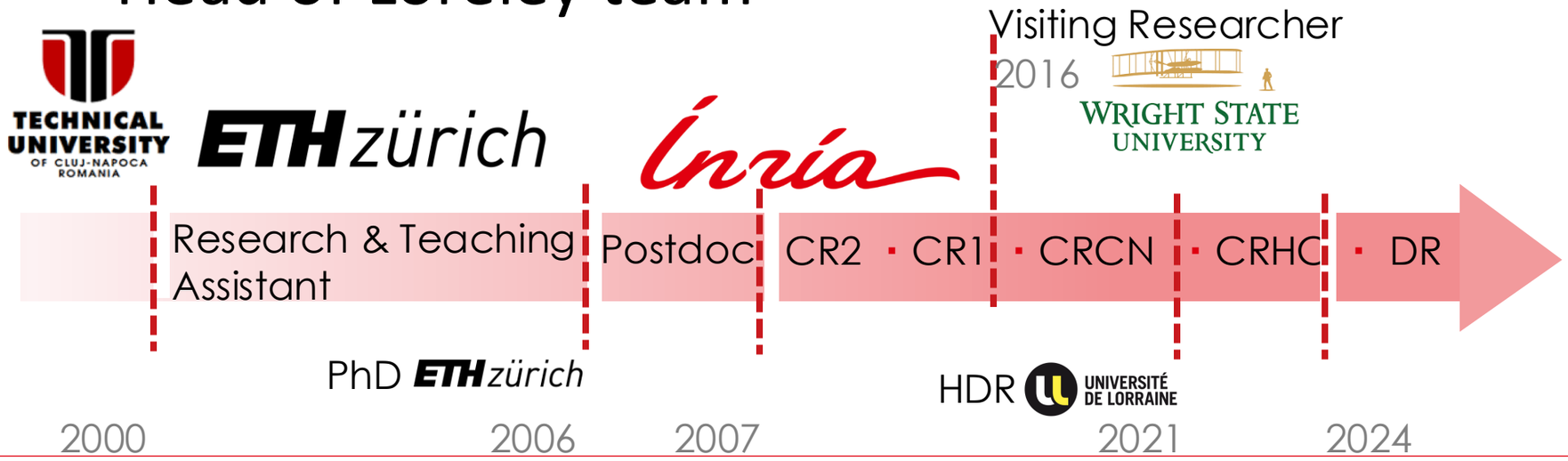
PEER-TO-PEER ARCHITECTURES FOR SCALABLE STORAGE AND EFFICIENT DISTRIBUTED LEARNING

Claudia-Lavinia Ignat
Director of Research, Inria, France
Head of Loreley team



Research domain: Distributed collaborative systems



Goal:

- Algorithms/solutions to allow **users** to collaborate on data in a reliable, secure and trustworthy manner

Scientific approach:

- Replication algorithms design, implementation and evaluation (simulations, **real collaboration traces**, **user studies**)
- Collaborative practices evaluation (real collaboration traces, **experimental game theory**, **empirical studies “in the wild”**)

