

Fórum GDI – Data Talk 20/04/2023

Data Science em Medicina: Privacidade na Era de Big Data e Inteligência Artificial

José Miguel Diniz

Médico Interno de Saúde Pública @ ACeS Lisboa Central

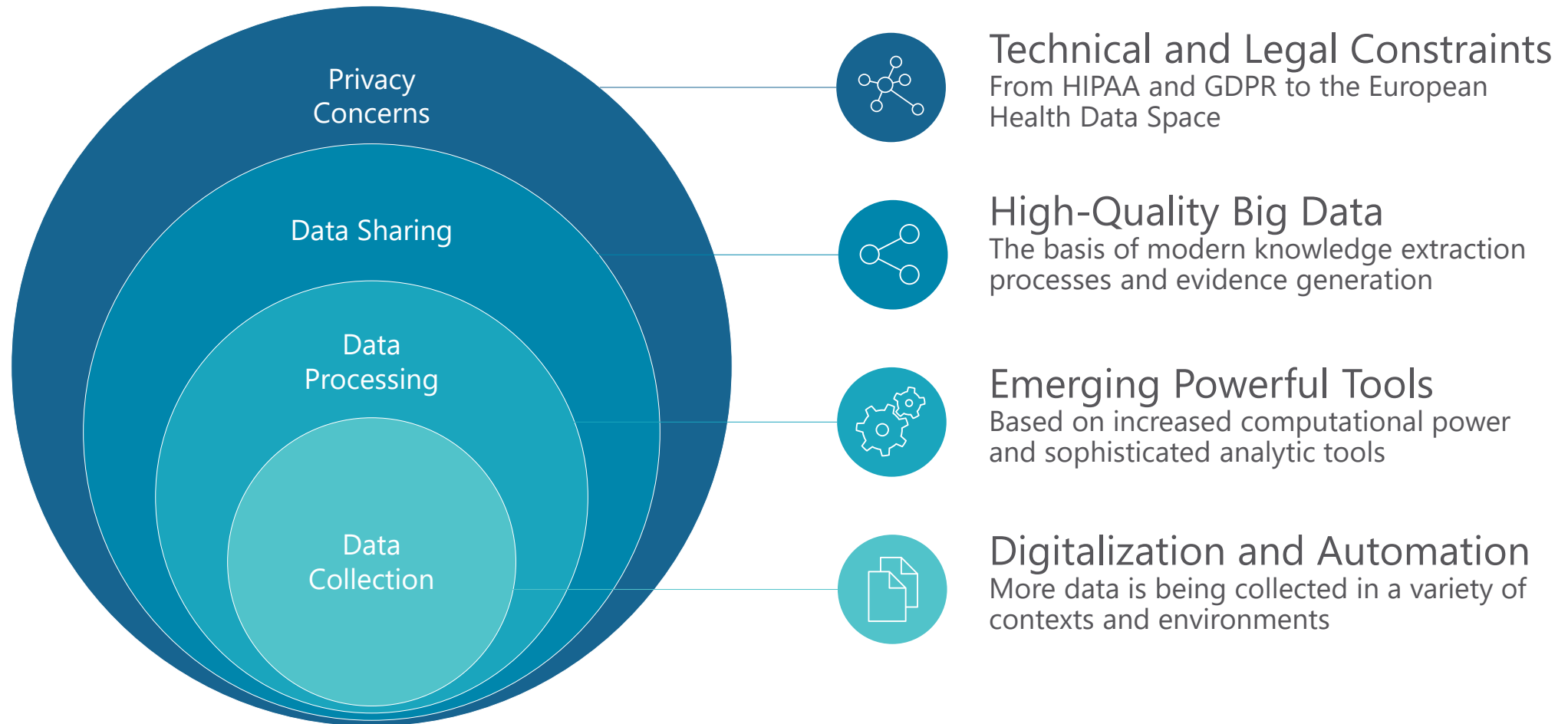
Candidato a PhD em Health Data Science @ FMUP

Docente Voluntário no Mestrado Integrado em Medicina @ FMUP

Investigador @ CINTESIS/FMUP

(Health) Data Science Paradigm

Status Quo and Challenges



Health (Data Science) Paradigm

Planning Meaningful Research



Underlying Principles

FAIR PRINCIPLES

The principles refer to three types of entities: data (or any digital object), metadata (information about that digital object), and infrastructure.

