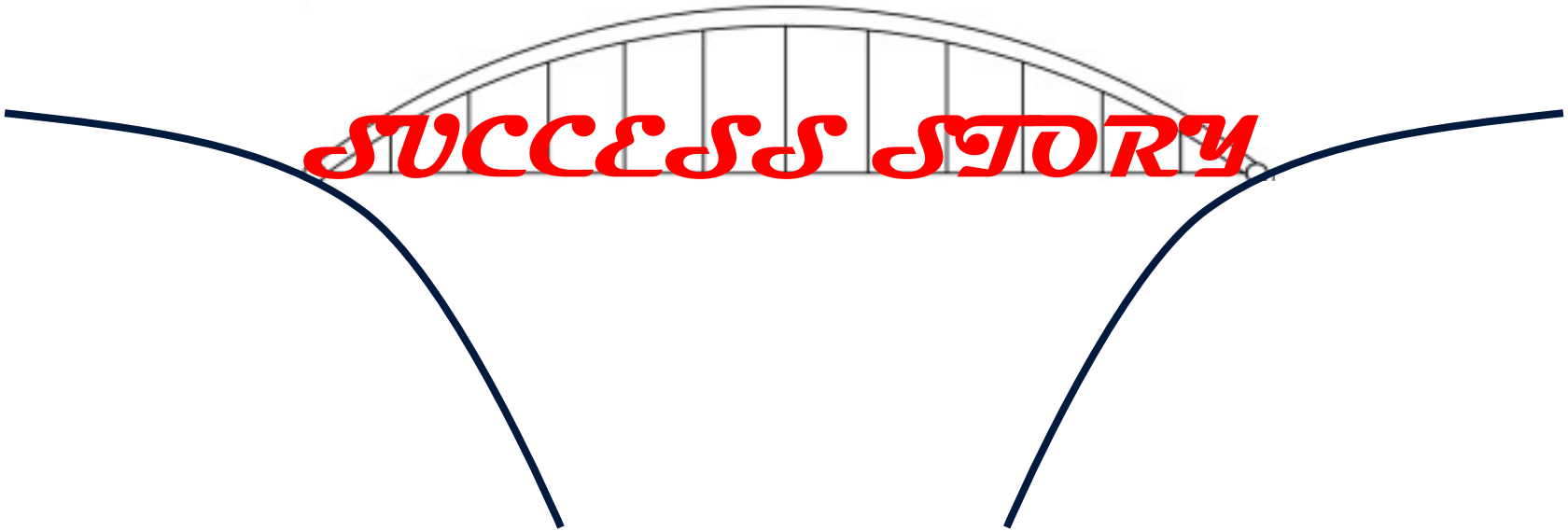
But…

Gartner’s hype cycle




EXPECTATIONS
Decentralized Identity
Digital Ethics
Data Breach Response
Blockchain for Data Security
SMPC
Synthetic Data
Consent and Preference Management
Subject Rights Requests
File Analysis
5G Network Security
Privacy by Design
Zero-Knowledge Proofs
Format-Preserving Encryption
Differential Privacy
Federated Machine Learning
Personification
Privacy Impact Assessments
Cloud Application Discovery
CASBs
Dynamic Data Masking
Homomorphic Encryption
Data Security Governance
Confidential Computing
ITRM Solutions
E-Discovery Software
Data Sanitization
Sovereign Cloud
Secure Instant Communications
Privacy Management Tools
Influence Engineering
Mobile Threat Defense
Cloud Data Protection Gateways
Data Classification
As of July 2021
Innovation Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateau of Productivity
TIME
Plateau will be reached:
< 2 yrs.
2–5 yrs.
5–10 yrs.
>10 yrs.
Obsolete before plateau

How to pass the trough of disillusionment?



TRL


DEPLOYMENT
DEVELOPMENT
RESEARCH
9
ACTUAL SYSTEM PROVEN IN OPERATIONAL ENVIRONMENT
8
SYSTEM COMPLETE AND QUALIFIED
7
SYSTEM PROTOTYPE DEMONSTRATION IN OPERATIONAL ENVIRONMENT
6
TECHNOLOGY DEMONSTRATED IN RELEVANT ENVIRONMENT
5
TECHNOLOGY VALIDATED IN RELEVANT ENVIRONMENT
4
TECHNOLOGY VALIDATED IN LAB
3
EXPERIMENTAL PROOF OF CONCEPT
2
TECHNOLOGY CONCEPT FORMULATED
1
BASIC PRINCIPLES OBSERVED

Technology push vs. requirements pull

Technology push

Requirements pull

Requirement pull = outcome thinking?

