

DATA SCIENCE WITH AZURE CLOUD SERVICES – INTERNSHIP AT DELOITTE



VIT
UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)

VELLORE ■ CHENNAI

www.vit.ac.in

Andrew John J, Internship @ Deloitte - USI

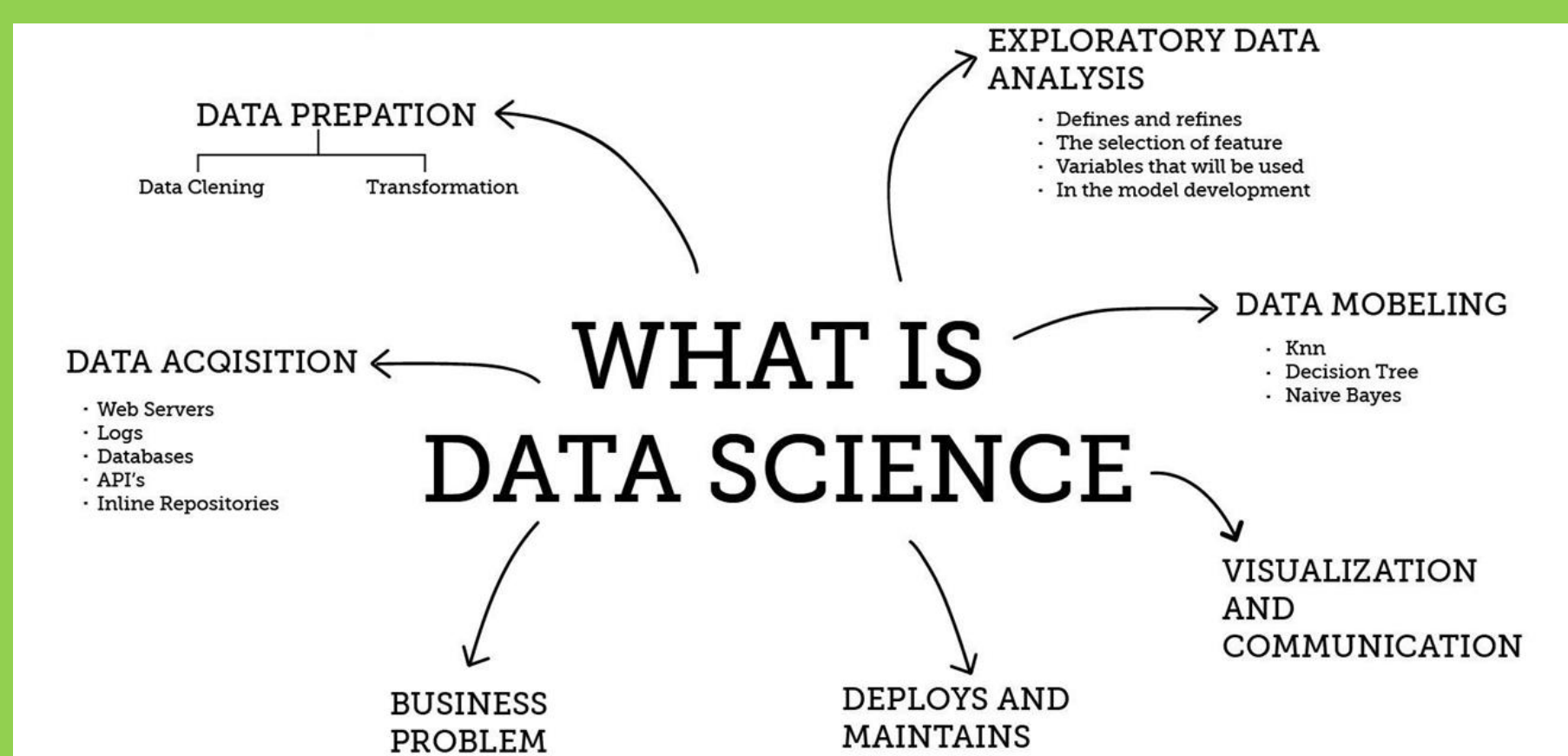
School of Electronics Engineering, VIT University, Chennai

ABOUT INTERNSHIP

- The Analyst - Internship at Deloitte is the position on which I have spent the most of my time.
- I was assigned to the following positions while learning various development tools, testing tools, techniques, and processes: Analyst – Specialist.
- As an Analyst, I developed and maintained solutions based on technologies such as Microsoft Azure Data Factory, Microsoft Azure Databricks and Dataiku.
- These technologies are more concentrated on Data Warehousing, Data Lake, and Data Lakehouse.
- My internship at Deloitte lasted from January 10th till June 3rd, precisely 21 weeks or roughly 5 months of time. My internship stipend was ₹25,000. HITEC City in Hyderabad, Telangana, was the place of my internship.

