



Hewlett Packard
Enterprise



LERNEN IM «SCHWARM» & WISSEN VERTEILEN

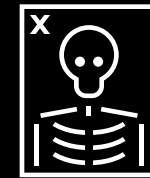
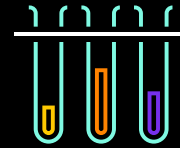
SWARM LEARNING FOR DECENTRALIZED AND CONFIDENTIAL CLINICAL MACHINE
LEARNING

Roger Fontana

November 2021

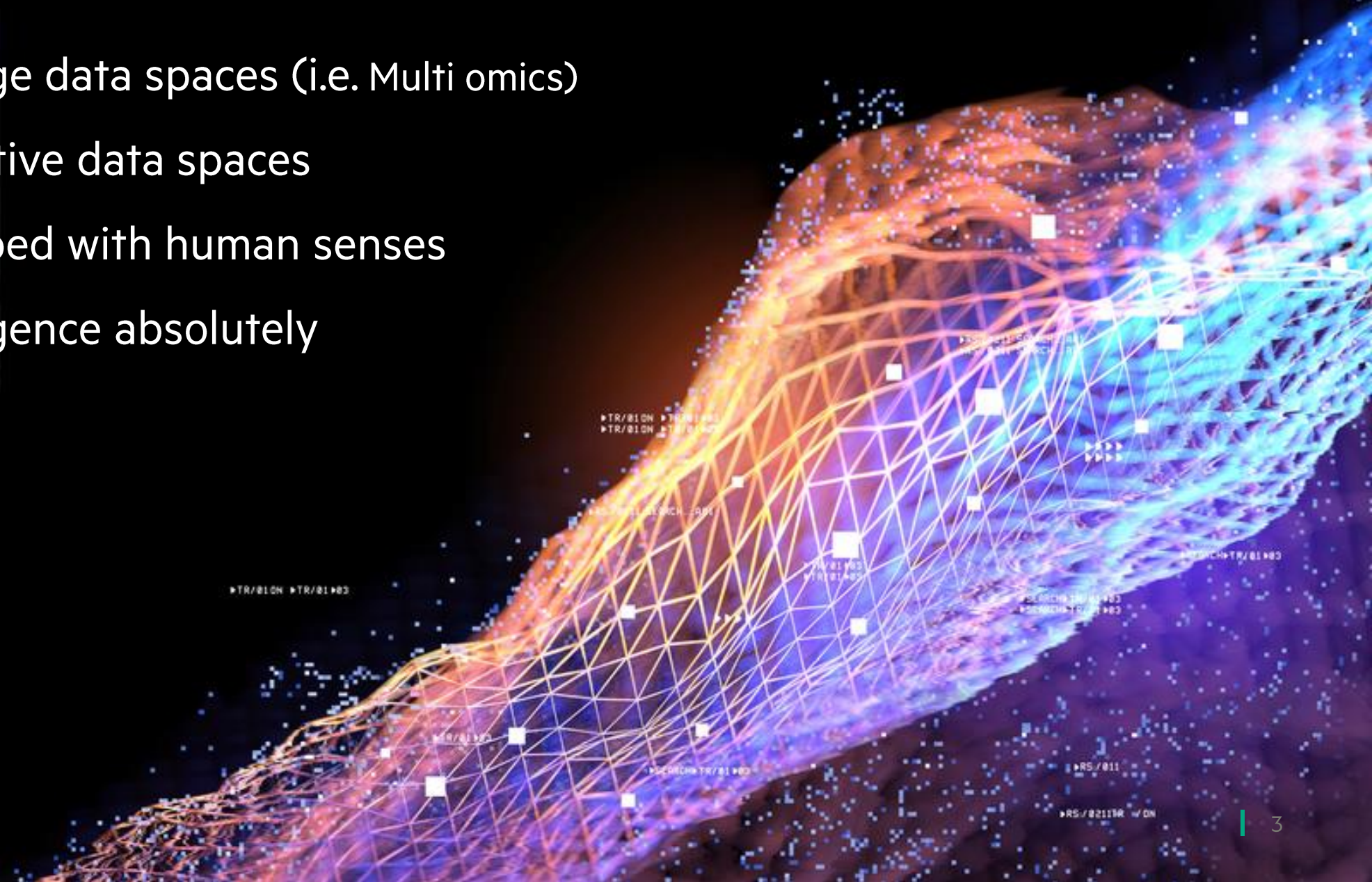
DOCTORS READ PATTERNS FROM DATA WHEN MAKING DIAGNOSES

- Medical history
- Physical examination
- X-ray images
- Blood values
- Laboratory values
- Pathology
- Computer tomography
- Magnetic resonance imaging
- ECG, EEG
- etc.



AI AND - PRIVATE - DATA SPACES

- Particularly large data spaces (i.e. Multi omics)
- Highly informative data spaces
- Not to be grasped with human senses
- Artificial intelligence absolutely necessary



AI/ML IN HEALTHCARE – IMPROVE MODELS – NEED FOR DATA

The GOOD

AI to predict breast cancer and personalize care

*“Rather than manually identifying the patterns in a mammogram that drive future cancer, the MIT/MGH team trained a deep-learning model to deduce the patterns directly from the data. Using information from more than **90,000 mammograms**, the model detected patterns **too subtle for the human eye to detect.**”*

The BAD

Doctors use algorithms that aren't designed to treat all patients equally

- "Otherwise, you're creating a scientific way of justifying the unequal distribution of resources."
- "an algorithm without data is useless"

