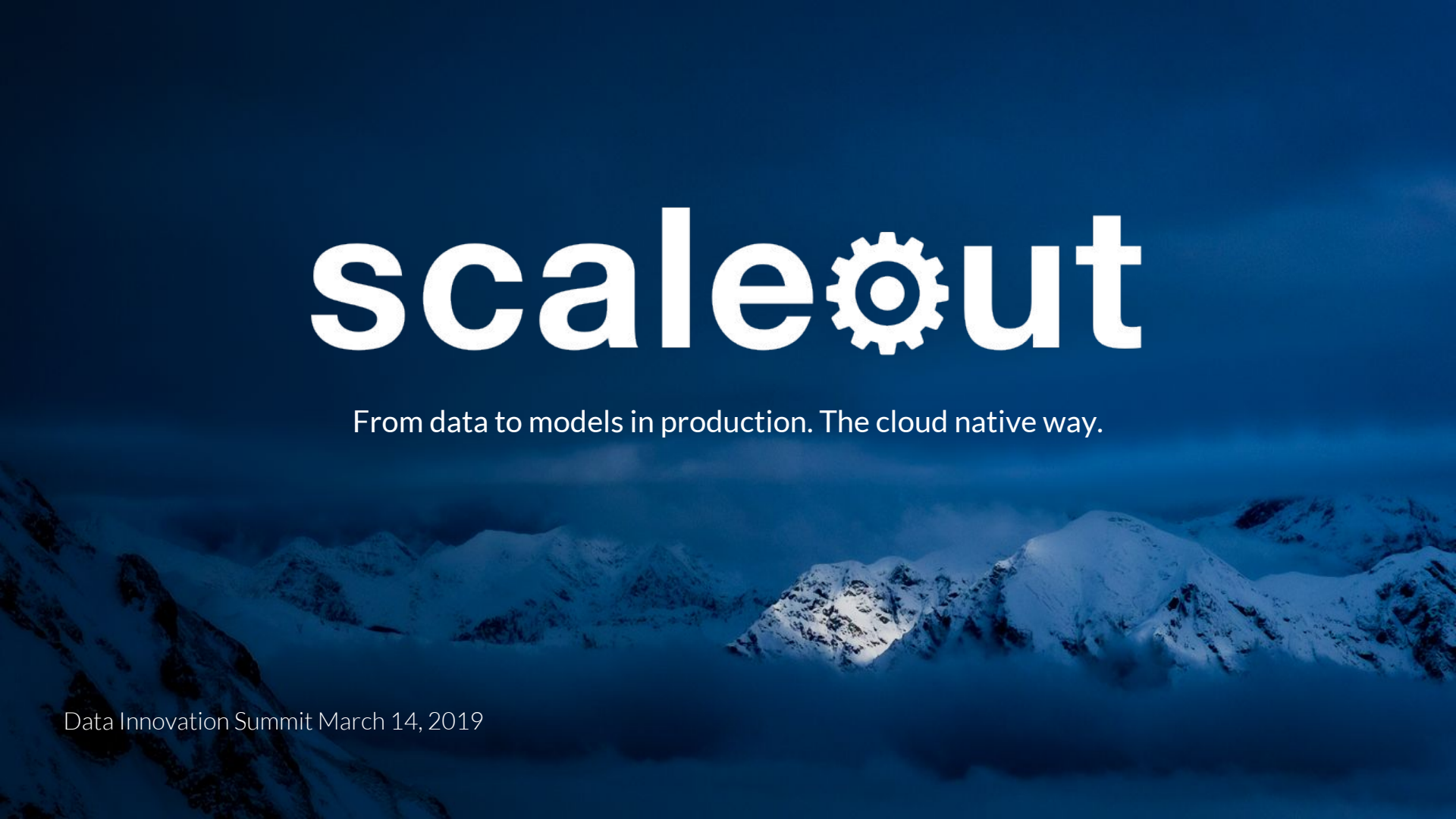

scaleout
From data to models in production. The cloud native way.
Data Innovation Summit March 14, 2019

