

# CASD-TeraLab

## Secure Access to Big Data

Alexandre Marty

[ [alexandre.marty@casd.eu](mailto:alexandre.marty@casd.eu) ]

# Outline

2

- ▶ CASD-TeraLab
- ▶ Use Cases
- ▶ Live Demo

# The Secure Data Access Centre

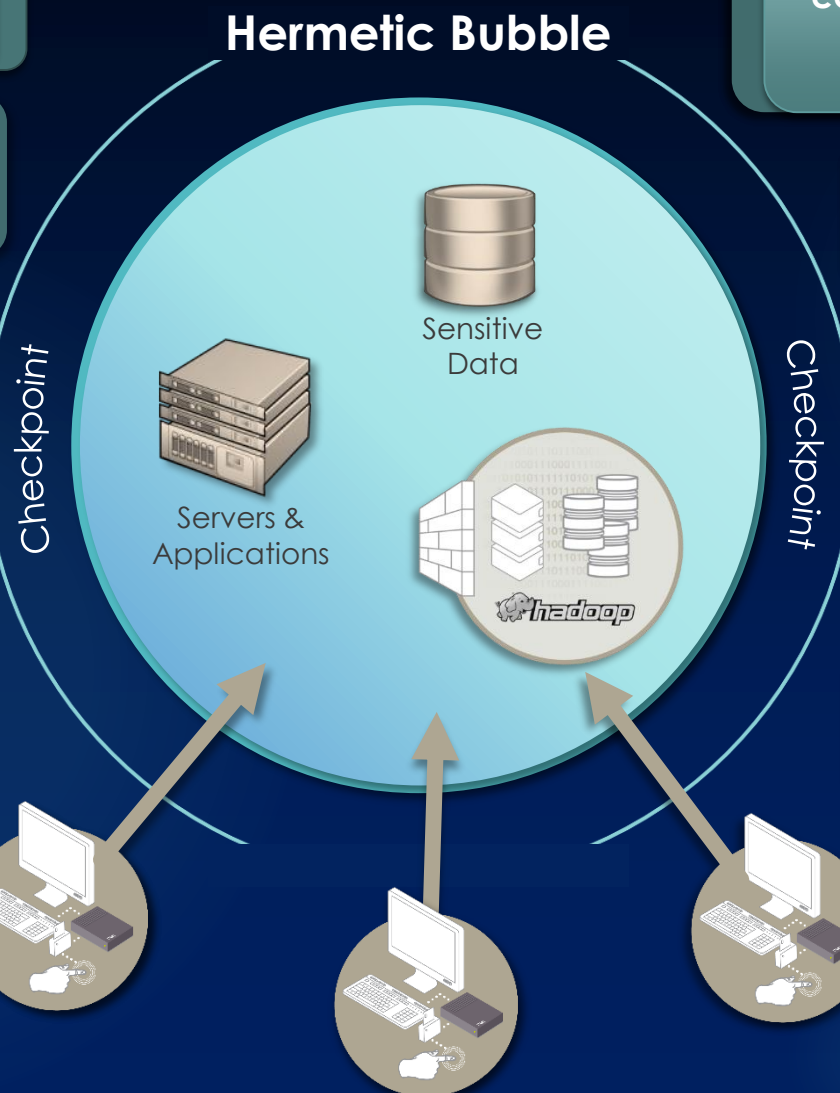
3

A group of **tightly-sealed secured servers**

User applications and processing are executed strictly within the Bubble.

Data Insertions/extractions are **controlled**. Users do not have Internet access from their workspace.

Sensitive data is hosted only within the Bubble.