01

Findable

Metadata and data should be easy to find for both humans and computers.

02

Accessible

Once the user finds the required data, they need to know how they can be accessed, possibly including authentication and authorisation.

03

Interoperable

Data usually need to be integrated with other data or to integrate with applications or workflows for analysis, storage, and processing.

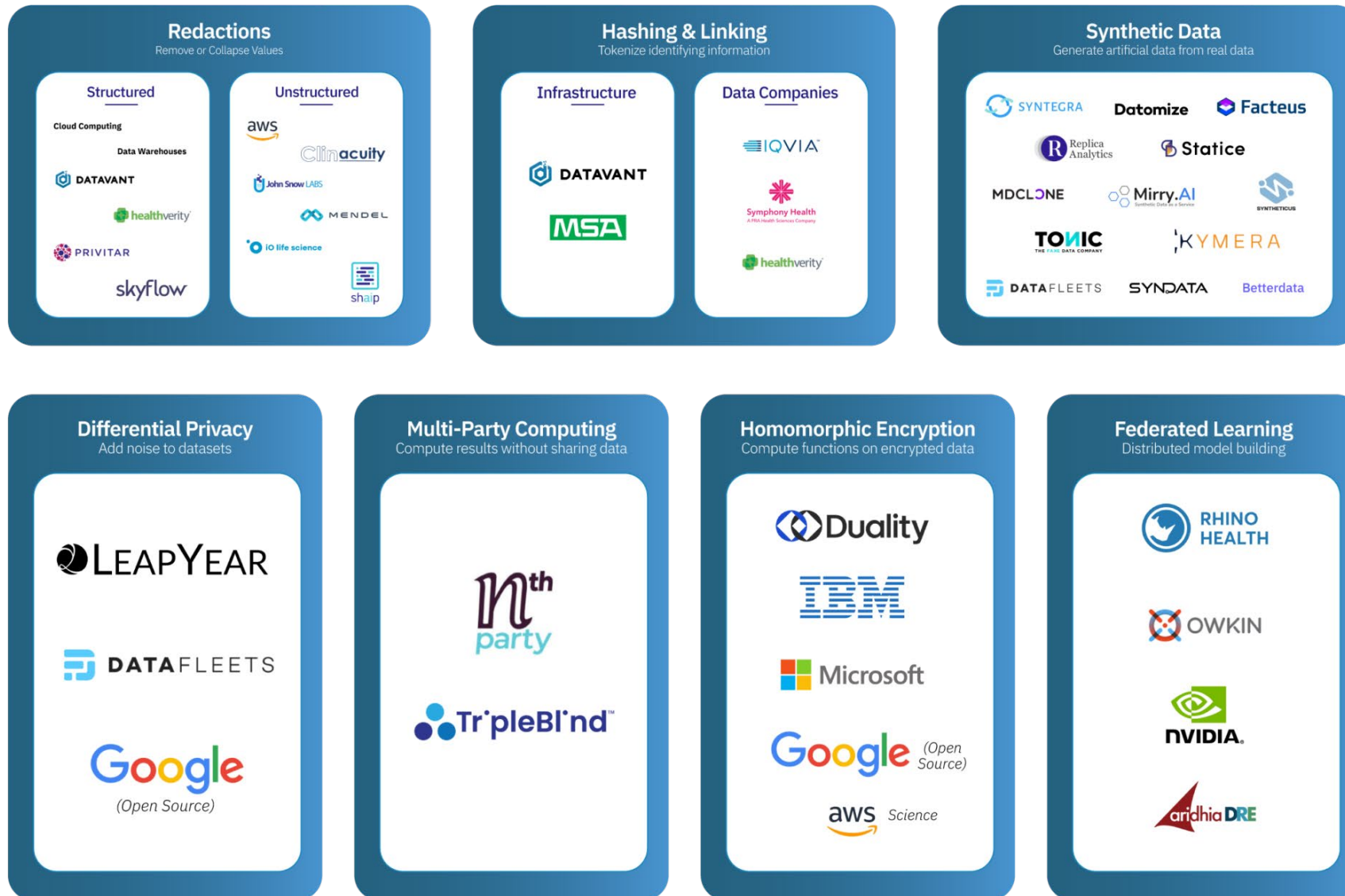
04

Reusable

Metadata and data should be well-described so that they can be replicated and/or combined in different settings.