# Data is becoming the most valuable asset

*"By 2030, companies that were early AI adopters could potentially double their cash flow, while non-adopters might experience around a 20% decline"*
*McKinsey, Sept 2018*

ComputerSweden
FROM IDG



1177-LÄCKAN SÄKERHET 2019-02-18 13:39

# 2,7 miljoner inspelade samtal till 1177 Vårdguiden helt oskyddade på internet

Computer Sweden avslöjar: alla telefonsamtal som ringts till 1177 sedan 2013 och som tagits emot av vårdentreprenören Medicall har legat helt öppet som ljudfiler på en oskyddad webbserver.

# World's Biggest Data Breaches & Hacks

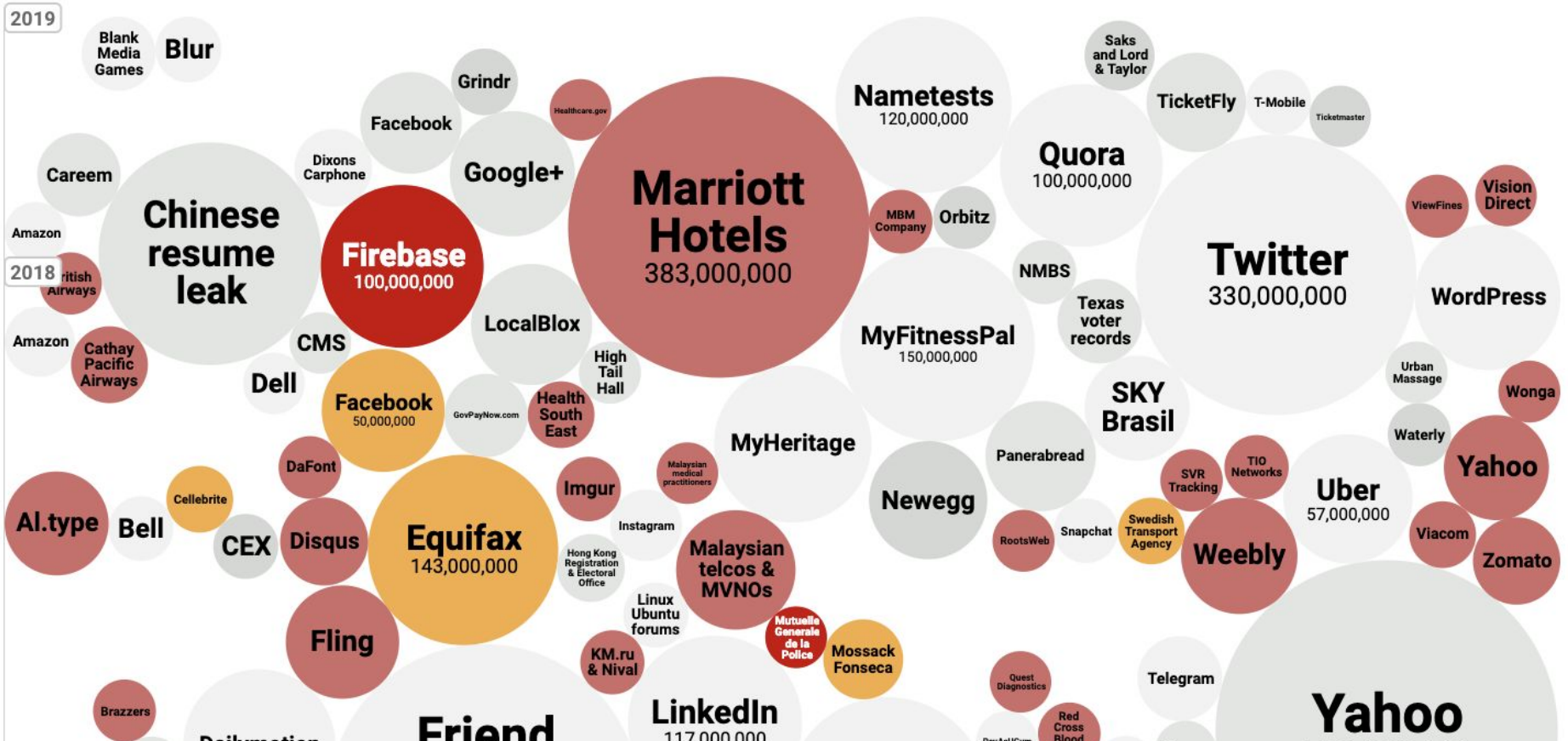
*Select losses greater than 30,000 records*