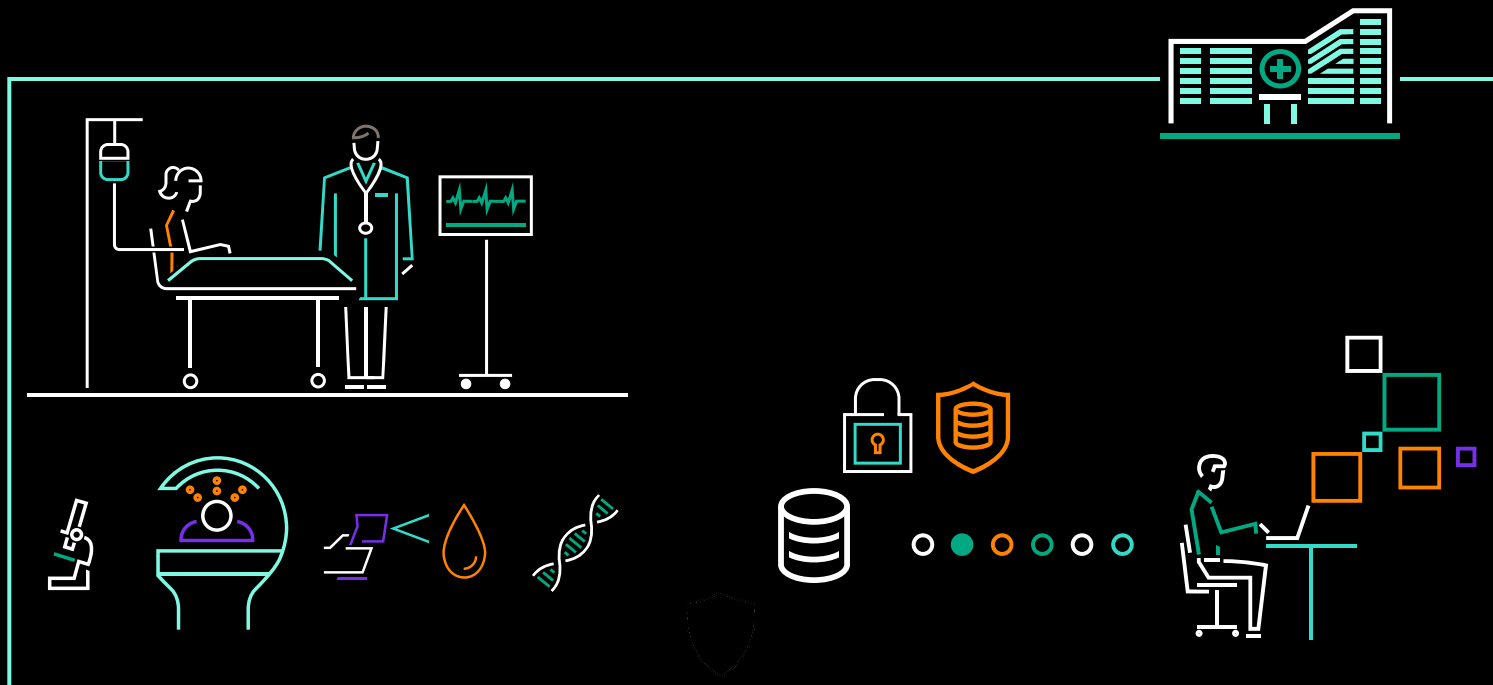
The NEED



to reduce BIAS



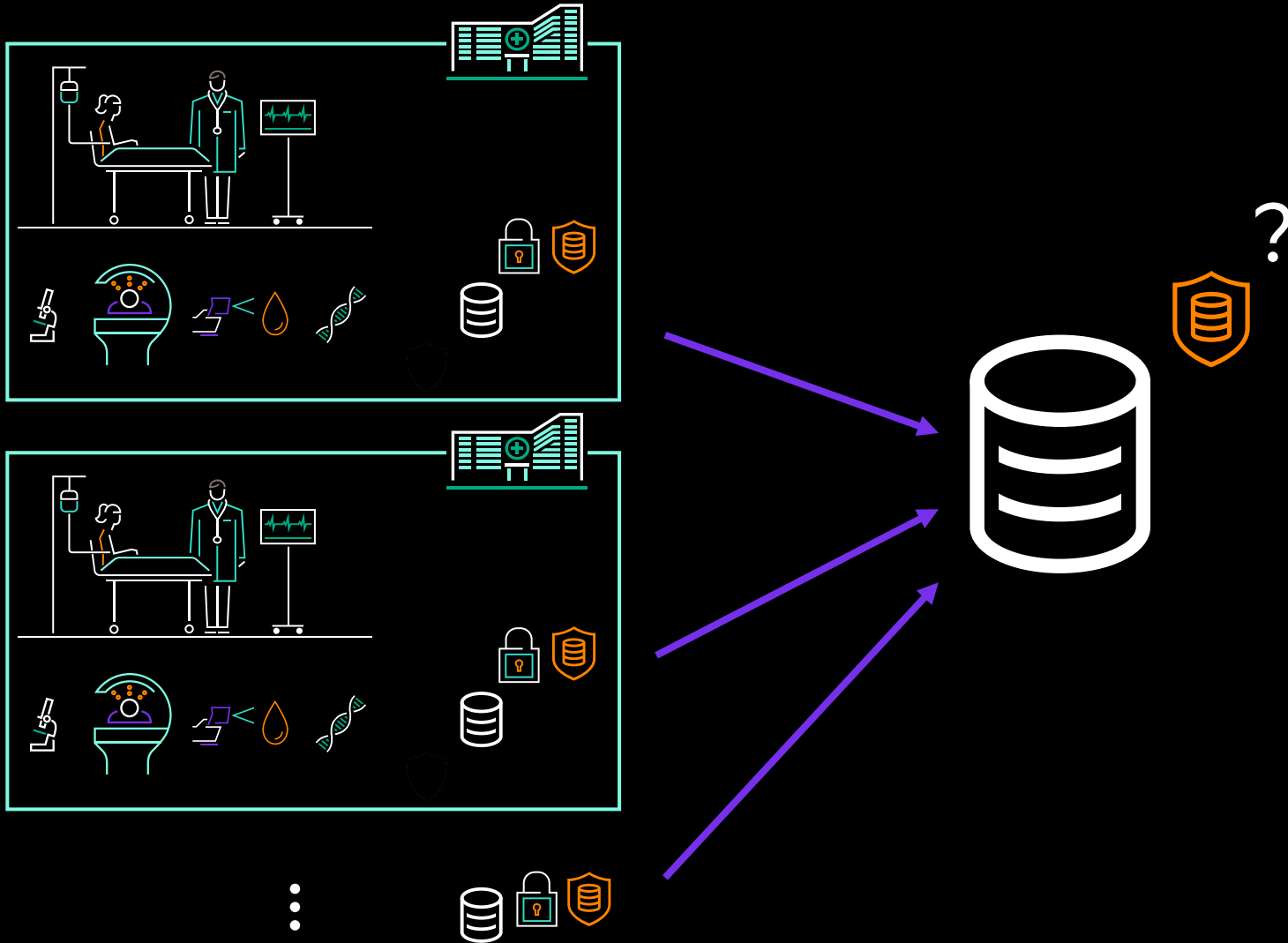
DATA IN HEALTHCARE – SINGLE



- limited Dataset
- my data
- my model
- cannot learn with others