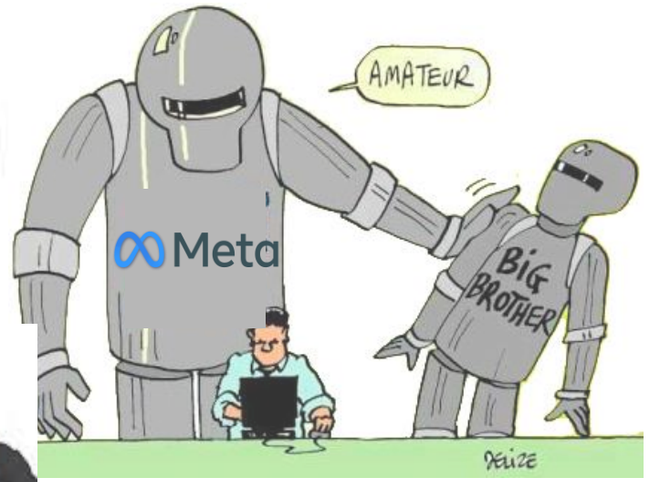
Collaborative system limitations

SCALABILITY



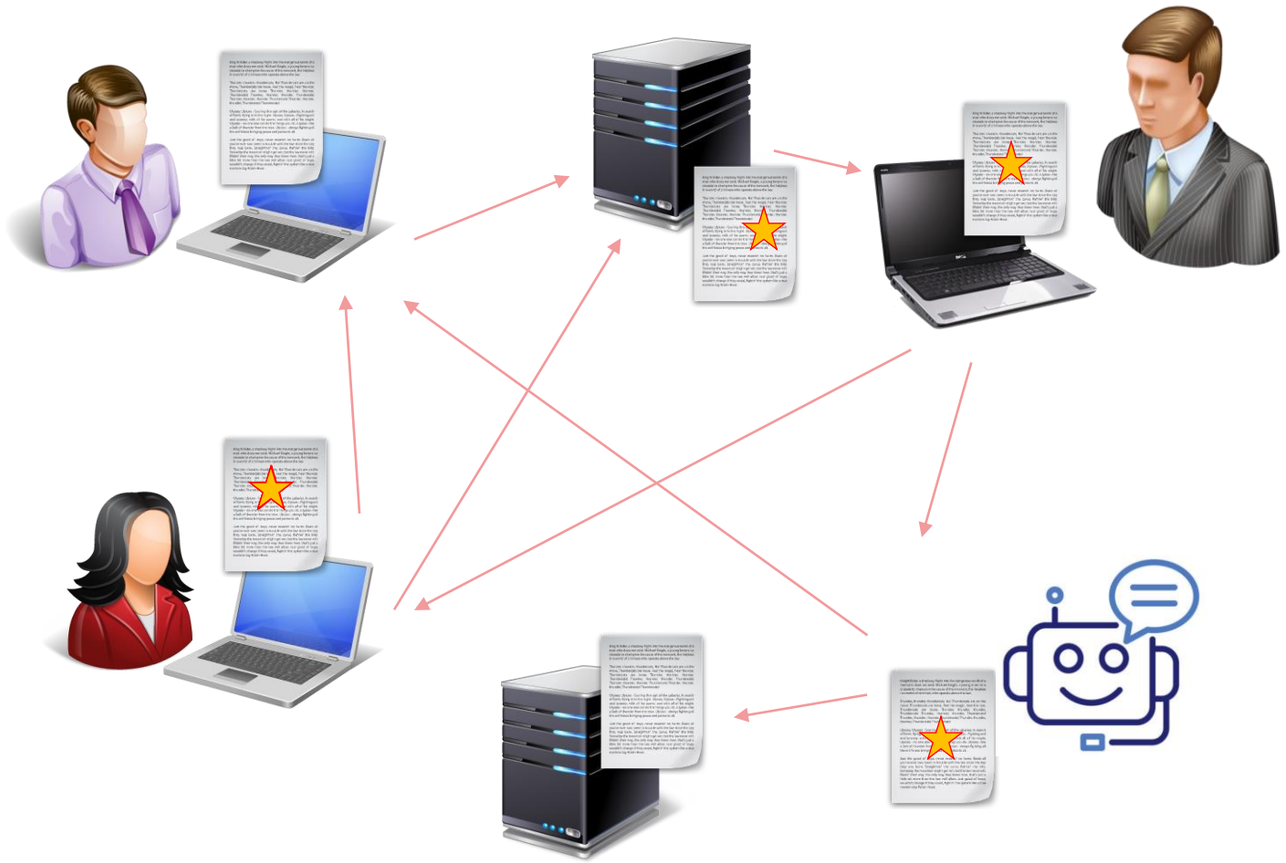


PRIVACY & OWNERSHIP

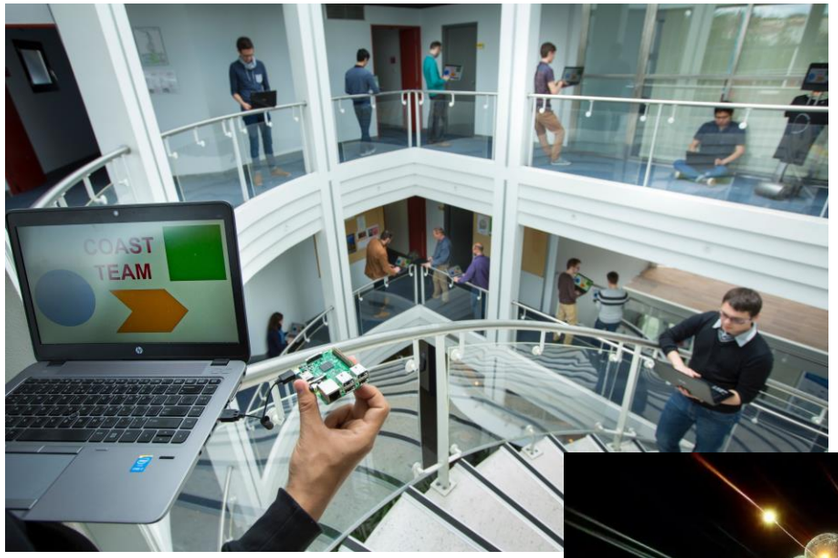


Scientific challenges

- Distributed/peer-to-peer collaborative systems without central authority
- Large-scale collaboration (from users to community of users)