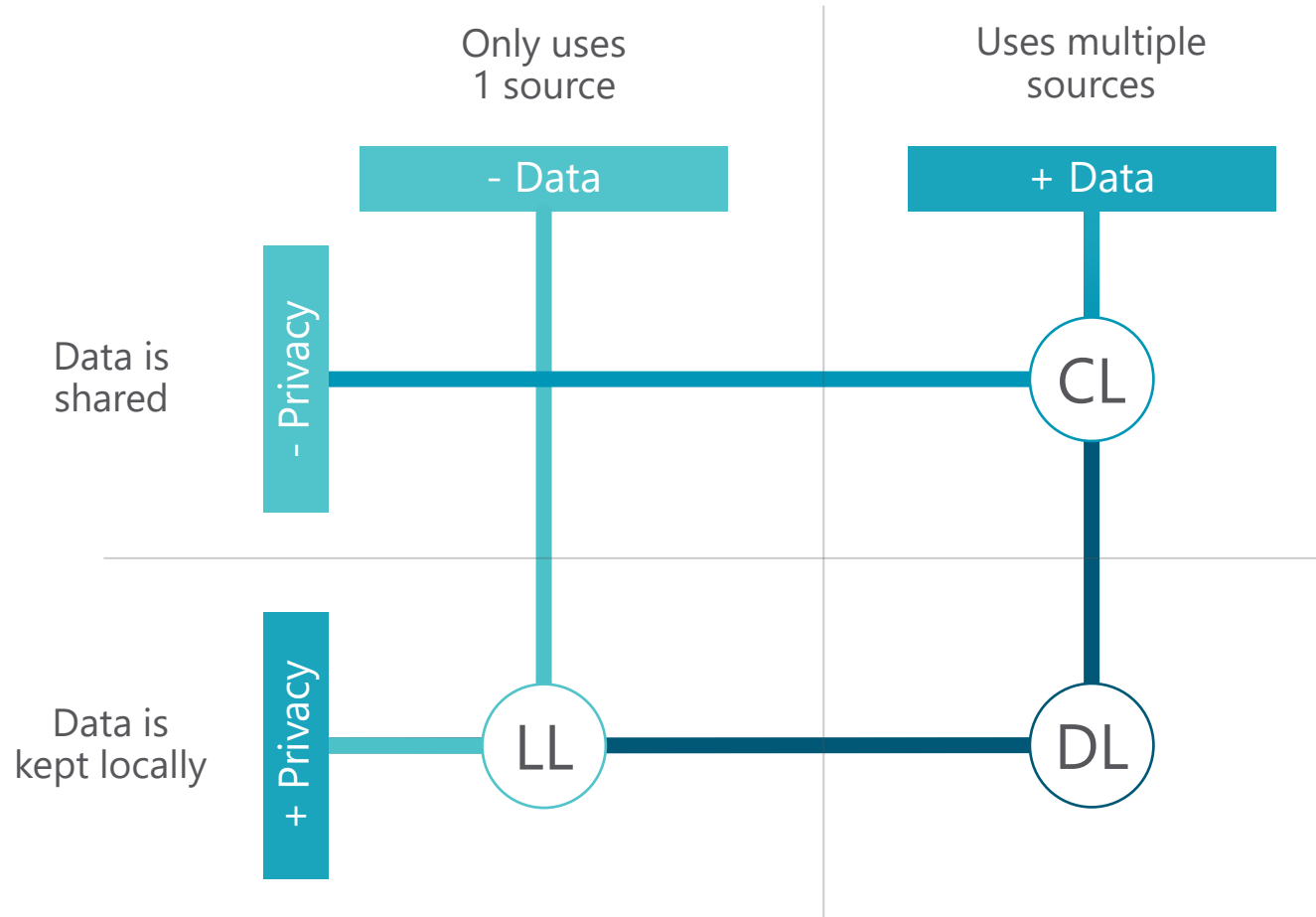
Personalization-Privacy Paradox



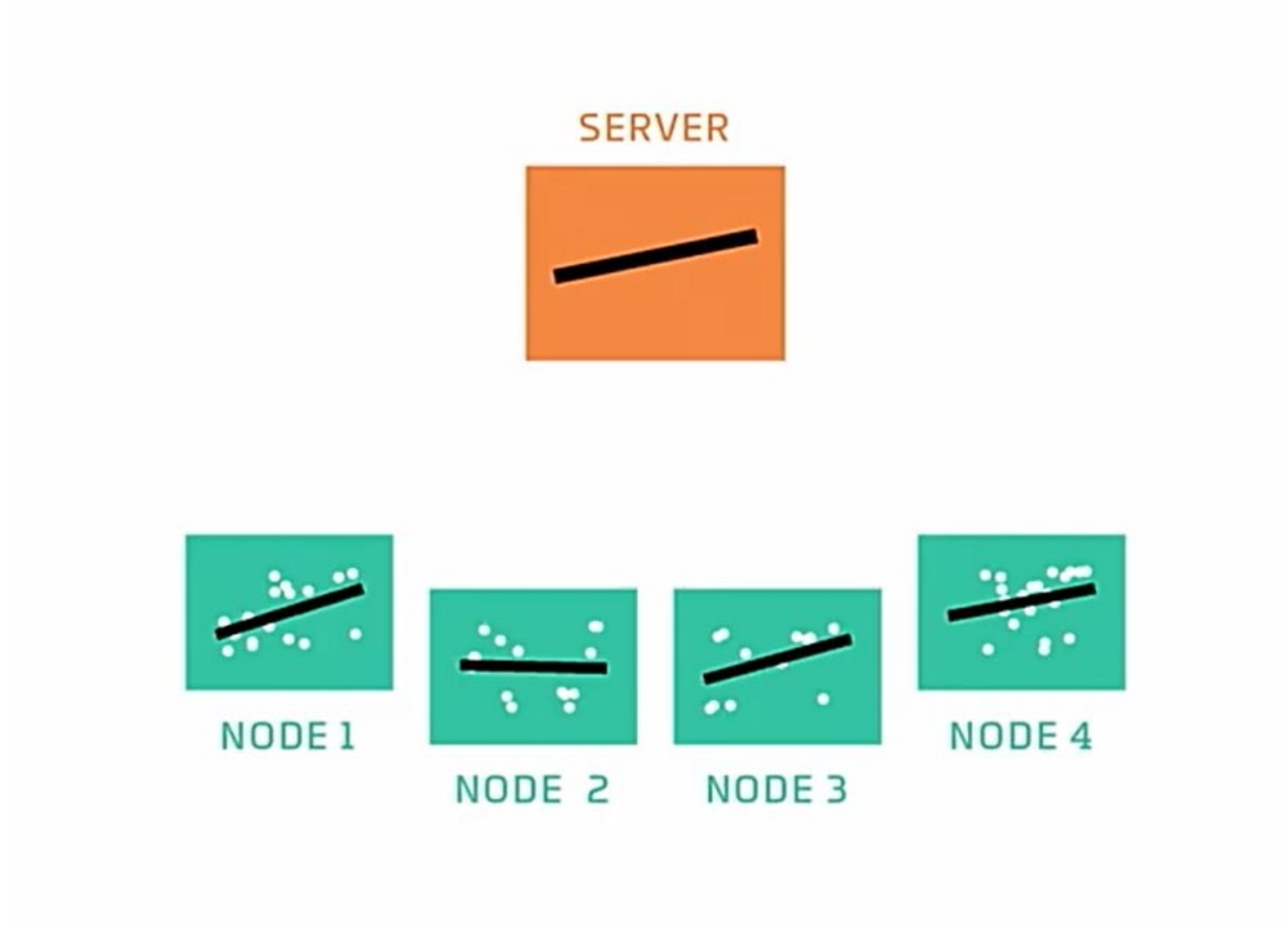


in <https://datavant.com/resources/blog/an-overview-of-approaches-to-privacy-preserving-data-sharing/>