DATA IN HEALTHCARE – SHARE DATA?



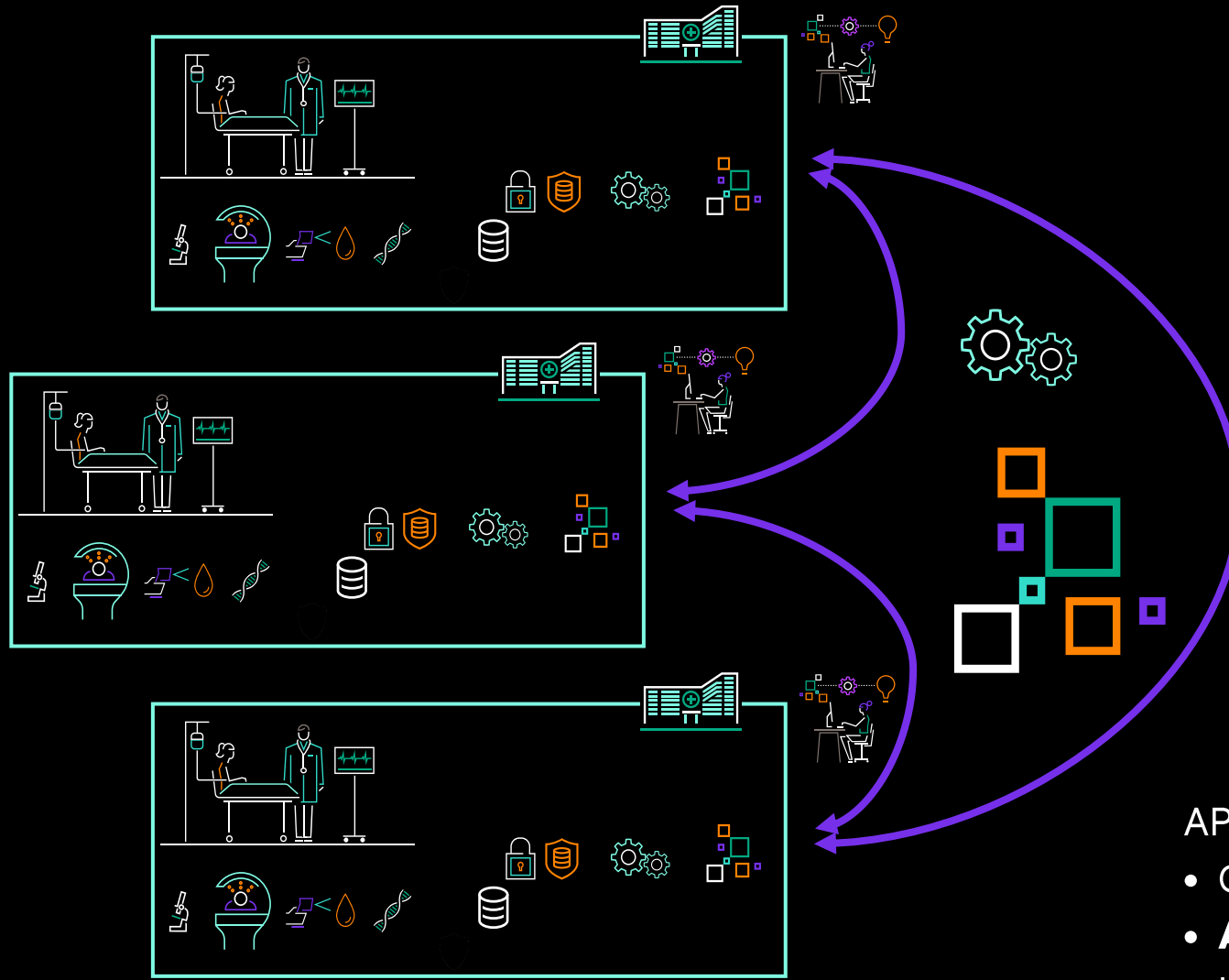
- Willing to share?
- Allowed to share?
 - GDPR
 - HIPAA
 - Consumer Privacy Act (CCPA)
- Agree on Data Sharing governance

