Project Design Phase
Technology Stack (Architecture & Stack)

DATE	29 October 2023
TEAM ID	NM2023TMID06909
TEAM LEAD	DINESH
TEAM MEMBERS	SYED ABDUL RAHMAN
	GUGANESH
	GOKUL
	RAJA PANDI
PROJECT NAME	DISSECTING THE DIGITAL LANDSCAPE: A COMPREHENSIVE ANALYSIS OF SOCIAL MEDIA

Technical Architecture:
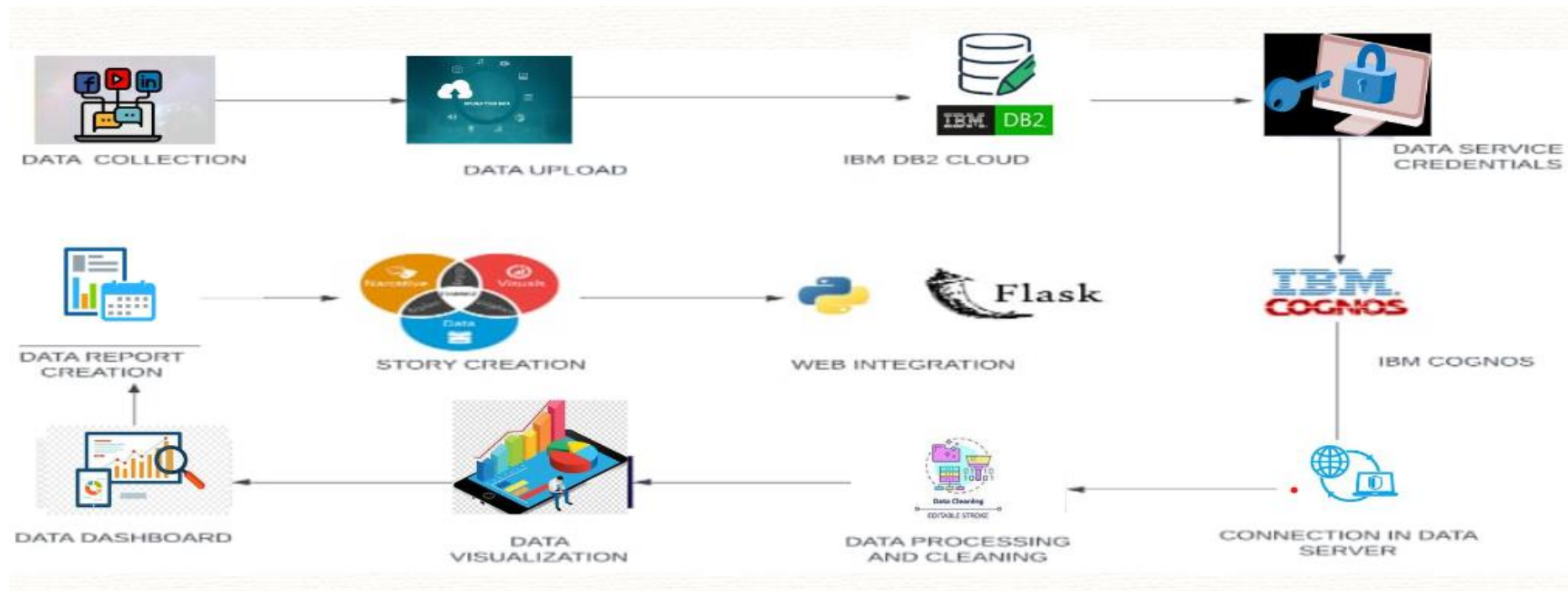
Technology Stack:

Table-1: Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	Web application for Exploratory Data Analysis (EDA) on social media	HTML, CSS, JavaScript, Flask
2.	Data Collection & Extraction from Database	Collect the dataset required for analysis.	IBM DB2, IBM Cognos
3.	Data Preparation	Data preparation involves cleaning, transforming, and structuring collected data for analysis.	Python
4.	Data Visualizations	Data visualizations to help users understand and explore the data.	IBM Cognos

5.	Database	Data storage for managing and storing structured data.	IBM Cloud, IBM DB2
6.	Cloud Database	Cloud-based database for traditional databases, offering scalability and accessibility.	IBM DB2, IBM Cloud
7.	File Storage	File storage is used for archiving and retrieving data files that may not fit into traditional databases.	IBM Cloud
8.	External API-1	External APIs are used to fetch additional data or services.	Twitter API
9.	External API-2	External APIs are used to fetch additional data or services	Twitter API
10.	Machine Learning Model	Machine learning models are used to gain insights from data.	Python
11.	Infrastructure (Server / Cloud)	provisioning and managing the infrastructure, including servers or cloud services	IBM Cloud, Flask

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Flask framework for building web application	Python web frameworks like Flask web framework
2.	Security Implementations	Security measures are essential to protect sensitive data and ensure data integrity.	IBM Cloud resources and IBM DB2
3.	Scalable Architecture	Essential to handle large datasets and accommodate future growth.	IBM Cloud Services, IBM Db2 on Cloud, and Flask Application
4.	Availability	Availability ensures that EDA results and tools are accessible to stakeholders for data-driven decision-making.	IBM Cloud Services, Scheduled EDA Reports and Collaboration Platforms.

S.No	Characteristics	Description	Technology
5.	Performance	Performance optimization is essential for efficient data exploration and analysis.	IBM Db2 Optimization, Data Indexing and Parallel Processing.