“People don’t want to buy a quarter-inch drill. They want a quarter-inch hole.” - Theodore Levitt (Harvard University), 1983

Outcome thinking

Outcome thinking

Outcome-Based Business Design in IoT-Enabled Digital Supply Chain Transformation

Paul Grefen
School of Industrial Engineering
Eindhoven University of Technology and
Atos Digital Transformation Consulting
Eindhoven, Netherlands
p.w.p.j.grefen@tue.nl, paul.grefen@atos.net

Anna Wilbik
Deptartment of Data Science and Knowledge Engineering
Maastricht University
Maastricht, Netherlands
a.wilbik@maastrichtuniversity.nl

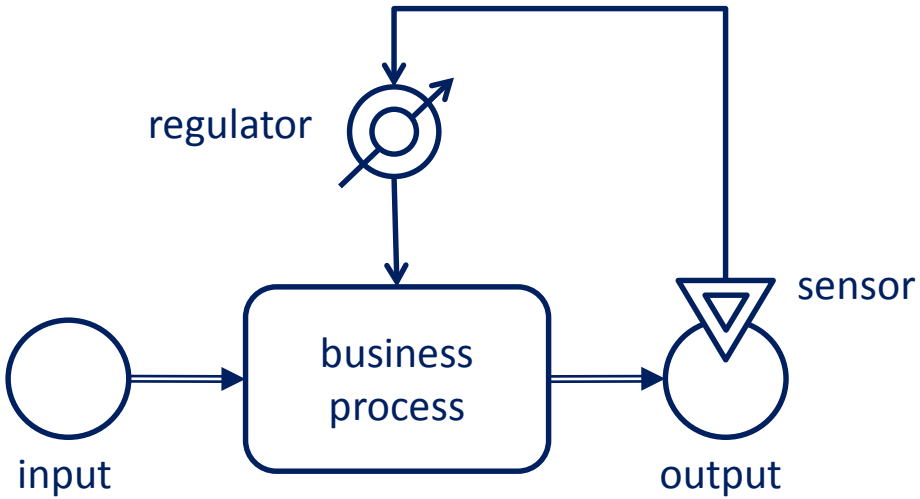
Frank Kuitems
Atos Digital Transformation Consulting
Eindhoven, Netherlands
frank.kuitems@atos.net

Menno Blanken
Atos Digital Transformation Consulting
Eindhoven, Netherlands
menno.blanken@atos.net

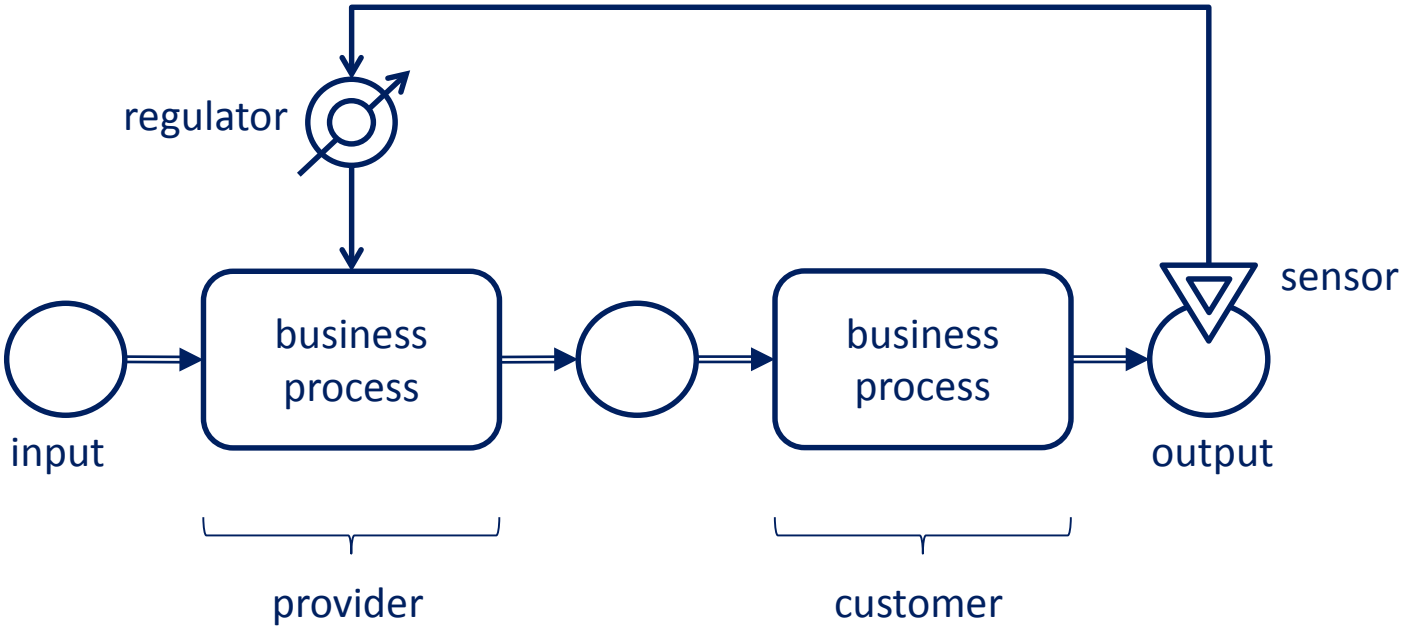
***Abstract*—In the current economy, we see a shift of focus from delivering products or services to delivering value or outcomes to customers, reflected in the concept of the outcome economy. The concept has been embraced by research and practice but lacks proper operationalization to make it fit for the digital transformation of supply chains. In this paper, we translate the concept into a cybernetic model and accompanying**

