Masters and PhD GRADUATE SCHOOL DS4H Digital Systems for Humans
Let's anticipate the digital transformation

UNIVERSITÉ CÔTE D'AZUR
ACADÉMIE D'EXCELLENCE
RÉSEAUX, INFORMATION, SOCIÉTÉ NUMÉRIQUE

Académie Réseaux, Information et Société Numérique
Ecole Universitaire de Recherche DS4H

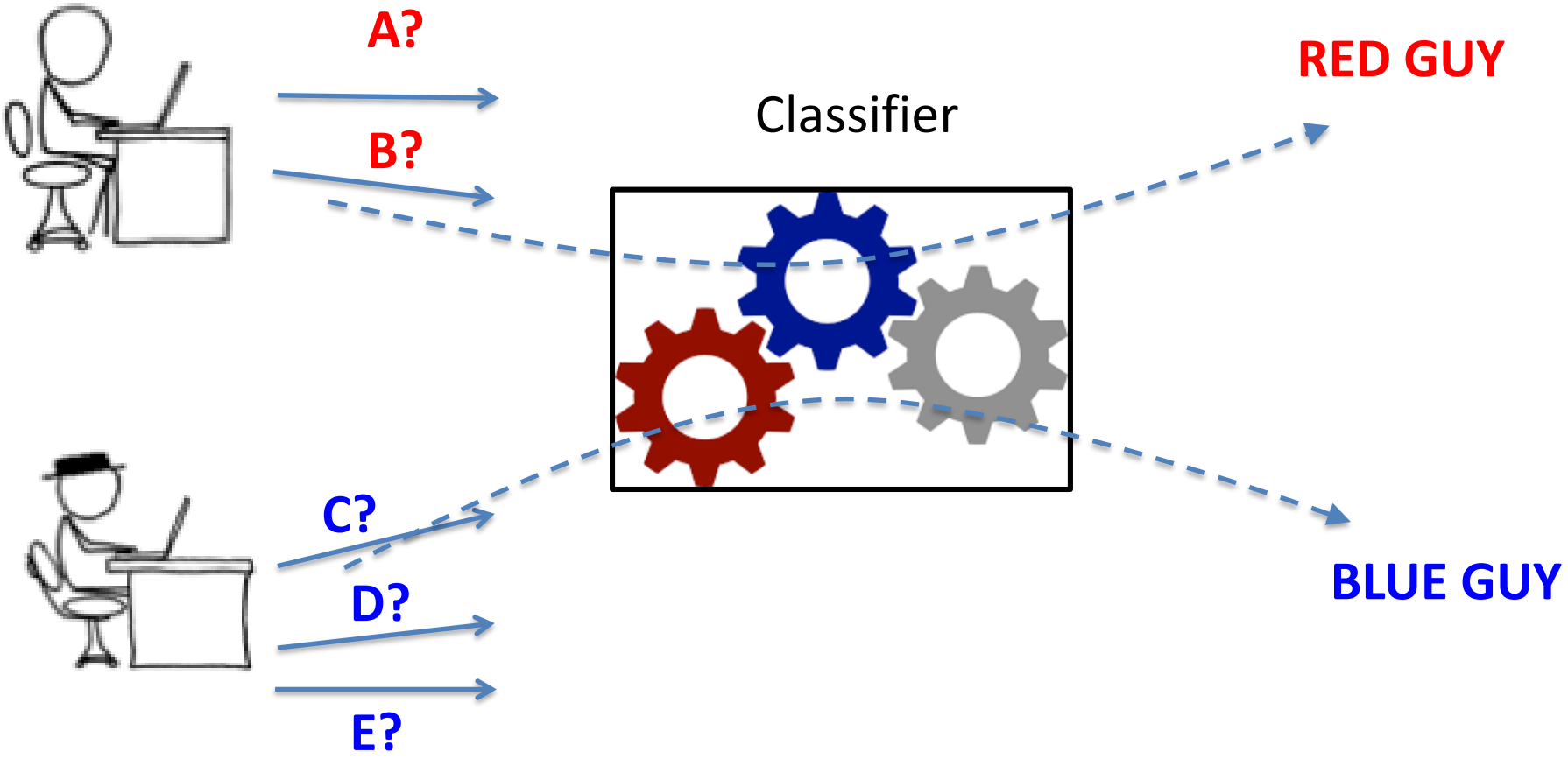
MYDATA: Muddle Your DATA

Charles Bouveyron, LJAD/Inria