Insertions

Extractions

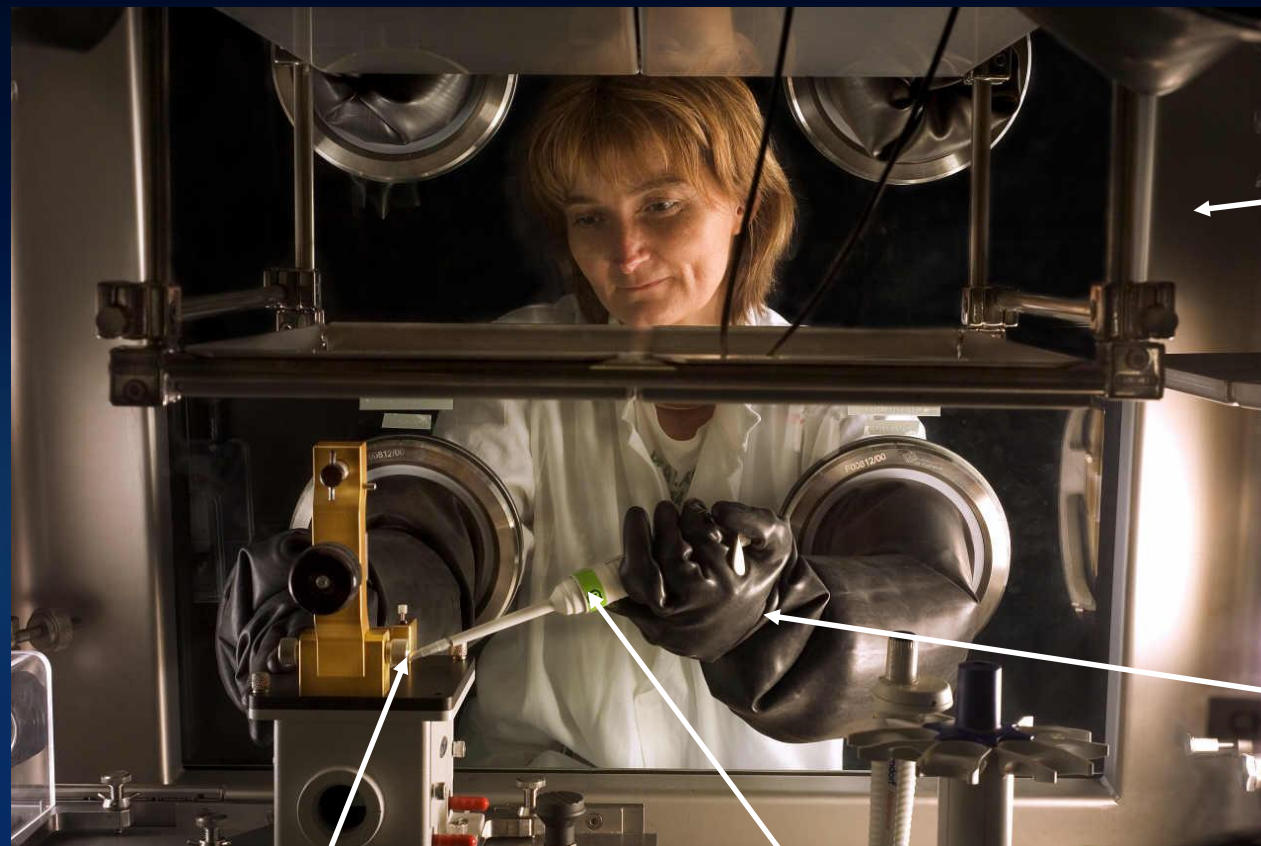
SD-Boxes are the only means of access to the Bubble.

Access occurs via the Internet by **encrypted channels**

Hadoop cluster is available for handling **Big Data**.

