# Functional Programming with Scala

Project Title: FP Project

Group Name: Best Team

Bastian Albaut

IG5_Polytech

November 19, 2024

# About the Project

- **MongoDB Integration:**
  - I was responsible for setting up the MongoDB database.
  - Established a connection between Scala and MongoDB.
  - Created Scala objects to manage MongoDB queries and updates.
  - Loaded and processed JSON data, then wrote the results into MongoDB.

- **Spark Integration:**
  - Used Apache Spark for distributed data processing.
  - Implemented a Scala-based Spark job to process and transform large datasets.
  - Loaded processed data into MongoDB for persistent storage.

# Challenges and Learning

- **Challenges Faced and Their Resolutions:**
  - Configuring the connection between Spark and MongoDB.
    - Resolved by setting up the correct MongoDB Spark Connector and ensuring compatibility with the Scala version.
  - Processing large JSON files efficiently.
    - Used Spark's distributed computation capabilities to manage data in chunks.

- **Learning Gained:**
  - Gained hands-on experience in integrating MongoDB with Spark and Scala.
  - Learned about distributed data processing and handling database connections in Scala.

- **Future Improvements:**
  - Optimize Spark jobs for better performance.
  - Implement data validation pipelines before storing data in MongoDB.
  - Explore using more advanced MongoDB query operations for analytics.