Michela Chessa, GREDEG
Giovanni Neglia, Inria

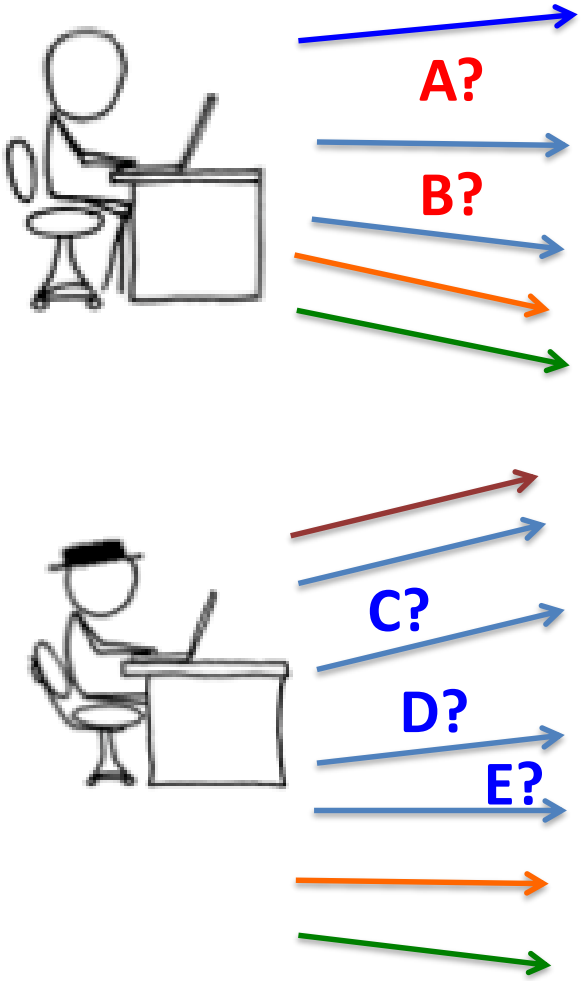
Motivation


A?
B?
C?
D?
E?
Google
YAHOO!
Netscape
msn
altavista
AOL
excite
Ask Jeeves
overture
Inktomi

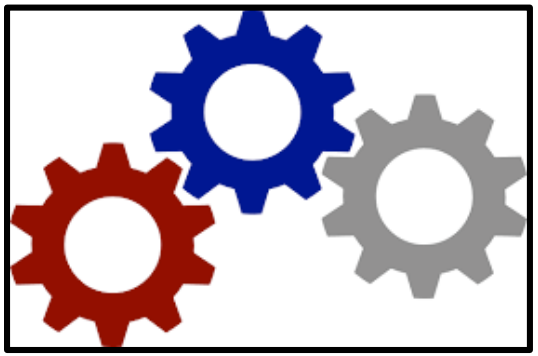
Motivation



First Goal:
investigate privacy by obfuscation



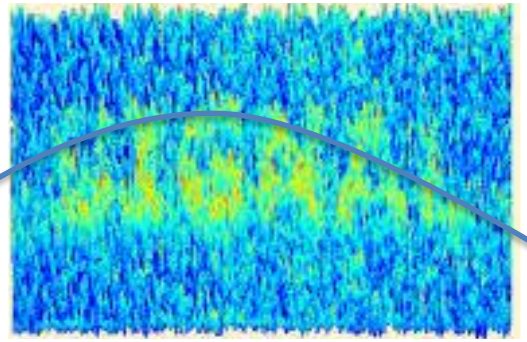
Classifier



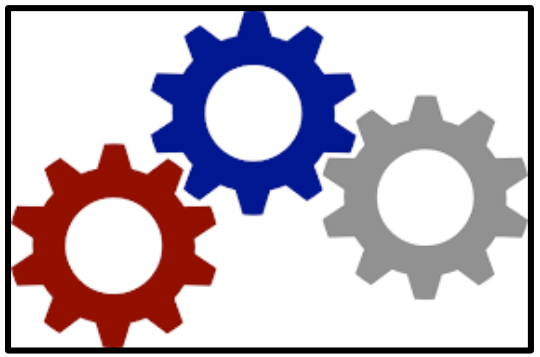
?