(updated 1st Feb 2019)

Interesting Story

Colour | YEAR | DATA SENSITIVITY | Filter

Search...

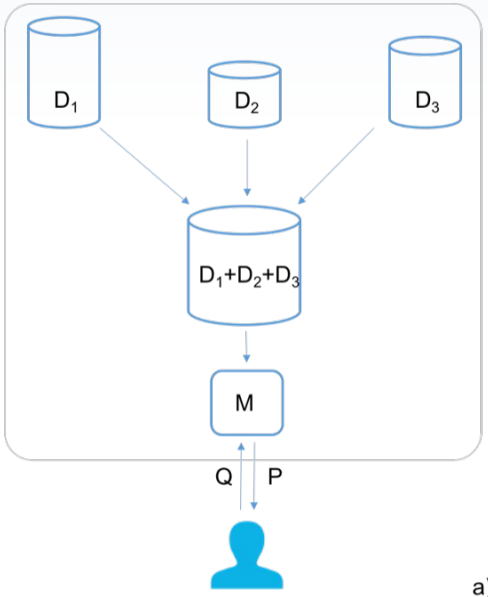


# Daniel Zakrisson
# co-founder and CEO of Scaleout

DIS 2019
-> Booth 34

# The Current Paradigm

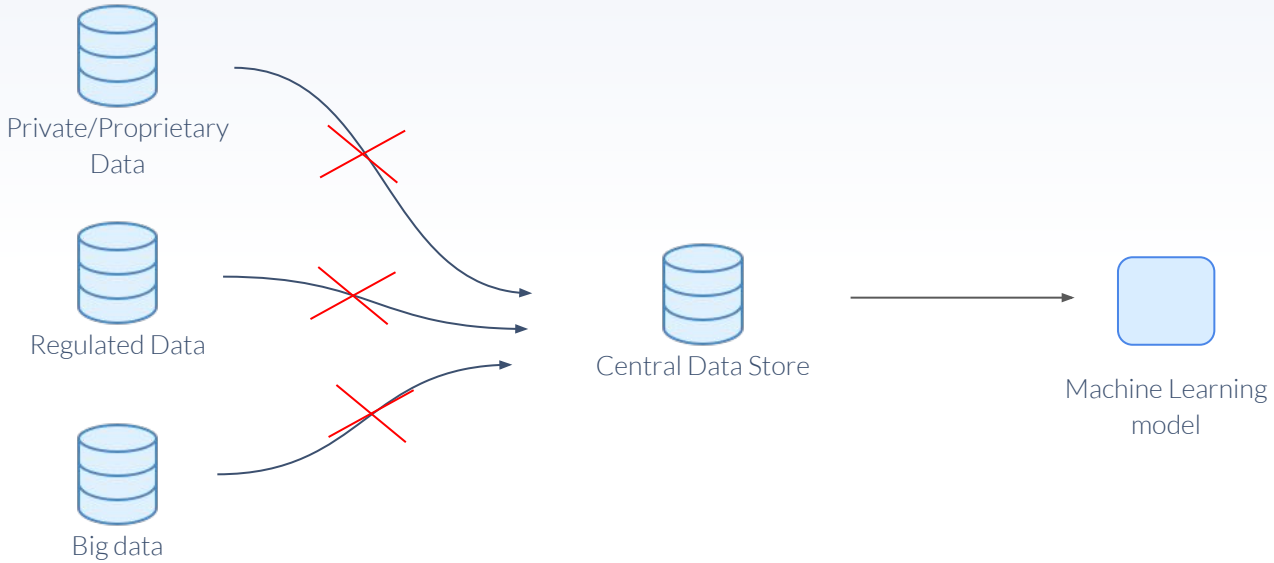
1. Centralise all data
2. Create ML model on centralised data



