CODE
BASICS

Data & AI

JARGONS

The Role of Data in AI

1 **Structure Data:** The type of data in which the data points follows a definite structure. For example: Data in excel is structured into rows and columns.

2 **Unstructured Data:** The type of data in which the data points do not follow a definite structure. For example: Image, audio and video files.

Data Infrastructure in a Company

1 **Data Distribution:** Providing the dataset to different consumers in a company based on their requirement and access privileges.

2 **ISO:** International Organisation of Standardization. This is internationally recognized and provides standard ways of representing data like date, weeks, currency which can be adopted by companies across the world to maintain consistency.

3 **Role Based Privileging:** Enabling access to data based on the role of an individual in the company.

Data Infrastructure in a Company

1 **POS:** Point of sales. Refers to the location where sales transaction takes place.

2 **API:** Application Programming Interface. A set of rules that allows different software applications to transfer data between each other.

3 **Web Scraping:** An automated way of extracting data directly from websites.

Data Storage and Transformation: Overview

1 **ETL:** Extract, Transform and Load. A set of steps through which data is collected, transformed and then stored in databases.

2 **ELT:** Extract, Load and Transform. A set of steps through which data is collected, stored in databases as it comes and then transformed as per the needs.

3 **Kafka:** A popular streaming platform used to process real time data.

4 **Graph Data:** A popular data structure used to represent information in the form of relationship between an entity and its attributes.

5 **MDM:** Master Data Management. A practice used in companies to ensure consistency and accuracy in data used across the organisation.

Data Distribution: Privacy, Ethics & Governance

1 **GDPR:** General Data Protection Regulation. It is a European Union regulation that aims to protect individuals' personal data by setting guidelines for its collection, processing, and storage, and granting individuals rights over their data.

2 **RBAC:** Role Based Access Control. This is same as role-based privileging where access to data or features is provided to individuals based on the role assigned to them.

3 **SDK:** Software Development Kit. A set of tools, libraries, sample code provided by a vendor to facilitate the development of application for a specific platform or operating system.