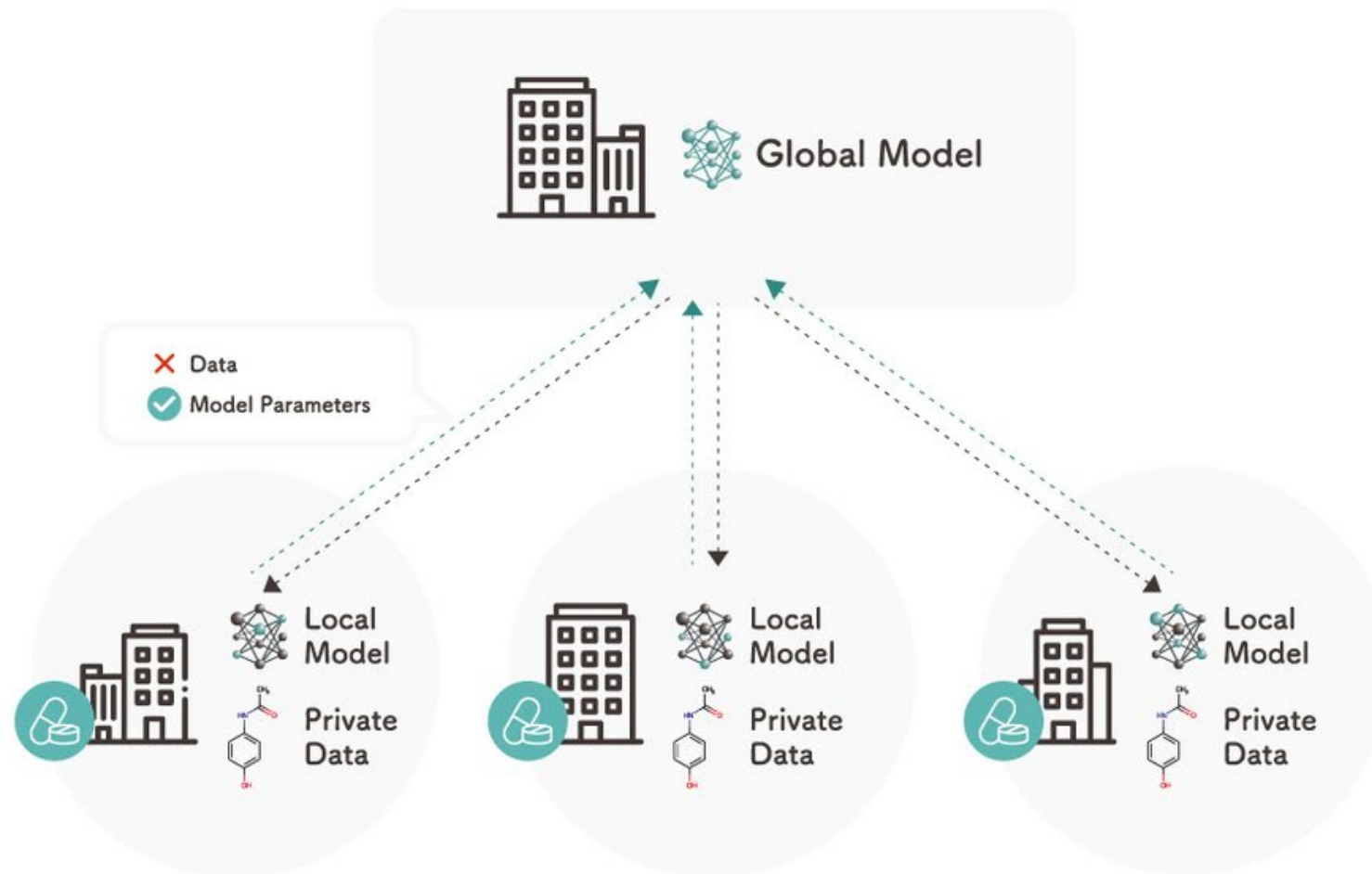
Distributed Learning Systems



- Data is analysed in a distributed way
- Only data from model parameters is shared
- Mathematically enforced privacy preservation, auditable data transactions



Decentralized Learning: A Federated Approach



Secur-e-Health

Privacy preserving cross-organizational data analysis in the healthcare sector



**Becoming the
trusted open science community
built with standardised health data
via a European federated network**

**EHDEN Academy: discover
our latest course for non-
experts**

**Voice of EHDEN season 3
episode 1 now available!**



Flower A Friendly Federated Learning Framework

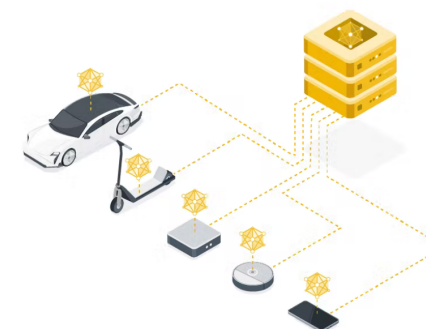
A unified approach to federated learning, analytics, and evaluation. Federate any workload, any ML framework, and any programming language.