DATA SCIENCE

- Data is the new oil. Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains.



AZURE DATA FACTORY & DATABRICKS

- Big data requires a service that can orchestrate and operationalize processes to refine these enormous stores of raw data into actionable business insights. Azure Data Factory is a managed cloud service that's built for these complex hybrid extract-transform-load (ETL), extract-load-transform (ELT), and data integration projects.
- Azure Databricks is a data analytics platform optimized for the Microsoft Azure cloud services platform. Azure Databricks offers three environments for developing data intensive applications: Databricks SQL, Databricks Data Science & Engineering, and Databricks Machine Learning.



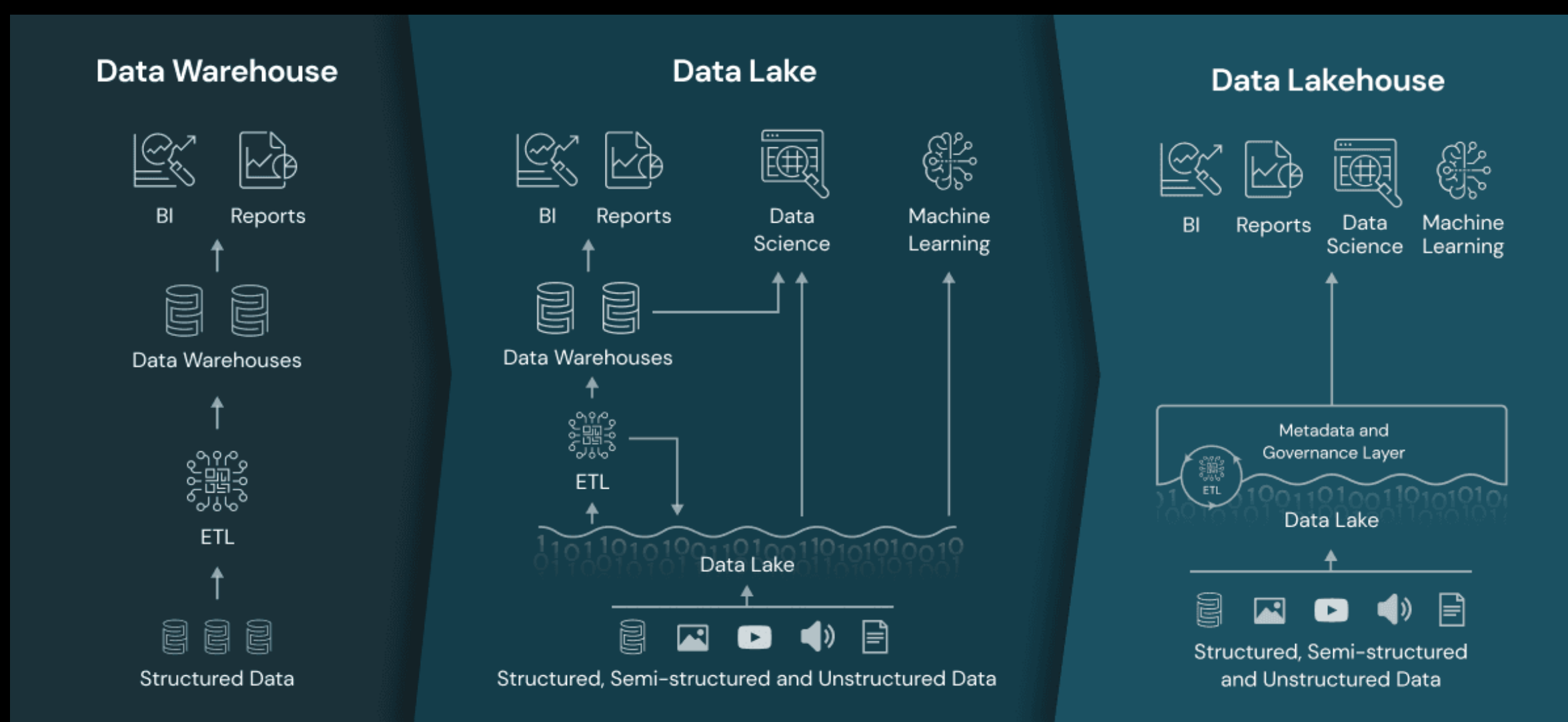
DELOITTE HYDERABAD OFFICE

In India, Deloitte has two entities: Deloitte India and Deloitte US-India (USI), which is a region within the Deloitte US organization. Deloitte India caters to clients within India, while Deloitte USI is an entity of Deloitte US that is geographically located in India and caters to clients of the US member firm.