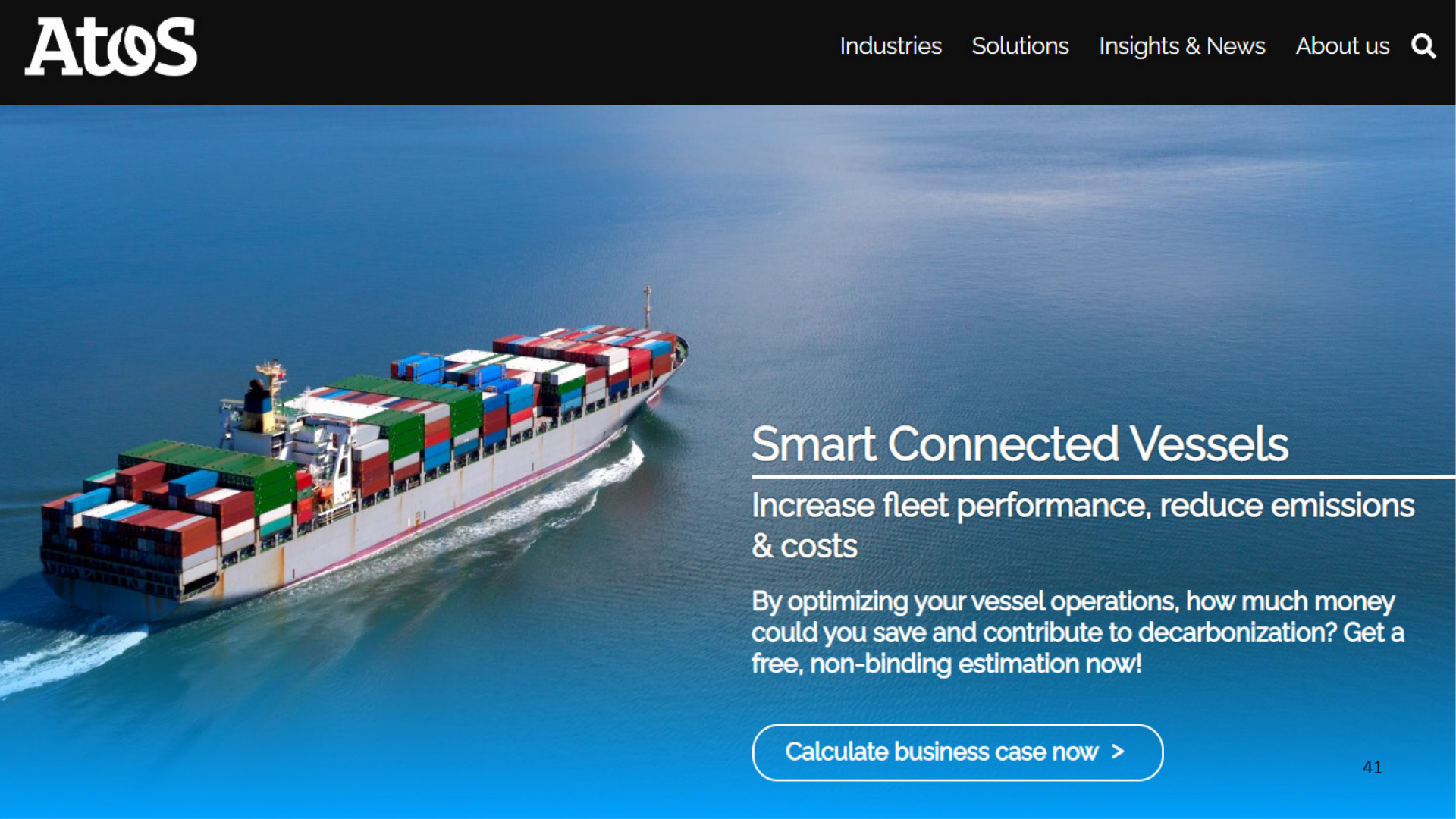
recent years [1-5]. Examples of this shift from selling products to selling outcomes can be found in many domains. An illustrative example is in the aircraft engine industry, where business models are explored where actual performance of engines is sold instead of the physical product [3]. In the transport and logistics domain, business models are explored where the effects of data analytics services on transport