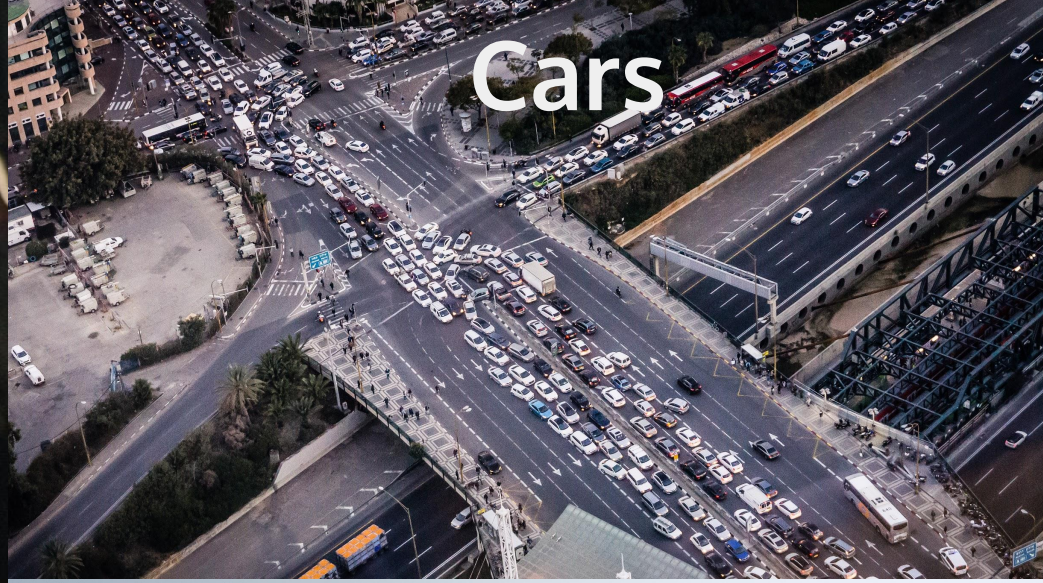

a)