SYNTHETIC NOISE

Second Goal: economic interaction

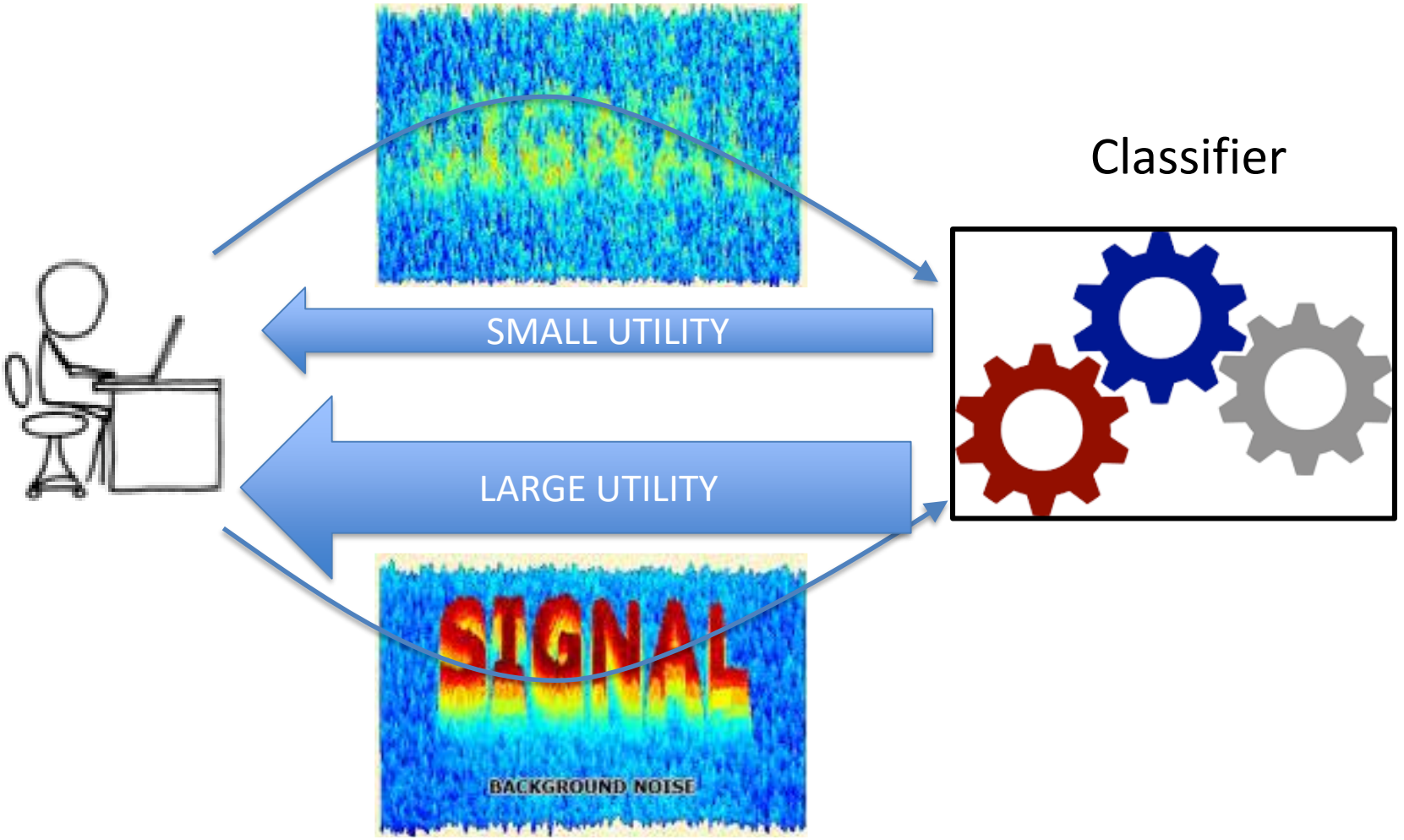


Classifier



SMALL UTILITY

Second Goal: economic interaction



Resources

Permanent researchers

- Charles Bouveyron, LJAD/Inria
- Michela Chessa, GREDEG
- Giovanni Neglia, Inria

