



WHITEPAPER

DATA ENGINEERING

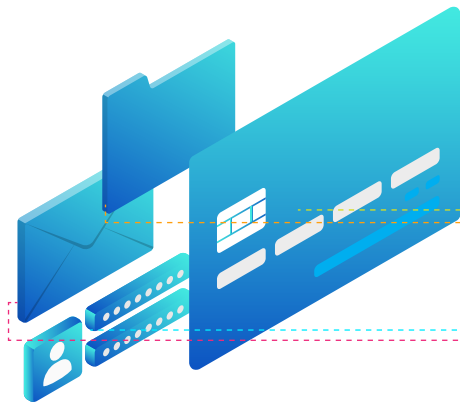
Unties the knot for effective
Artificial Intelligence

Data Engineering

A Quick Overview

Data Engineering is a process with a rational focus on the data for some practical application. It is the process of collecting data, validating its authenticity and ensuring that it is relevant to a business objective. The data should encompass a truth and present a correlation to the task

at hand. Data pipelines indicate the journey and processes that data undergoes within an organization. Data engineers are responsible for creating those pipelines.

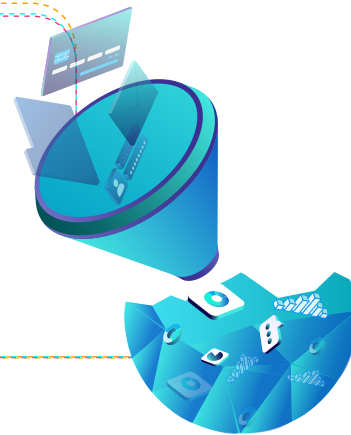


Data flows from multiple sources

When trends change, they create a pressing necessity for modifications from the point of production till the delivery of goods and services. Apart from this, data also flows from multiple sources into an organizational system in real-time, making it mandatory for organizations to respond rapidly to progressive conditions. It becomes imperative for a business to thoroughly comprehend such business patterns and take data-driven decisions to successfully meet challenges.

Valuable insights to accelerate growth

The next step of progress for organizations would be to invest in intelligent solutions. This should enable them to scale, integrate data analysis and subsequently obtain valuable insights to accelerate business decisions and growth. This modern approach to storing data would help them create efficient data



Transformed into decision making information

Most enterprises store volumes of data in sophisticated ways but fail to make optimum use of the knowledge and insights lurking underneath. A 2017 HBR report on "What's your data strategy?" states that raw data, such as customer retention rates, sales figures and supply costs, is of limited value until it has been integrated with other data and transformed into information that can guide decision making. Today, enterprises are neither mature nor ready to accept that they are unable to productionize their analytics initiatives.



Data Cruise

From the source to the AI Platform



Harvesting of data

The data that is stored or hidden in a conventional manner is extracted for a specific task.

Channelizing of data

Screen the data to identify the authenticity of the source, its significance and relevance in the present context.



Integration of data

Bringing data with similar attributes under a common group.

Application of data

Apply the data thus extracted to a real world business situation.



Enterprises use custom and manually coded approaches to modernize data pipelines. Let when an organization chooses an infrastructure, it relies on customized code bases and data engineers who are adept at coding skills. Customized coding reduces development costs by 20%.

Data engineers face the challenge of building effective data pipelines while simultaneously ensuring the right mix of data governance and data integrity. Data pipelines should enlighten and empower data scientists, analysts, AI and machine learning

Before setting foot on building a data pipeline, data engineers must understand that the market and technologies have reached a point, compelling organizations to



Unveil the data for use by a broader set of users without creating a need for any additional training or hiring expensive experts from outside.



Make their data system management less complex without compromising on its agility.



Scale resources effectively without losing an iota of data.



The data of today must serve to meet the needs of all, not a few. While keeping a close watch on how to guard and govern sensitive data, one must also bring it to a platform where it is accessible, usable and consumable by all.

In order to handle data that is historical, near to real time and real time; extract data that is buried in the abyss of information; assess its authenticity and put it to use for a situation at hand, an AI system is mandatory.

Accurate insights with clean data inputs

The insights derived by an AI system solely depends on the data inputs given to it. The significance of Data Engineering increases manifold in this context. In order to mine the appropriate data from the sea of useful and pretentious information, a prudent data engineering mechanism needs

A scalable, innovative engineering platform is what the data science and the AI team look forward to in order to productionize. The power of multi-cluster and multi-tenant, machine learning, artificial intelligence (AI) and business intelligence is unleashed with such a platform.

It would permit the user to easily spin up and down a cluster, ingest and process large volumes of data and control expenditures.



If the data is sensitive and needs to be strictly guarded, the data engineering solution must be flexible enough to allow the client to opt out of Cloud services and choose the same design as an On-Prem solution too.

An effective data engineering platform should have the following Value Drivers :



Provide the ability to adopt the required data for any given task.



Be an industry agnostic platform that is horizontally scalable to cater to the data engineering needs.



Possess the dynamic ability to build a use case or industry specific Data Lake.



Test the predictive model even before productionizing the power to foresee the prediction accuracy.



Productionizing efficacy



DATA FOR ALL

In a digitally borderless world, it is data that reigns. Although an organization generates data within its frontiers in huge proportions, not every data is useful at all times. It takes an effective data engineering platform to tell the difference between what is needed at the moment and what needs to be pushed into oblivion for now.

Data, artificial intelligence and analytics can help enterprises change the facade of business models, customer experiences and business operations. Organizations have, thus, started basing their decisions – including market capitalization, business development, capturing the market share - on data.

And high-level administrators of organizations have realized the significance of data in making the difference between success and failure. A McKinsey report “Advanced Analytics: Nine insights from the C-suite” states that more than 50 percent of CEOs consider themselves as the primary leader of the analytics agenda. Industry leaders unanimously agree that AI and true Data Engineering can capture nearly 1 trillion dollars in total economic profit across various emerging industries.



Phone: 972-756-1212 (Ext: 110)

Mail: marketing@intellisofttech.com