scaleout

# But we cannot move the training data!




Smartphones


Cars

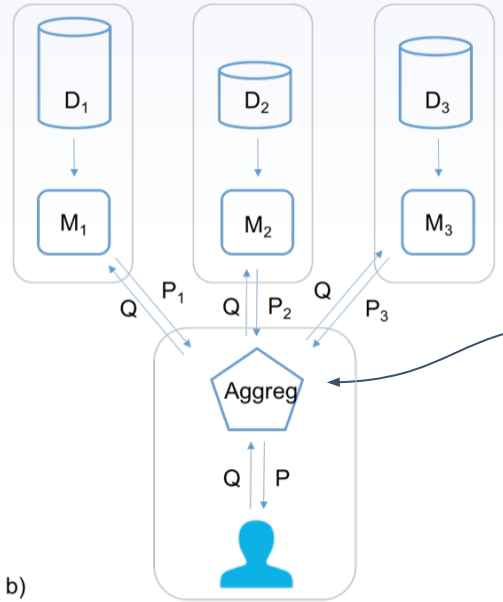

Healthcare


Predictive
maintenance

# How can organizations collaborate on building machine learning models without giving up ownership of their data?

# Federated machine learning: the idea

1. Train a local machine learning model on local data



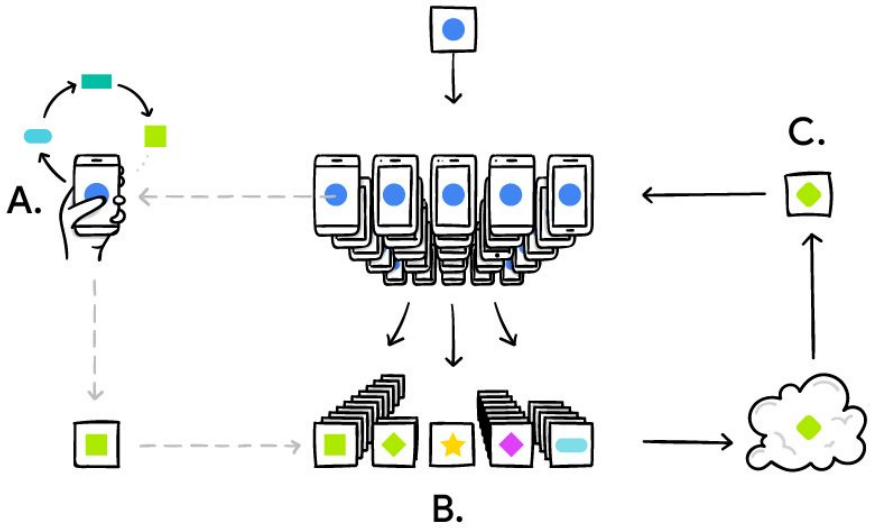

2. Aggregate local models to one central model

scaleout

# First big paper, FedML on gboard

Local model for search suggestion, with context and whether suggestion was clicked

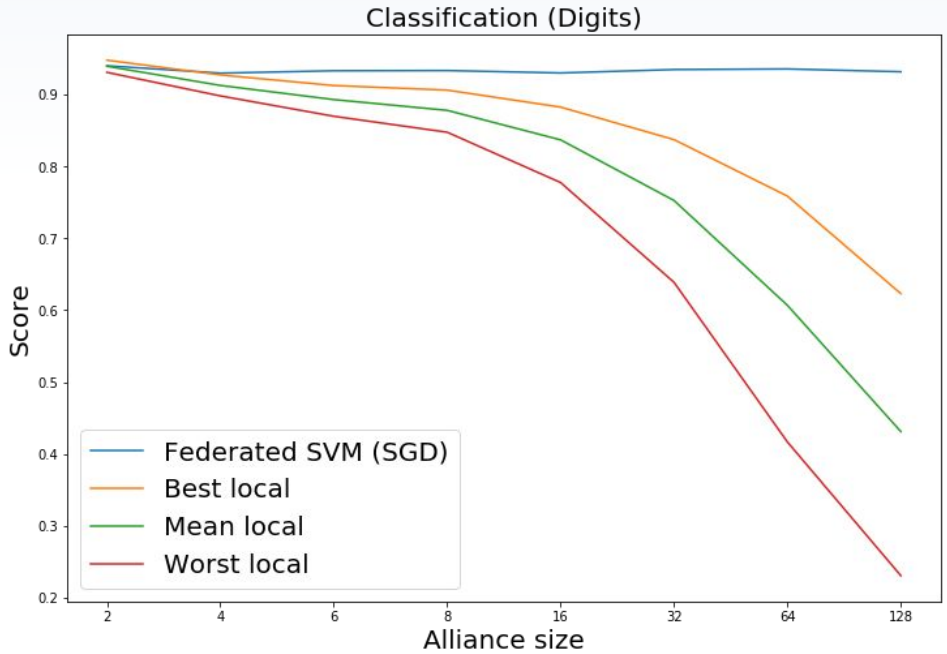
On device the history is processed, and then only a model update is suggested to Google




https://ai.googleblog.com/2017/04/federated-learning-collaborative.html

# Key benefit of federated learning

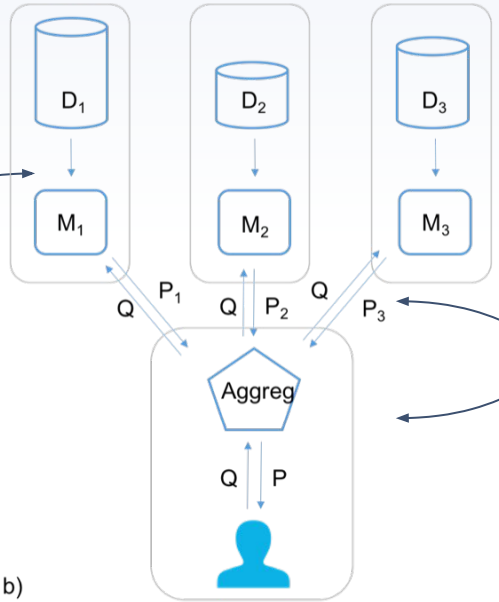
Promises to let parties collaborate to build stronger models than what could be attained by any of the parties in isolation.