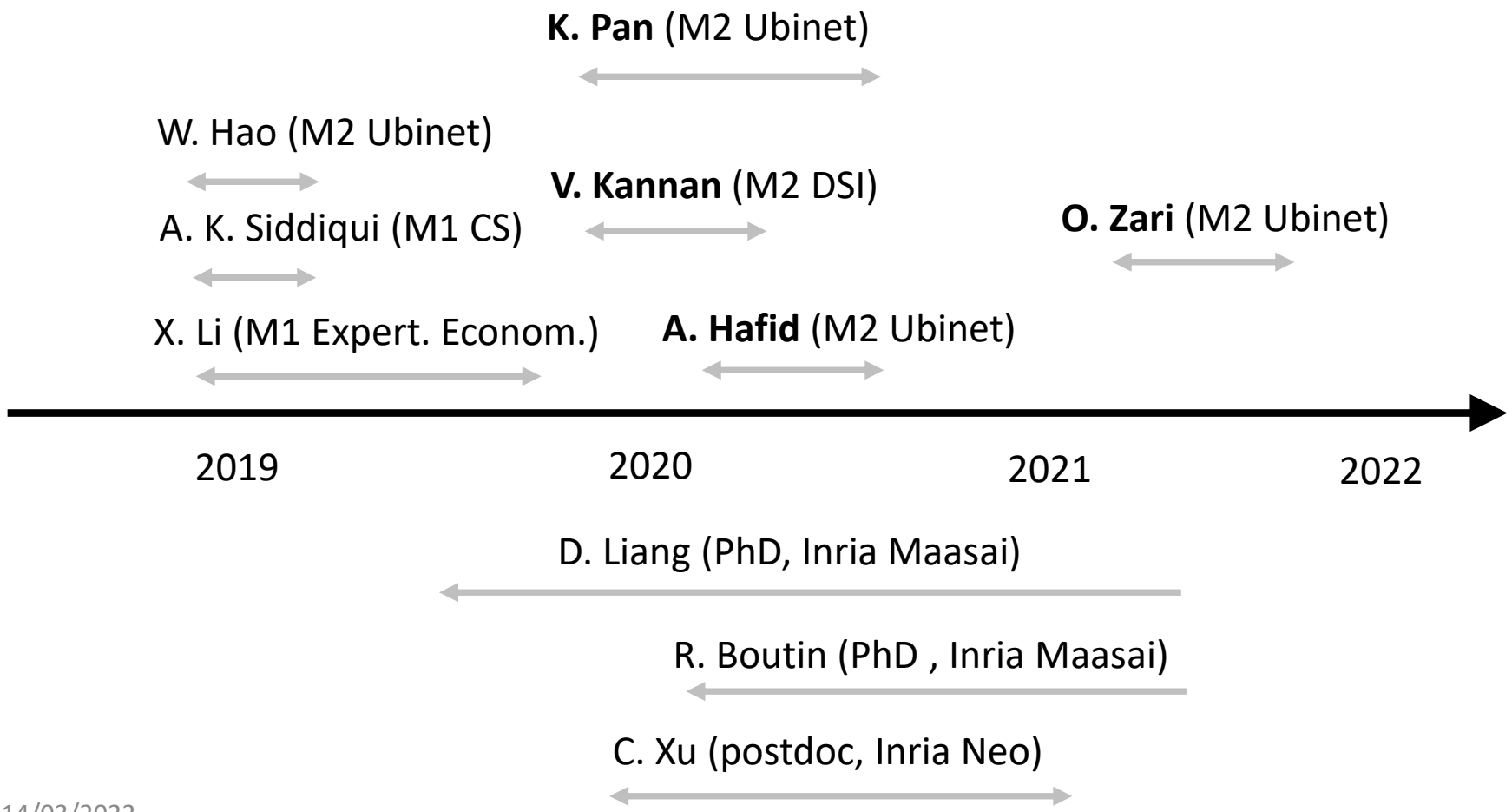
Funding (Sept. 2018 – Dec. 2020, ext. Dec. 2021)

- 27720 euros from Academy 1
  - 4 six-month internships
  - 2 PCs
  - Dissemination (this budget was reoriented to material)