# The SD-Box

4



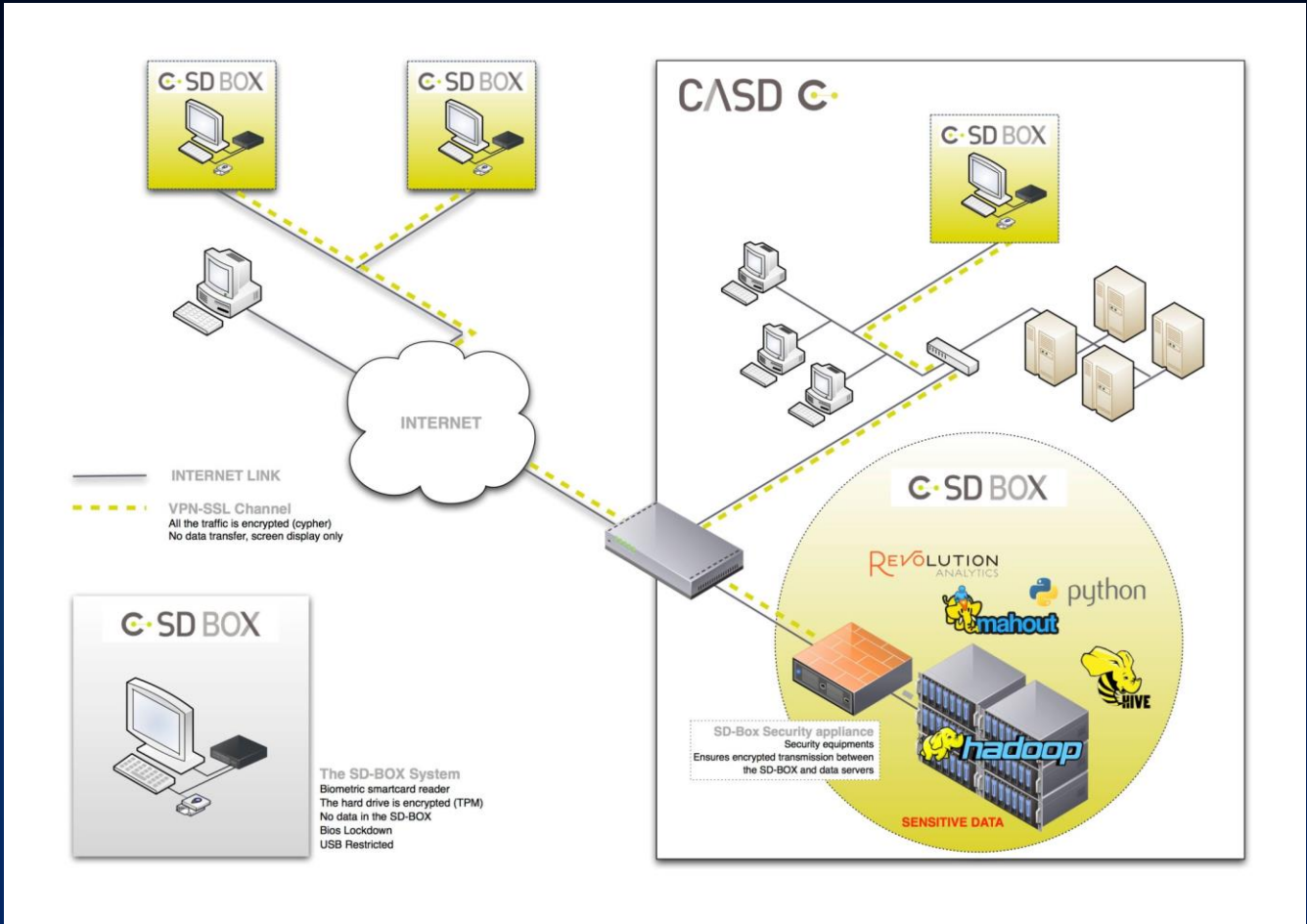
← The bubble

← The SD-Box

← The data

← The tools

# CASD Architecture



- ▶ Publicly funded Big Data & Data Science platform
- ▶ Open to:
  - ▶ R&D and teaching projects, proof of concepts
  - ▶ Public and private sectors
- ▶ Everything for big data and data science:
  - ▶ Powerful and scalable infrastructure
  - ▶ Hadoop-based with all Hadoop tools, Spark
  - ▶ Extensive tools for scientists (Python, R, SAS, machine learning...)
- ▶ Turnkey solution with full support and maintenance

# Use Cases 1

7

- ▶ Electricity transmission network data
  - ▶ Impressive variety of data sources
  - ▶ Development of innovative apps
- ▶ Health data
  - ▶ Requires high confidentiality
  - ▶ About 250 TB generated each year
- ▶ Involvement in European projects

# Use Cases 2

- ▶ Scanner data for the Consumer Price Index
  - ▶ Daily sales data from 4 French major distribution companies
  - ▶ Very detailed data: products, stores...
  - ▶ 5.7 billion rows, 1 TB
- ▶ Tabular data anonymization
  - ▶ With INSEE methodology team
  - ▶ Anonymize large data sets, too large for Tau-Argus



# Live Demo

9

# For More Information

10

- ▶ [www.teralab-datascience.fr](http://www.teralab-datascience.fr)
- ▶ [casd.eu](http://casd.eu)
- ▶ [alexandre.marty@casd.eu](mailto:alexandre.marty@casd.eu)