DATA LAKEHOUSE

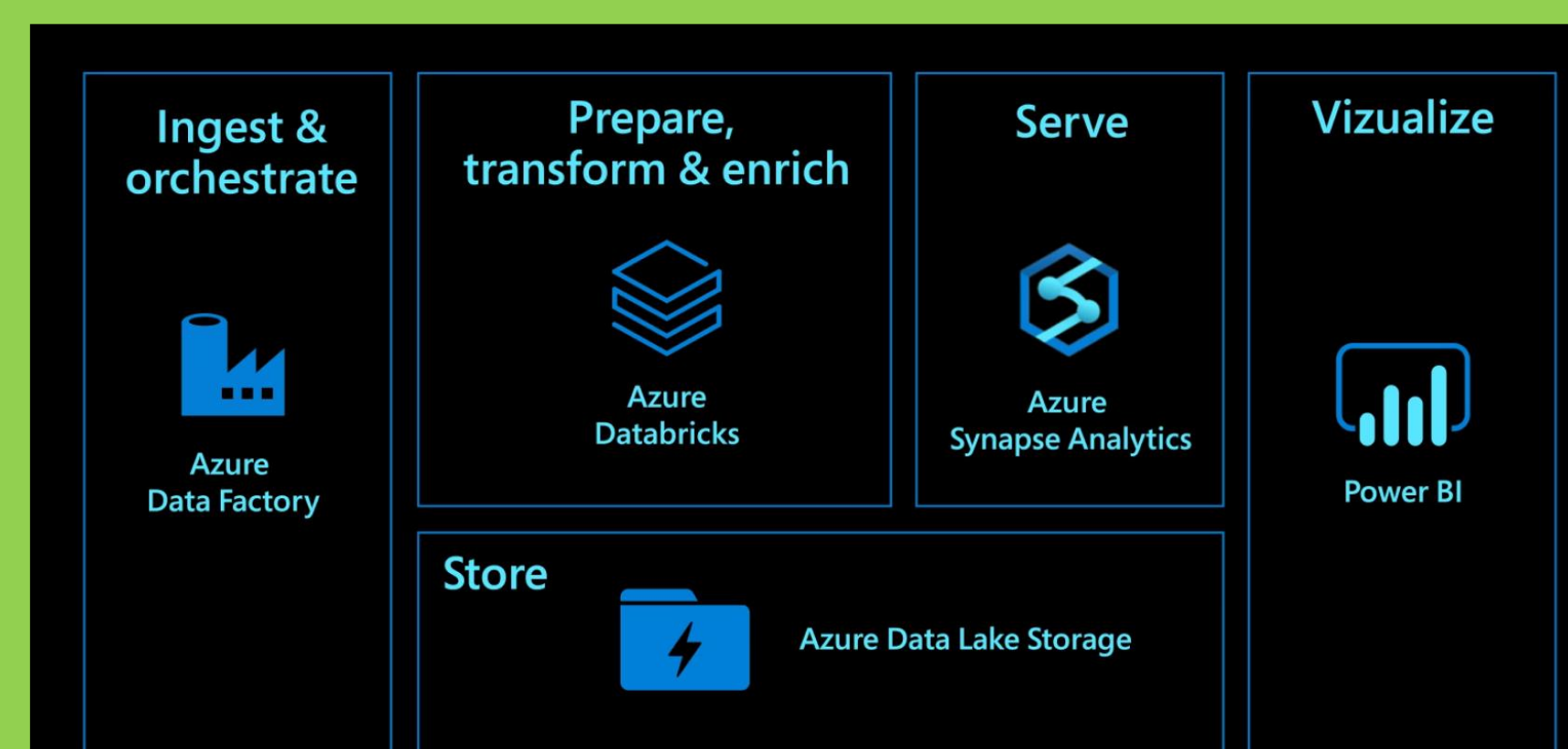
Data Lakehouse is a new, big-data storage architecture that combines the best features of both data warehouses and data lakes. A data lakehouse enables a single repository for all our data (structured, semi-structured, and unstructured) while enabling best-in-class machine learning, business intelligence, and streaming capabilities.



Deloitte.



BIG DATA ANALYTICS AND AI WITH OPTIMIZED APACHE SPARK



DATAIKU

•Dataiku Data Science Studio (DSS) provides a centralized data platform that moves businesses from scale and traditional analytics to Enterprise AI in their data journeys, empowering self-service analytics while also enabling the operationalization of machine learning models in digital environments.



CONCLUSION

- I am grateful for the internship at Deloitte – USI, which has provided me with invaluable corporate and technical expertise and experience. Technologies such as Data Science and Data Engineering are in-demand skillsets in today's world, and the need for this expertise is increasing
- Artificial Intelligence, the Internet of Things, and Deep Learning are just a few of the cutting-edge technologies covered by Data Science. Data science's influence has grown dramatically as a result of its success and technical advancements. The value of acquiring and collecting data is critical because it allows businesses to predict and affect our purchase behaviors. As a result, it wields significant power through its purchasing power.