Andreas Hellander, CSO Scaleout and Assoc. Prof, UU. Manuscript in preparation.

# Differential privacy & Homomorphic encryption in FedML

Differential privacy
- add noise to data (protects against inference attacks)



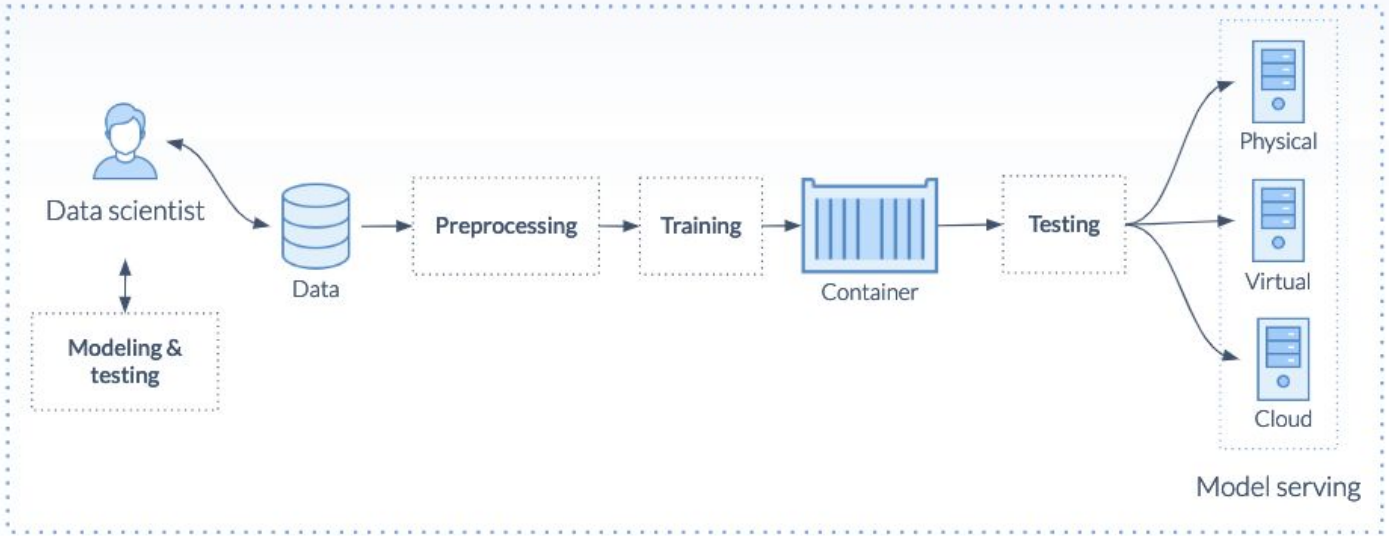

b)

Homomorphic encryption
- encrypts model updates before transit
- perform aggregation on encrypted data