Single-step business control model

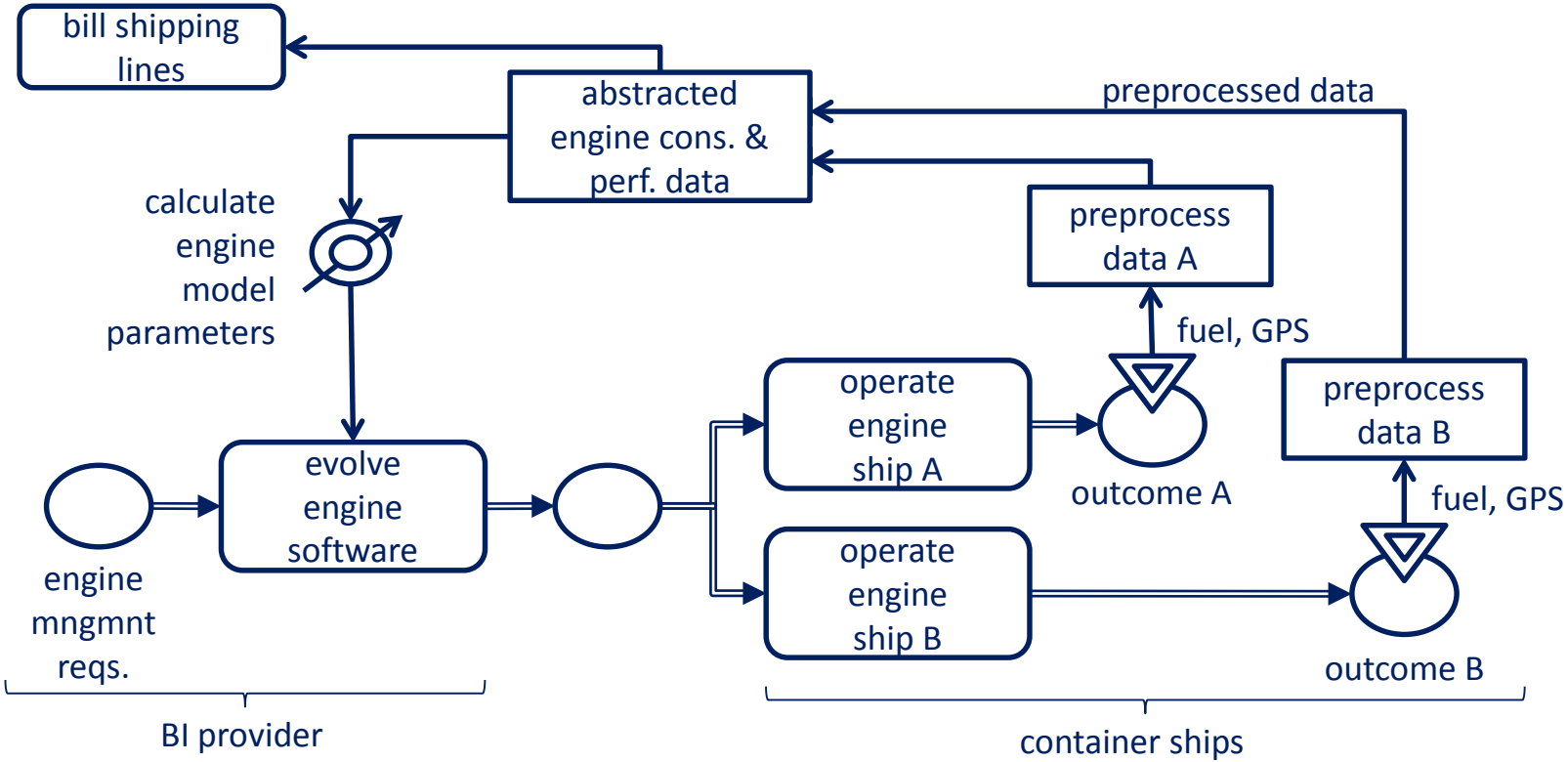


Multi-step business outcome control model

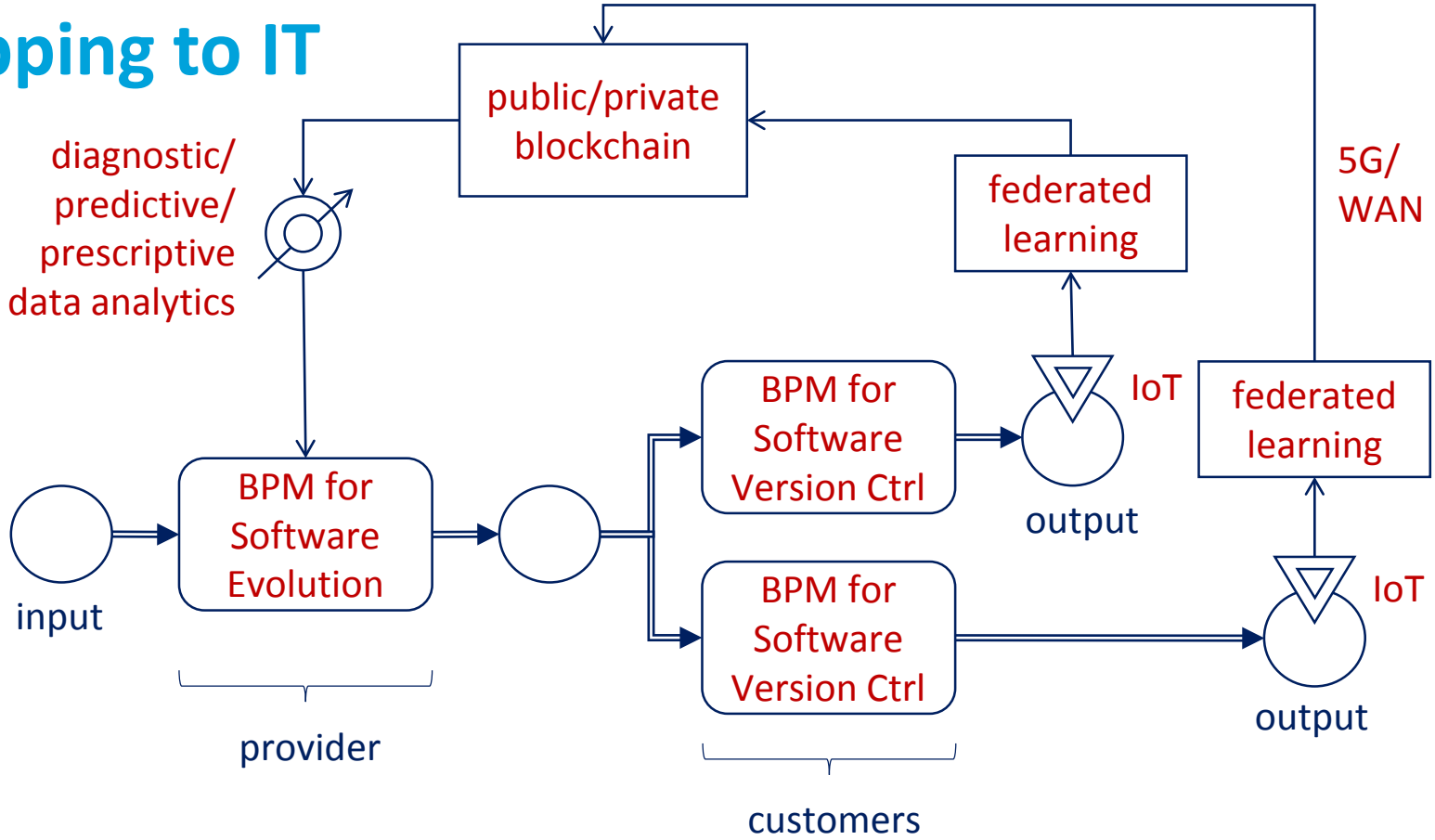



Atos
Industries
Solutions
Insights & News
About us
Smart Connected Vessels
Increase fleet performance, reduce emissions & costs
By optimizing your vessel operations, how much money could you save and contribute to decarbonization? Get a free, non-binding estimation now!
Calculate business case now >