Loreley application domains



Collaborative editing

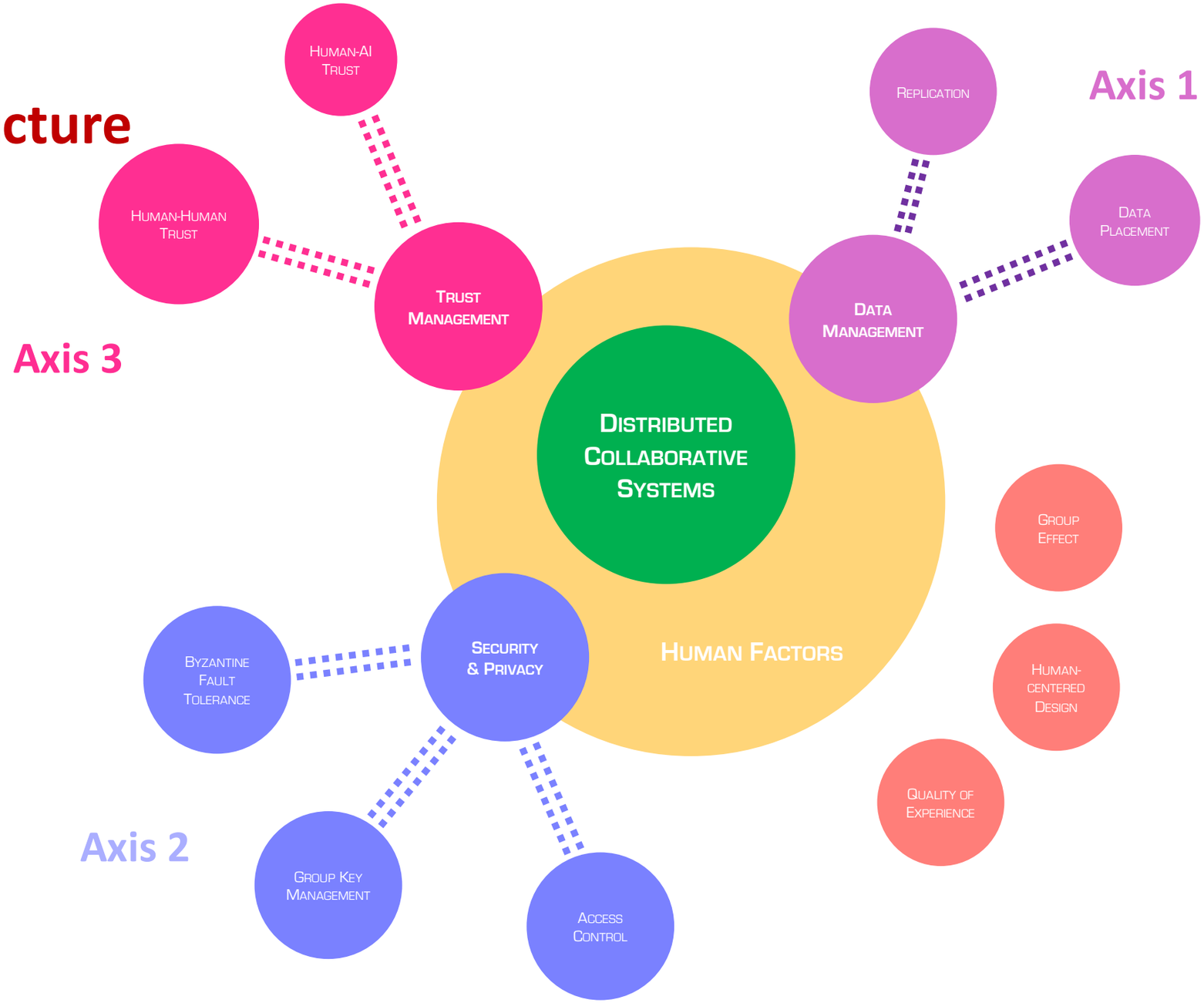


Peer-to-peer storage



Crisis management

Loreley structure



Impact

Web 3.0 paradigm

- secure sharing mutable data
- how users adopt peer-to-peer collaboration, how trust is built between them