Challenges & Difficulties

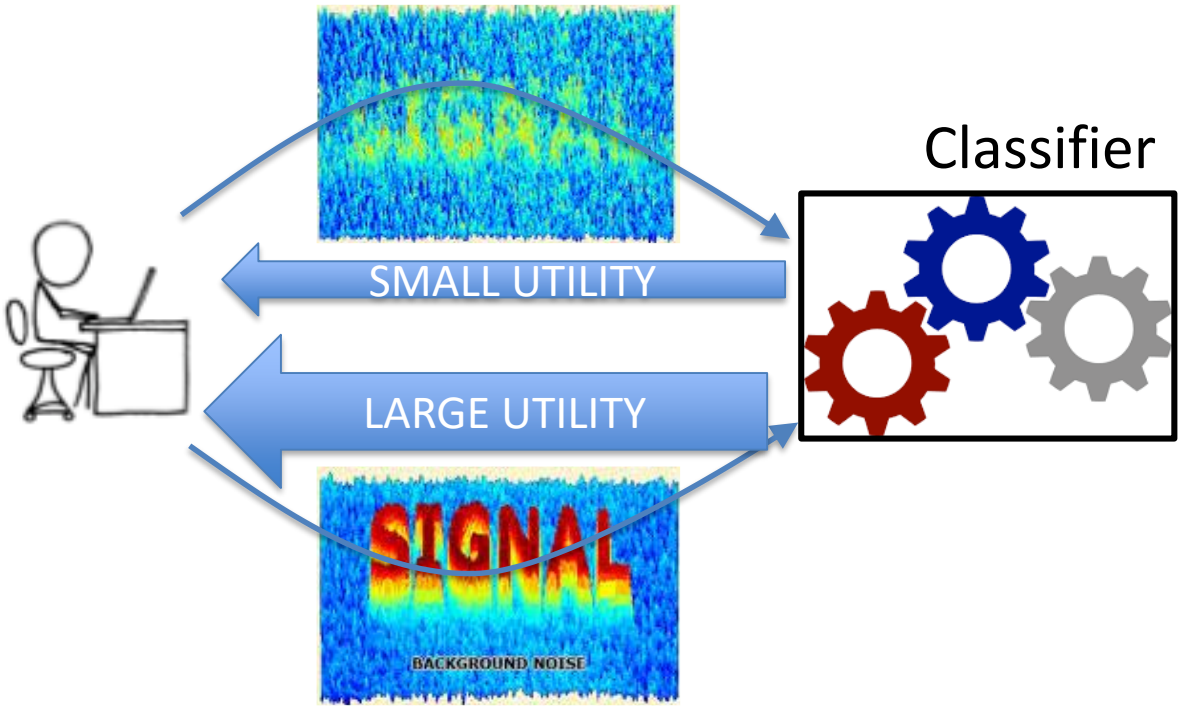
- New topic for 2 out of 3
- New collaborations
  - to bootstrap through master students' internships
- 3IA (since 2018)
- COVID-19

Other people involved



Scientific contribution 1 (M.C.)

- Study users' strategic behaviour to protect their personal data in a centralized and decentralized setting

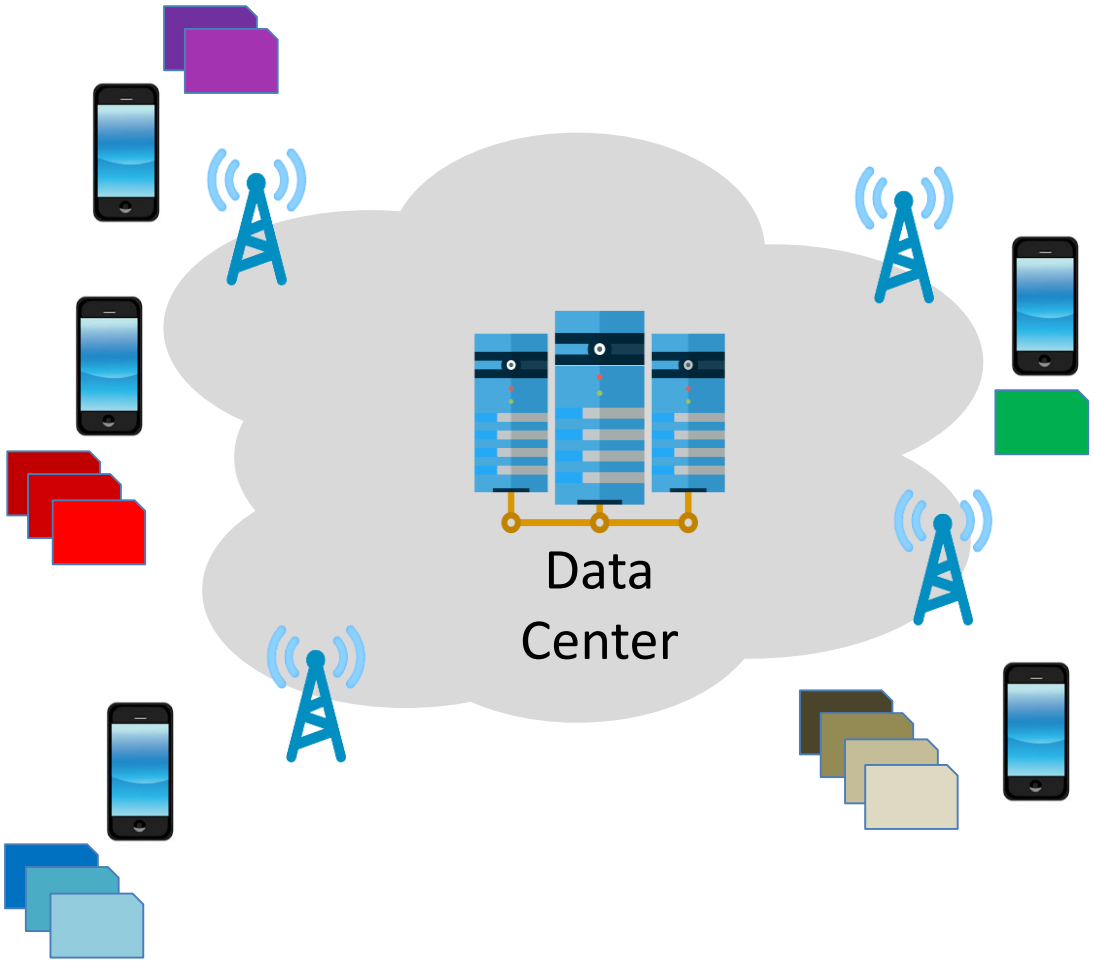


Scientific contribution 1 (M.C.)