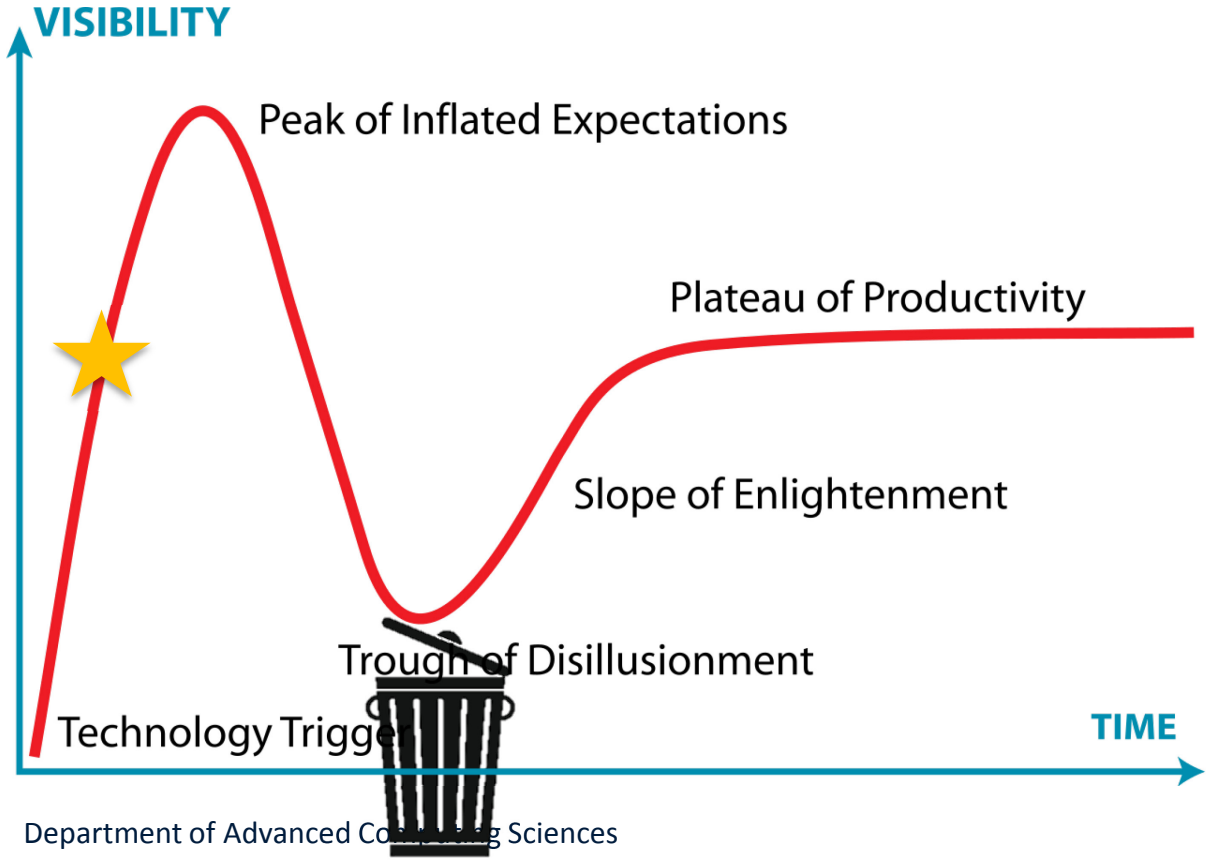
Smart connected vessels, extended conceptual view



Mapping to IT



Conclusion and outlook




TREND
HYPE

Get in touch

Anna Wilbik
a.wilbik@maastrichtuniversity.nl

Paul-Henri Spaaklaan 1
6229 EN Maastricht

Learn more about us at
www.maastrichtuniversity.nl/dacs