scaleout

# Putting ML models in production

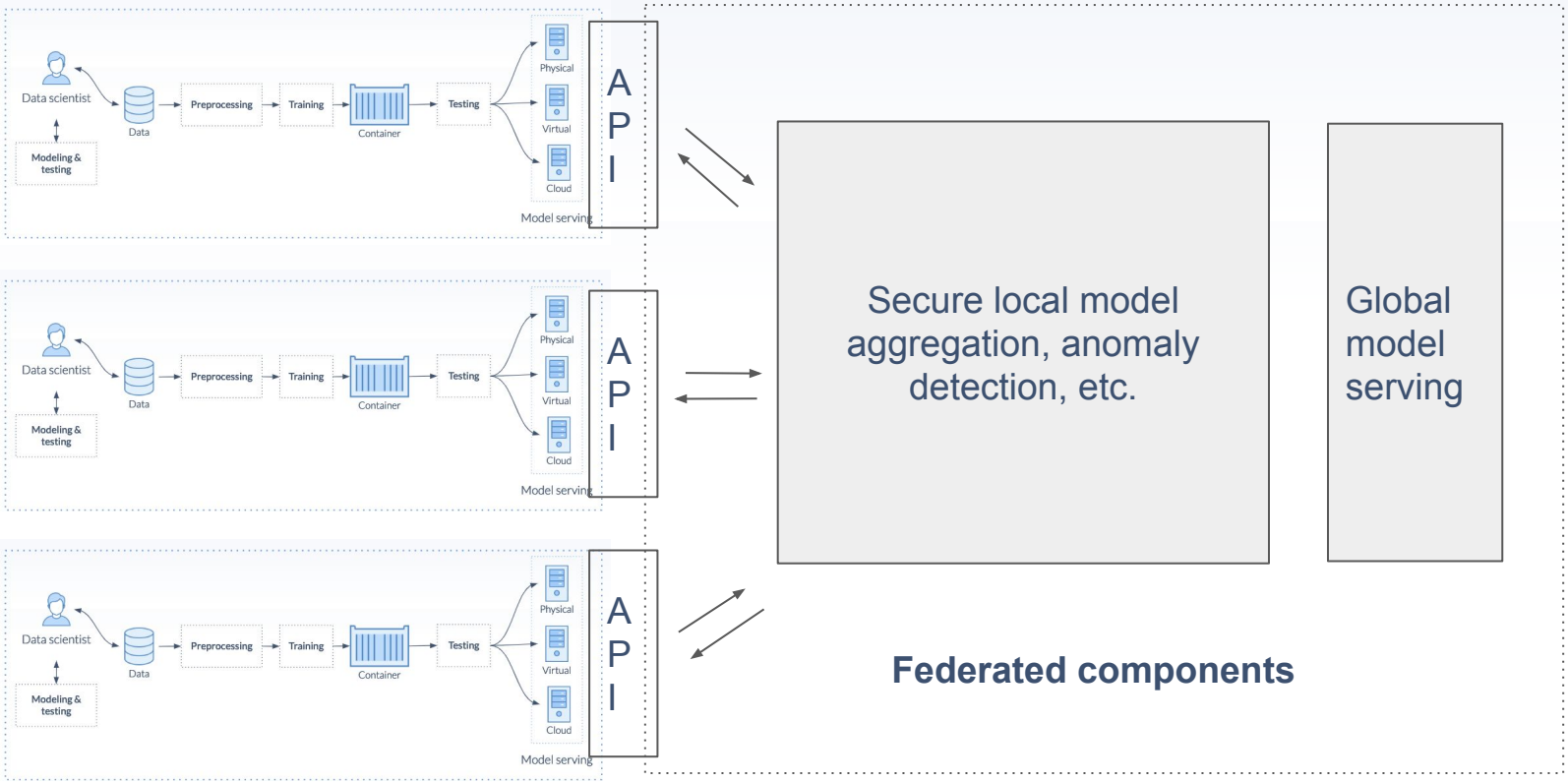
Normal pipeline




Scaleout Lean AI addresses this challenge

scaleout

# Federated learning




scaleout

# Scaleout: Committed to bridging the gap between FedML research and production-grade systems.

# Non-Disclosive Conformal Prediction

- First public demo on March 22 at GE TestaCenter in Uppsala.
- MVP based on Non-Disclosive Conformal Prediction.
    - Shared nothing architecture
    - Conformal classification (prediction sets based on confidence level)
    - Use case from molecular safety assessment in drug discovery

Ola Spjuth, lead scientist AI at Scaleout, and Assoc. Prof. at UU.

https://arxiv.org/pdf/1806.04000.pdf

scaleout


Thank you!

Email: daniel@scaleout.se
LinkedIn: linkedin.com/in/danielzakrisson

Data Innovation Summit March 14, 2019