- Study users' strategic behaviour to protect their personal data in a centralized and decentralized setting
- Current result: new revelation mechanisms in the centralized setting, which are both incentive compatible (user has no interest to cheat) and budget balanced in expectation (no money exchange among central authority and users).

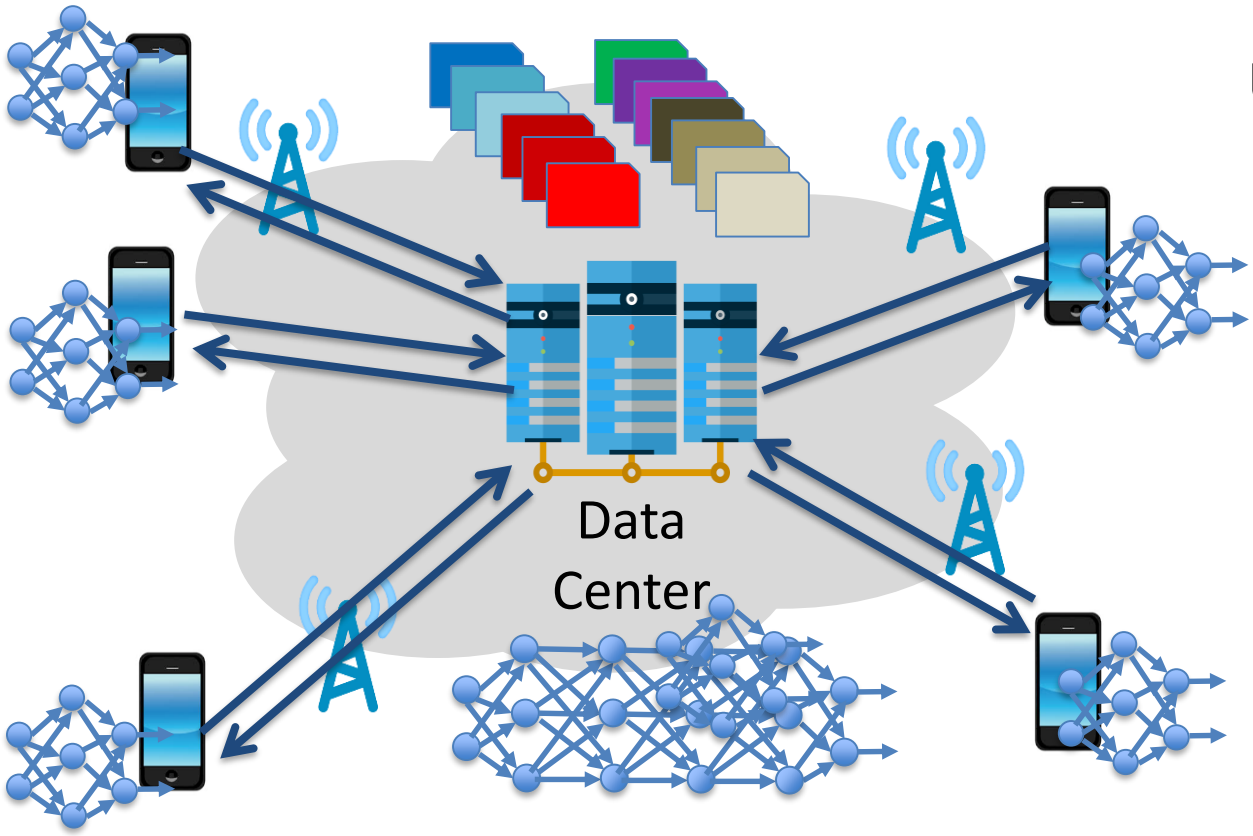
- Michela Chessa, "A Shapley-based Groves mechanism: When the mechanism designer plays the wise man," Operations Research Letters, Vol. 47(6), 2019

Scientific contribution 2 (G.N.)