- Data privacy
- Data movement & duplication

- **Data sharing, open data, centralized data** are concepts that are **diametrically** opposed to medical traditions



SWARM LEARNING IN HEALTHCARE – CHANGE APPROACH



- share insight / results
- do not move data
- do not duplicate data
- keep privacy & sovereignty
- enlarge dataset

- decentralized learning

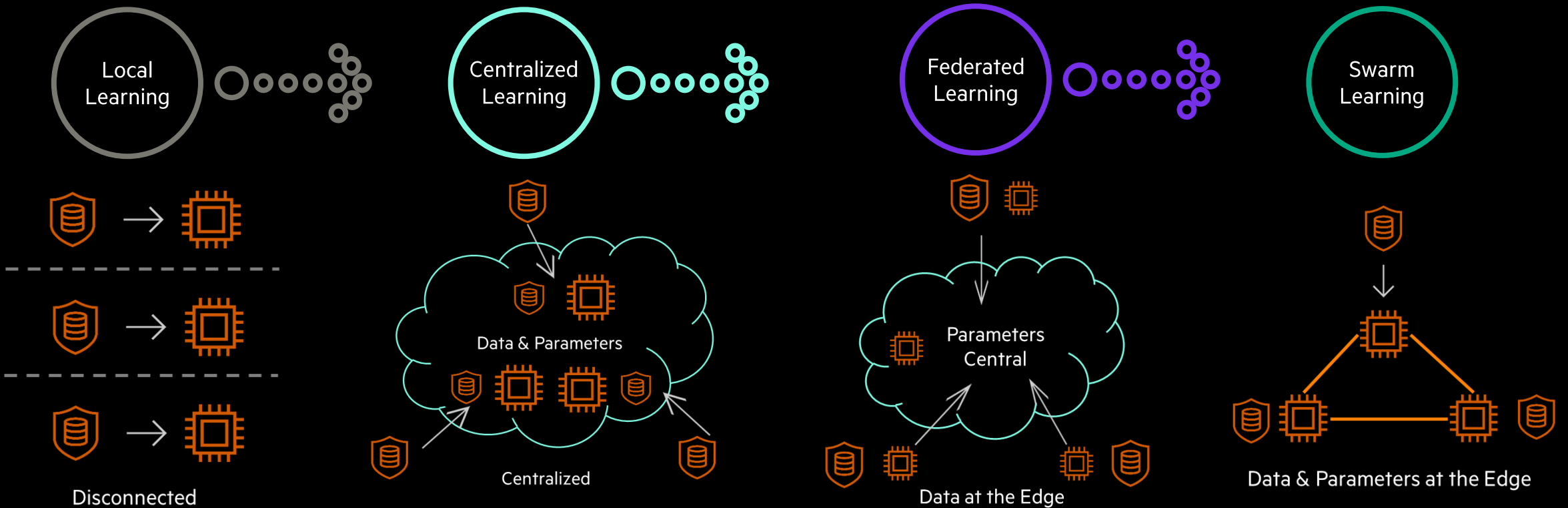
From classical medicine to systems medicine (pattern recognition at scale)

- AI as the basis for precision medicine, in line with medical traditions and habits
- make private data clinically usable

APPROACH

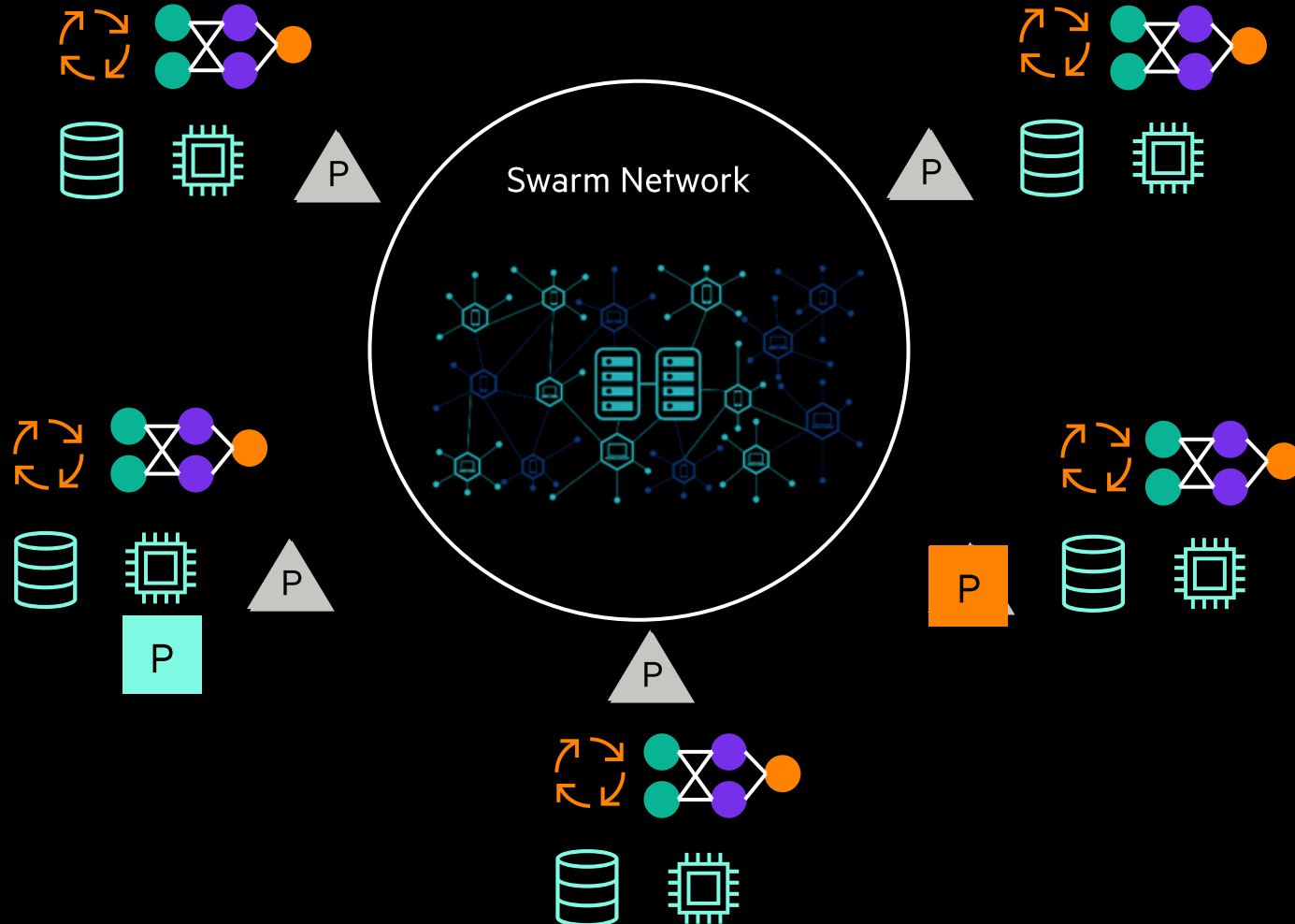
- Customer has (one or many) AI Model(s)
- **Automated decentralization** with Swarm Learning aaS
- keep **privacy** and **sovereignty** on a worldwide swarm