- *Coordination of défi Alvearium with Hivenet*
- Coordination of PILOT (Practices and Infrastructure for long term collaboration) of PEPR eNSEMBLE
- Coordination of IPCEI DXP (Data Exchange Platform) with Amadeus
- Participation to TrustInCloudS of PEPR Cloud
- *Participation to défi CUPSELI with Hivenet (follow-up of Alvearium)*

Peer-to-peer (P2P) cloud storage

Alvearium - Collaboration with Hivenet



Goal:

- Provide data storage through a P2P network rather than a centralized set of data centers
- Users bring computing resources to the network

Advantages:

- Sustainable
- Secure & private
- Data sovereignty
- Affordable

Research questions

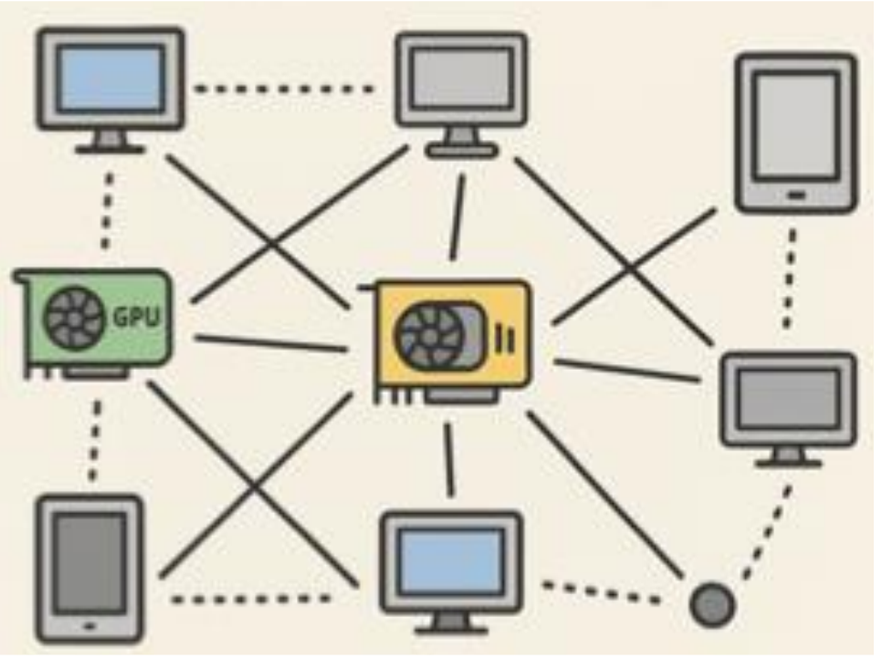
- How to place data in the P2P storage in order to ensure high availability?
- How to manage data that can be modified by different peers?
- How to ensure a P2P storage in a trustless environment? How to deal with Byzantine nodes? How to handle Sybil attacks?
- How to define a peer-to-peer security mechanism (E2E encryption, group key management)

Decentralized and Sustainable Computation

Cupseli – Collaboration with Hivenet

Motivation

- Exploit the unused computing power available on user devices (GPUs, CPUs, ..)
- Target workloads like AI inference and fine-tuning, which can be distributed efficiently even over less powerful networks



Challenges

- Hardware heterogeneity and limited GPU memory
- Low network bandwidth
- Data confidentiality

Research topics

Frugality: efficient training on limited and volatile resources

- Memory and communication frugality
- Dynamic resource management
- Compression and sparsification

Security and confidentiality in distributed AI

- Software-level privacy
- Hardware-level privacy
- Adversarial robustness

Large-scale computing on intermittent resources

- Dynamic scheduling and fault tolerance
- Resilient distributed training and inference
- Predictive allocation and frugal algorithms for volatile devices
- Containerization and orchestration for transient clusters

Thank you! Questions ?