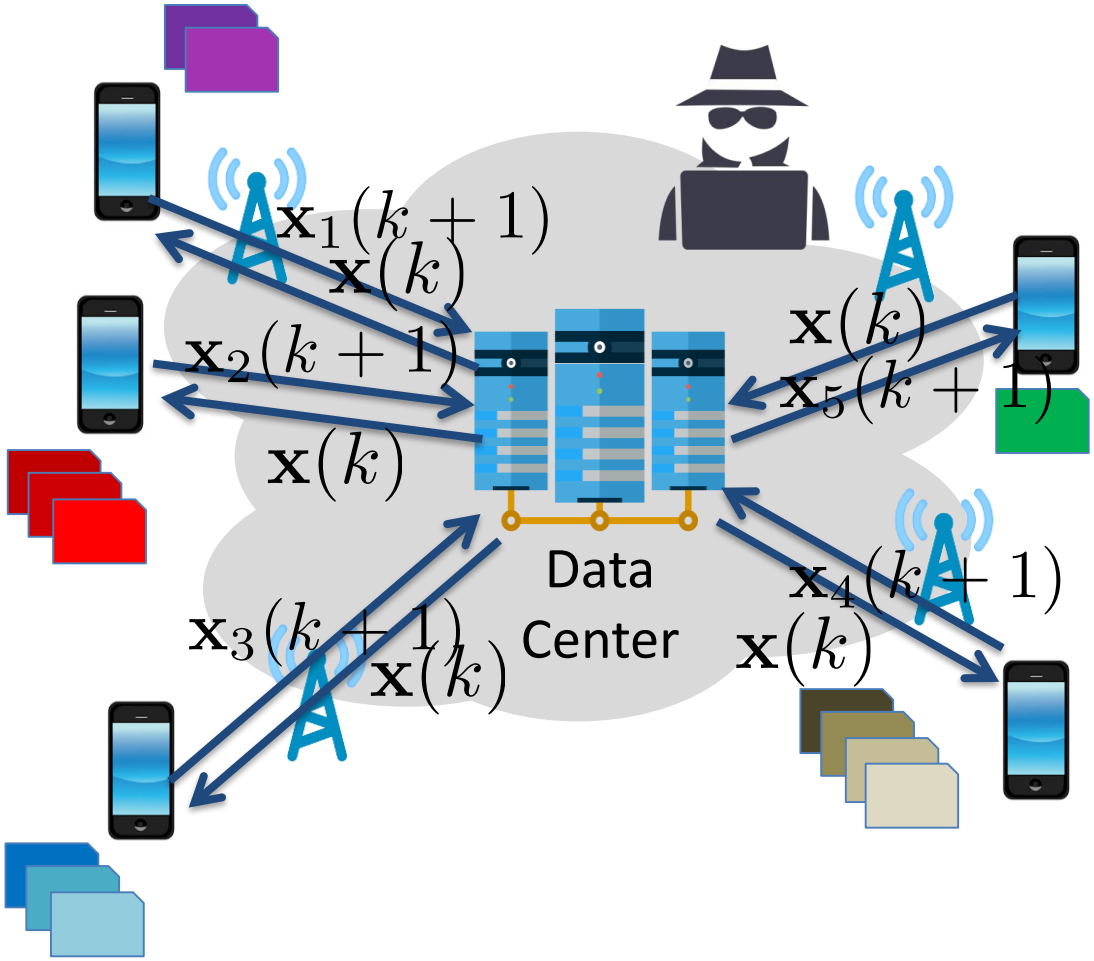
Usual ML Training

Scientific contribution 2 (G.N.)



Usual ML Training

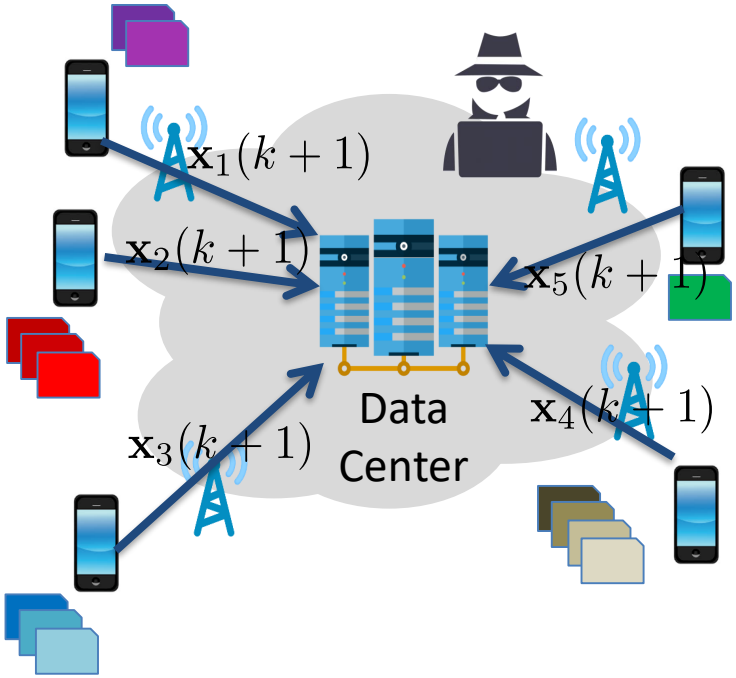
Scientific contribution 2 (G.N.)