MACHINE LEARNING FRAMEWORK JOURNEY



Swarm Learning enables privacy-preserving, collaborative machine learning by treating all participants equally

SWARM LEARNING – FLOW



1. Register

Nodes register to Swarm Network and receive ML model

2. Train

Nodes train the model on local data for a time-window (epoch)

3. Merge

Nodes share and merge the trained models

4. Repeat

Repeat 2 & 3 until desired accuracy is achieved

DOES IT WORK?



Article

nature

Explore content ▾ Journal information ▾ Publish with us ▾

nature > articles > article

Article | Open Access | Published: 26 May 2021

Swarm Learning for decentralized and confidential clinical machine learning

Stefanie Warnat-Herresthal, Hartmut Schultze, [...]Joachim L. Schultze

Nature 594, 265–270 (2021) | Cite this article

29k Accesses | 378 Altmetric | Metrics

Abstract

Fast and reliable detection of patients with severe and heterogeneous illnesses is a major goal of precision medicine^{1,2}. Patients with leukaemia can be identified using machine learning on the basis of their blood transcriptomes³. However, there is an increasing divide between what is technically possible and what is allowed, because of privacy legislation^{4,5}. Here, to facilitate the integration of any medical data from any data owner worldwide without violating privacy laws, we introduce Swarm Learning—a decentralized machine-learning approach that unites edge computing, blockchain-based peer-to-peer networking and coordination while maintaining confidentiality without the need for a central coordinator, thereby going beyond federated learning. To illustrate the feasibility of using Swarm Learning to develop disease classifiers using distributed data, we chose four use cases of heterogeneous diseases (COVID-19, tuberculosis, leukaemia and lung pathologies). With more than 16,400 blood transcriptomes derived from 127 clinical studies with non-uniform distributions of cases and controls and substantial study biases, as well as more than 95,000 chest X-ray images, we show that Swarm Learning classifiers outperform those developed at individual sites. In addition, Swarm Learning completely fulfils local confidentiality regulations by design. We believe that this approach will notably accelerate the introduction of precision medicine.

[Swarm Learning for decentralized and confidential clinical machine learning | Nature](#)