`pip install flwr`

[Copy](#)

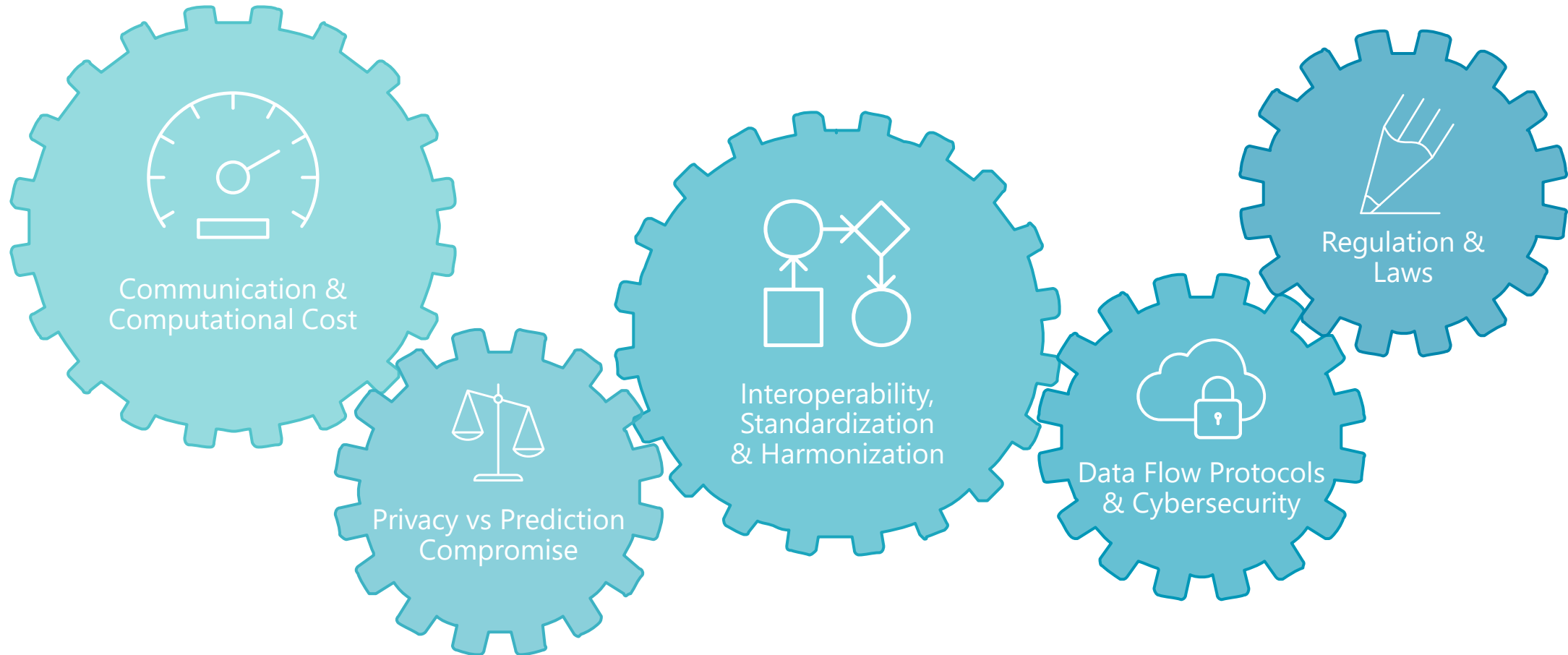
[Documentation](#)

[GitHub](#)



Discussion

Active Challenges & Upcoming Work



Conclusions

- Despite growing global health challenges, **research that fosters and empowers planning** will be of the uttermost relevance.
- To do so, we often need high quality Big Data to leverage powerful algorithms, which are internally and externally valid.
- To solve the conflict between privacy and better performing health data models, new **privacy preserving technologies** should be adopted.

Fórum GDI – Data Talk 20/04/2023

Data Science em Medicina: Privacidade na Era de Big Data e Inteligência Artificial

José Miguel Diniz

Médico Interno de Saúde Pública @ ACeS Lisboa Central

Candidato a PhD em Health Data Science @ FMUP

Docente Voluntário no Mestrado Integrado em Medicina @ FMUP

Investigador @ CINTESIS/FMUP

jmdiniz.com/linkedin

jmdiniz.med@gmail.com