**A metadata based Spark data ingestion framework with emphasis on data DQ profiling**

Prepared by :

Abhishek Singh

[abhishekvsingh@outlook.com](mailto:abhishekvsingh@outlook.com)

**Contents :**

1. Project Architecture
2. Simple Twitter data extraction
3. Metadata based framework
4. Scope of DQ implementation in the project
5. Code explanation