Research Highlight

nature medicine

Explore content ▾ Journal information ▾ Publish with us ▾ Subscribe

nature > nature medicine > research highlights > article

RESEARCH HIGHLIGHT | 10 June 2021

A safer way to share health data

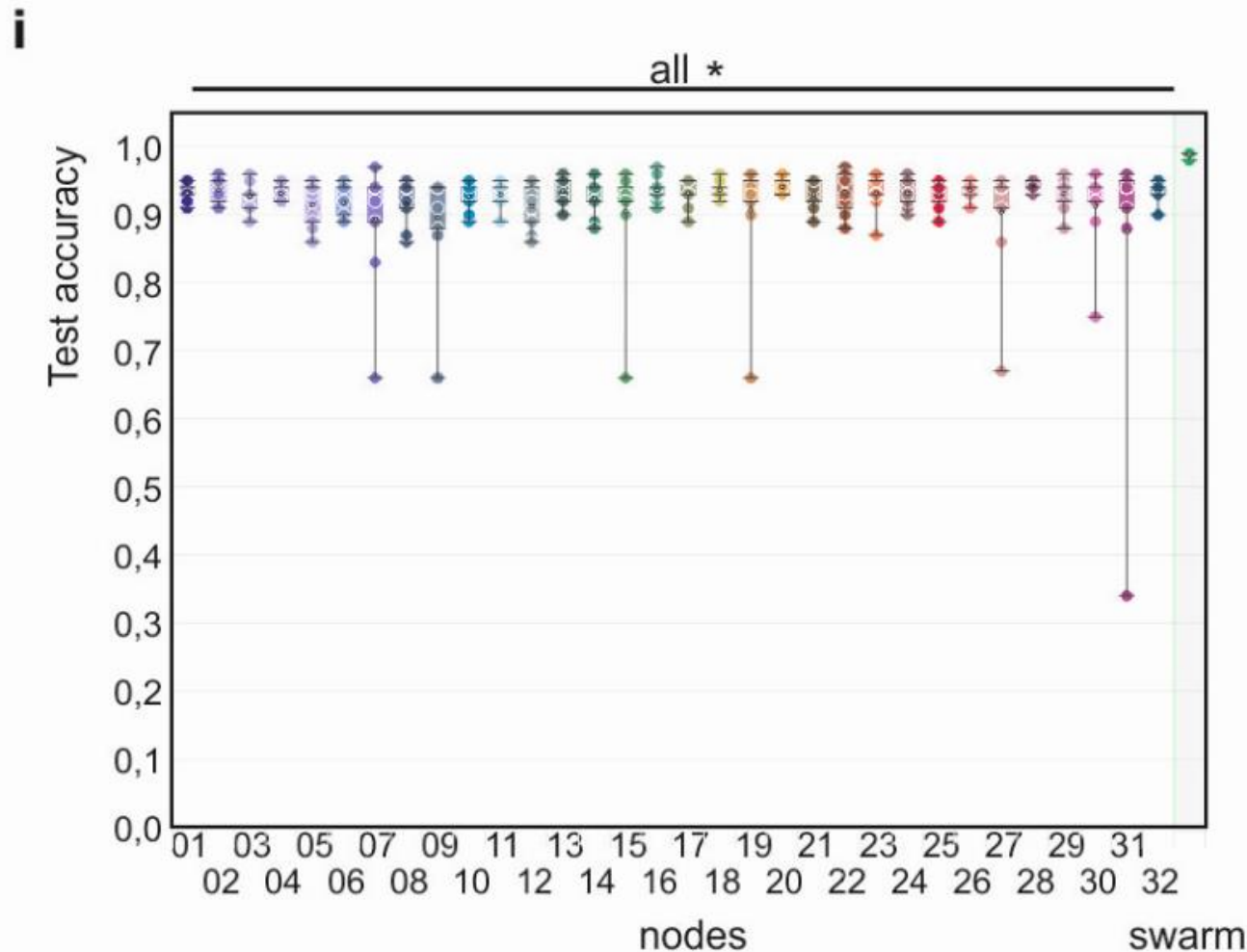
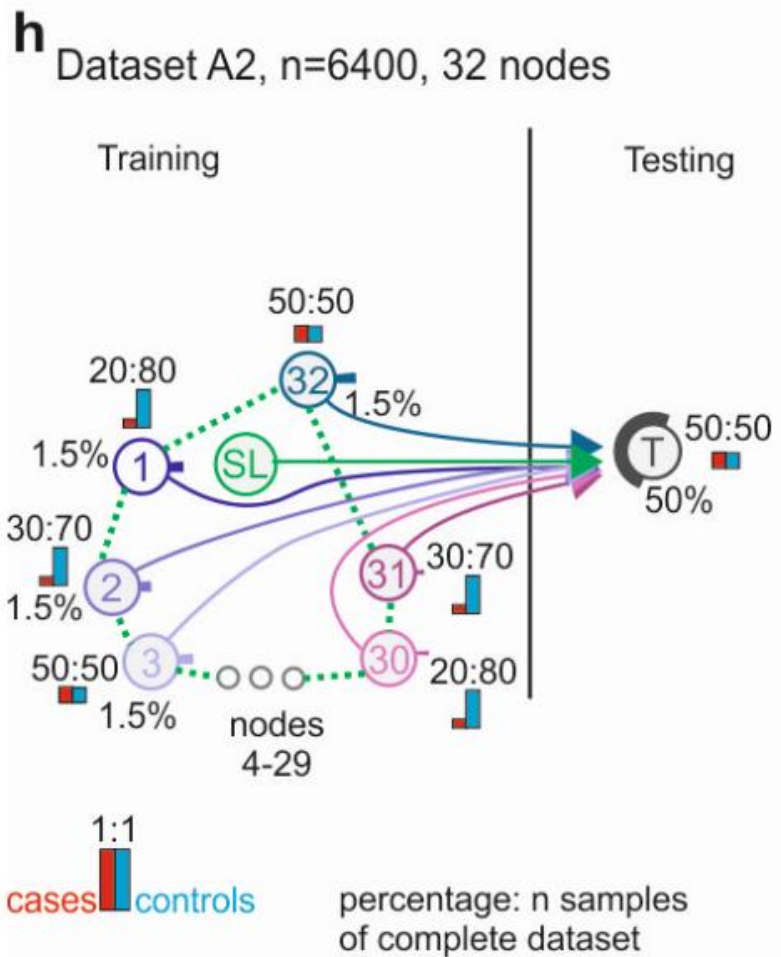
A 'swarm learning' technique enables AI-based diagnosis with enhanced data security and confidentiality.

Karen O'Leary

Artificial intelligence (AI)-based tools are becoming ever more proficient at clinical decision-making.