Federated learning

- Train ML models keeping data local (transfer costs and privacy concerns)

Scientific contribution 2 (G.N.)

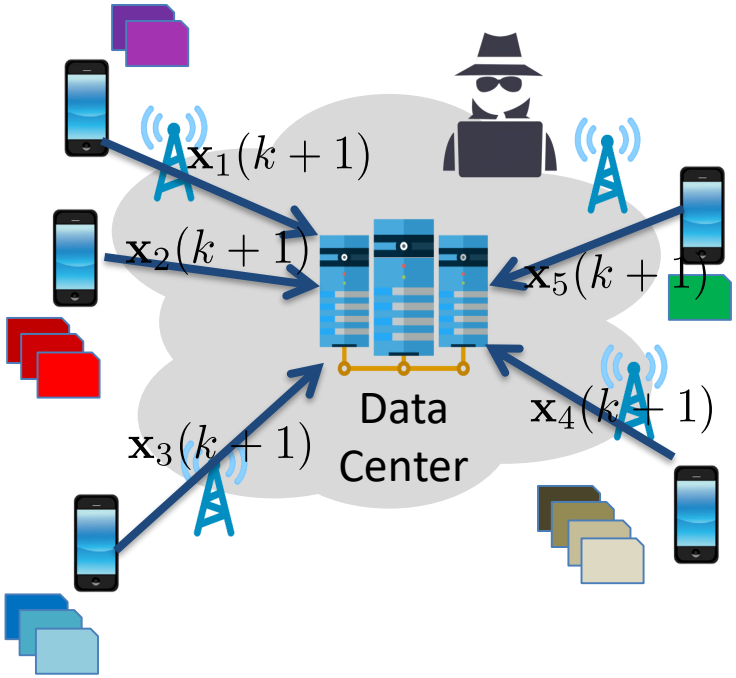


Membership Inference Attack

- Understand if a given record was in the training dataset
- Proposed more efficient attack
  - on CIFAR100:
    - 4 p.p. more accurate
    - 3 orders of magnitude less memory space
    - 5 orders of magnitude less calculations

Oualid Zari, Chuan Xu, and Giovanni Neglia. “Efficient passive membership inference attack in federated learning”. In: NeurIPS workshop on Privacy in Machine Learning (PriML). Virtual., Dec. 2021.

Scientific contribution 3 (G.N.)



New Local Model Reconstruction Attack

- A single record may not reveal much about the user, its local model may reveal much more!
- Proposal of this new framework and of practical attacks

Chuan Xu and Giovanni Neglia. “What else is leaked when eavesdropping Federated Learning?” In: CCS workshop Privacy Preserving Machine Learning (PPML). Virtual, Contributed talk. Nov. 2021.

Scientific contribution 4 (C.B.)

- Development of statistical models combined with deep architectures for visualization & clustering of networked data.
- Idea: adding noise in the (deep) latent space may be a convenient way to introduce privacy for network clustering techniques.

- Publications to be submitted in the next couple of months.

Pedagogical contributions

- a new M2 course on “Stockage et protection des donnees” for the master Expertise Economique, titled, taught by M. Chessa and T. Jobert since 2020–2021

- a new M2 course on “Federated Learning and Privacy” for UCA master Data Science and Artificial Intelligence, taught by G. Neglia and C. Xu, to start in 2022–2023

Transfer toward Industry

- Payback network [Nov. 2019 – Dec. 2020], two full consulting days on differential privacy provided by G. Neglia

- Caelin Kaplan's industrial PhD thesis [Sept. 2021 – ongoing] with SAP Labs. Co-supervised by G. Neglia, with C. Xu (UCA) and A. Santana De Oliveira (SAP Labs)

- Ilias Driouich's industrial PhD thesis [Jan. 2022 – ongoing], Amadeus. Co-supervised by G. Neglia, with C. Xu (UCA), F. Giroire (CNRS), and E. Thomas (Amadeus)