MUSKETEER: Machine Learning to Augment Shared Knowledge in Federated Privacy-Preserving Scenarios

Dr. Luis Muñoz-González – Imperial College London (l.munoz@imperial.ac.uk, @luismunozgz)

IDSA Winterdays, Paris, 4th December 2019

“This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824988”.
MUSKETEER: Project overview

- **H2020 project**
  - Research and Innovation Action
  - ICT-13-2018-2019: Supporting the Emergence of Data Markets and Data Economy


- **Budget**: 4,380,335€

- **Consortium**: 11 partners including:
  - 3 Academic institutions
  - 8 Industrial partners

- **Website**: [http://musketeer.eu/](http://musketeer.eu/)
Consortium

- **11 participants** from 7 European countries:
  - 3 academic institutions
  - 2 SMEs
  - 2 large technological companies
  - 1 industrial association (IDSA)
  - 2 companies in the car manufacturing industry
  - 1 hospital
Motivation and Objectives

**Motivation:**
- Need of new ways to **preserve privacy** while still allowing data or model sharing among multiple data owners.
- Lack of trusted and secure environments for **data sharing inhibits data economy**.
- Legality, privacy, trustworthiness, data value and confidentiality hamper the free flow of data.

**MUSKETEER's goals:**
- Create machine learning models over a variety of **privacy-preserving scenarios**.
- Ensure **security** and **robustness** against external and internal threats.
- **Enhance data economy** by boosting sharing across domains.
- Provide a **standardized** and **extendable architecture**.
- Demonstrate and validate in two different industrial scenarios: Smart manufacturing and healthcare.
Privacy-Preserving Machine Learning

- **Federated Machine Learning** enables the creation of shared machine learning models without sharing datasets.
- In MUSKETEER we aim to provide different **Privacy Operation Modes** according to users’ privacy and security requirements:
  - Data cannot leave users’ facilities.
  - Data can be stored in a trusted external cloud server.
  - Encryption of datasets (ML models learn from encrypted data).
Security of Machine Learning

- Machine Learning algorithms are **vulnerable** and can be the objective of attackers.
- Attackers can compromise data collection or exploit the system weaknesses at test time.
  - Poisoning Attacks
  - Evasion Attacks
  - Backdoors
- In MUSKETEER we aim to investigate and develop more **secure federated machine learning** algorithms:
  - To mitigate external threats.
  - Detect malicious and faulty users in the platform.
  - Detect users providing poor quality data.
Data Value Estimation

Musketeer in the IDS Ecosystem

- Align MUSKETEER platform with IDSA Reference Architecture Model:
  - Ensuring Privacy Operation Modes compatibility with IDSA framework.
  - Investigate the interoperability of the MUSKETEER architecture with IDSA.

- Ensure MUSKETEER platform interoperability with existing data platforms using IDSA framework.

- Boost data sharing and data market places within IDSA framework.
Questions?

- MUSKETEER website: http://musketeer.eu/
- Twitter: @H2020Musketeer
- LinkedIn: H2020_MUSKETEER
- Dr. Luis Muñoz-González:
  - Email: l.munoz@imperial.ac.uk
  - Website: https://www.doc.ic.ac.uk/~lmunozgo
  - Twitter: @luismunozgz