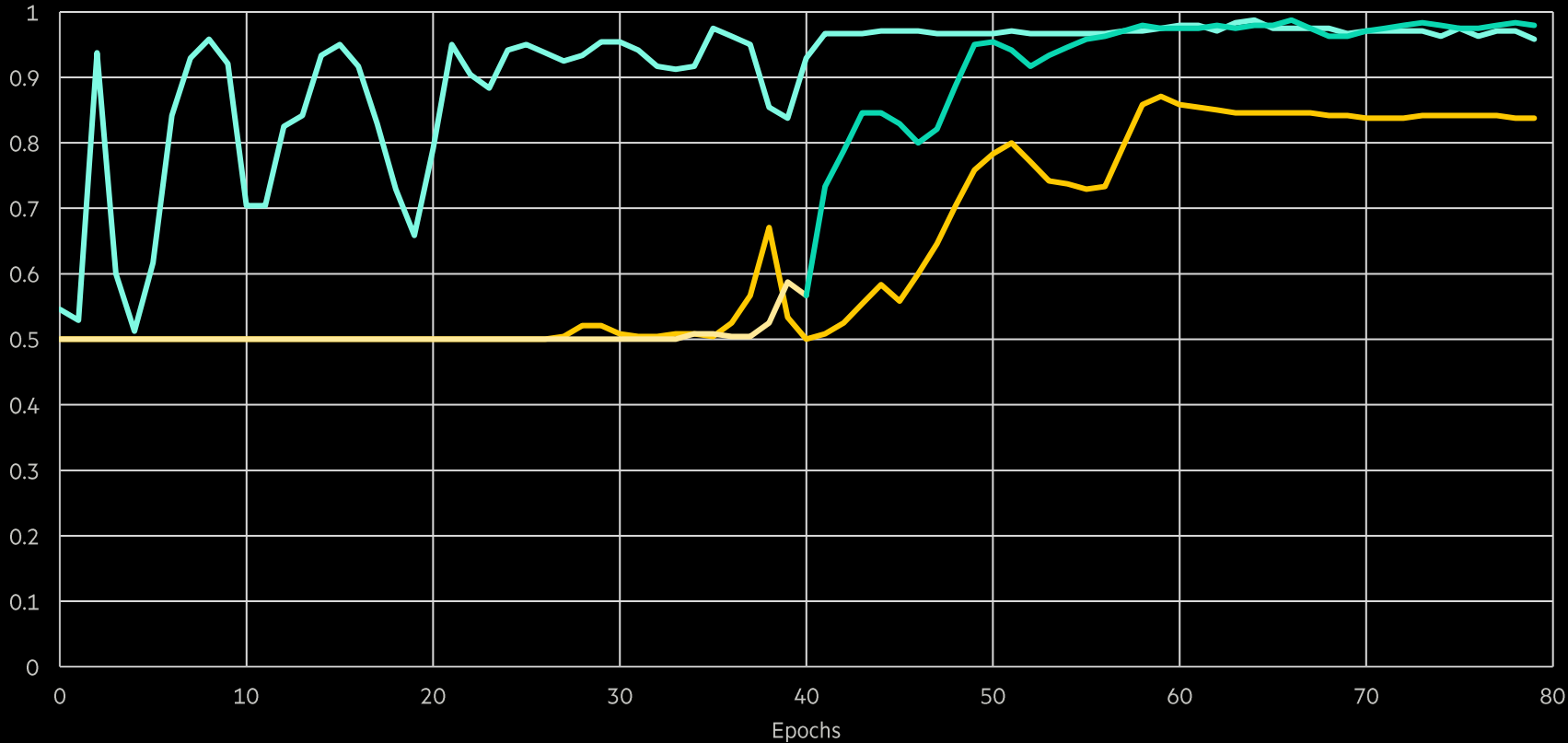
[A safer way to share health data \(nature.com\)](#)

MULTIPLE NODES



DYNAMIC NODE ADDITION – CONTINUOUS LEARNING

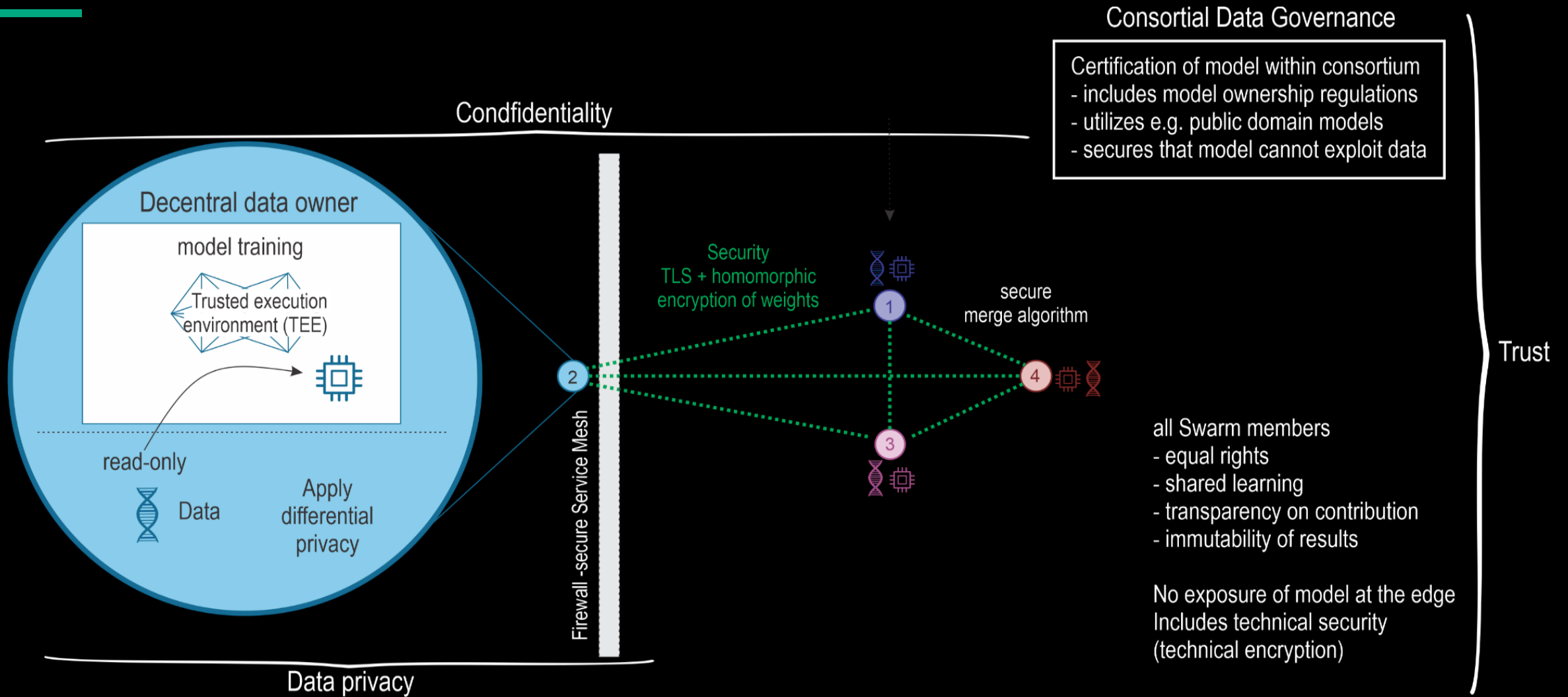
Swarm Learning with dynamic node addition



- (a) train with 3 low prevalence nodes and hence achieves lower accuracy
- (b) train with 3 high prevalence nodes and achieves higher accuracy
- (c) 6 node scenario, starts with 3 low prevalence nodes, 3 additional nodes added around 40th epoch. Learning improves at this point and achieves higher accuracy



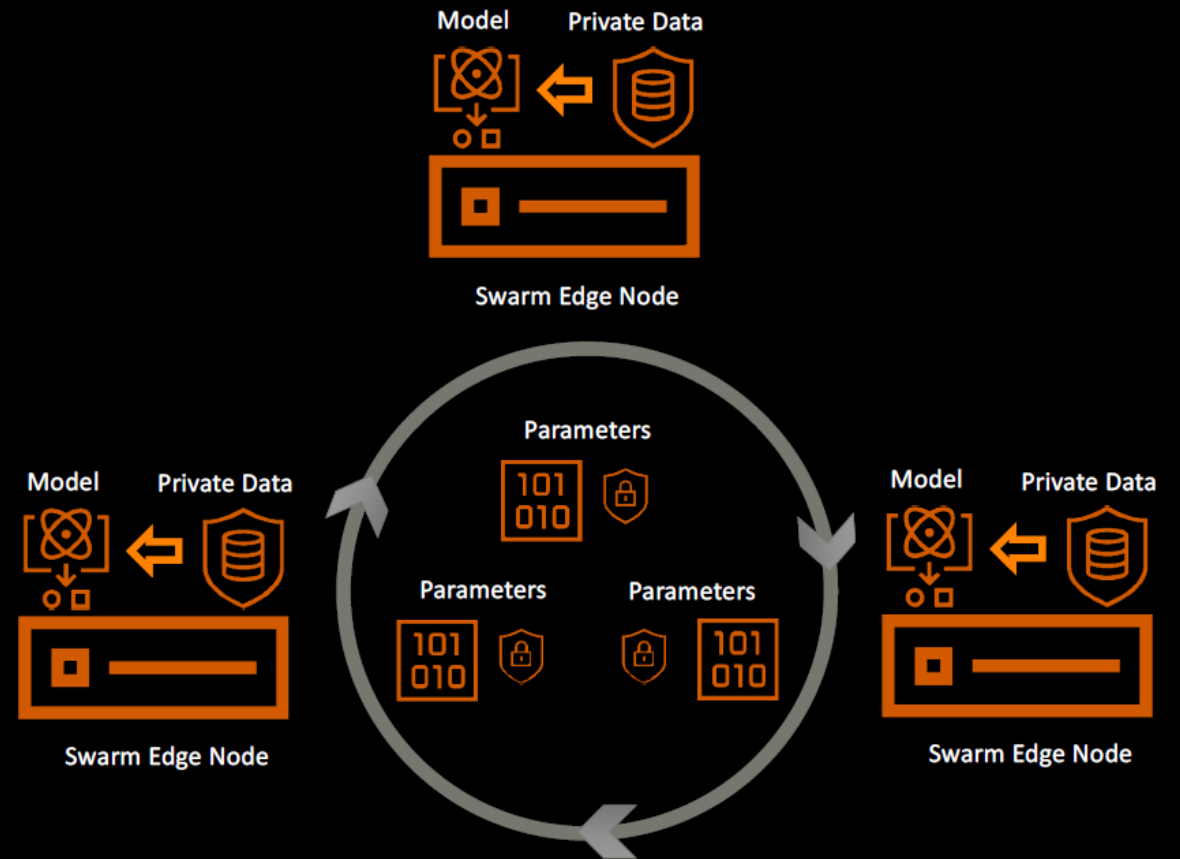
CONSORTIAL DATA GOVERNANCE



SWARM LEARNING

Democratic Machine Learning

- Equal and like-minded partners in the network
- Data protection and data security applied directly at the source
- Larger dataset accessible with more data to be learned on
 - Less susceptible to bias in machine learning
 - Higher resiliency



[GitHub - HewlettPackard/swarm-learning](https://github.com/HewlettPackard/swarm-learning)

THANK YOU

Roger Fontana

roger@hpe.com

+41 79 401 08 83

