Exploring the Synergy between Generative AI, Data and Analytics in the Modern Age

Pan Dhoni [1]

[1]Five Below Inc

October 31, 2023

Abstract

In the year 2023, a heightened sense of curiosity and apprehension pervaded the landscape of generative artificial intelligence (AI), particularly in the wake of the unveiling of the ChatGPT product by OpenAI. This pivotal moment sparked a flurry of discussions that predominantly revolved around the role of data in shaping the trajectory of generative AI. As researchers and organizations alike delved into this innovative realm, a pronounced inclination toward investigating its potential applications emerged. Notably, organizations swiftly recognized the transformative potential of generative AI in bolstering productivity across various sectors.

At the heart of these deliberations lies the profound significance of data. With data as the focal point, a compelling exploration began to unfold, with researchers keenly scrutinizing the ramifications of integrating generative AI within the domain of data and analytics. This research initiative was driven by an intrinsic desire to uncover the ways in which generative AI could be harnessed to enhance and streamline analytical processes.

In this context, the present research undertook a comprehensive investigation, employing a multifaceted approach. Leveraging various social media platforms as a primary source of insights, the research embarked on a journey to discern the prevailing sentiments, concerns, and expectations surrounding generative AI tools. This was further complemented by the execution of proof-of-concept (POC) endeavors, which not only enabled hands-on experience with generative AI tools but also facilitated a nuanced comprehension of their practical implications.

The culmination of these efforts yielded a series of noteworthy findings. Principally, it was discerned that enterprises stand to gain substantial benefits from embracing the capabilities of generative AI within the domain of data and analytics. The integration of generative AI tools offers the potential to revolutionize productivity, propelling organizations toward novel insights and expediting analytical processes. Concurrently, a strategic partnership with generative AI entities emerged as a salient consideration for safeguarding intellectual properties. Collaborative engagements between companies and generative AI providers became imperative to navigate the evolving landscape of data-driven innovation.

In conclusion, the year 2023 ushered in a period marked by intense curiosity and apprehension surrounding generative AI, catalyzed by the introduction of ChatGPT and its ensuing discussions. The centrality of data within this discourse propelled researchers and organizations toward an exploration of generative AI's potential applications, notably in the realm of data and analytics. Through a comprehensive research endeavor encompassing social media insights, POC experimentation, and practical insights, it became evident that the integration of generative AI could usher in transformative enhancements to productivity and analytical processes. In parallel, collaborative endeavors with generative AI entities emerged as a strategic imperative, offering a dual advantage of innovation and intellectual property protection. This research underscores the compelling need for enterprises to harness generative AI's capabilities, thereby positioning themselves at the vanguard of data-driven progress.

Exploring the Synergy between Generative AI, Data and Analytics in the Modern Age

Pan Singh Dhoni (ORCID: 0009-0006-7143-5353)

Abstract—In the year 2023, a heightened sense of curiosity and apprehension pervaded the landscape of generative artificial intelligence (AI), particularly in the wake of the unveiling of the ChatGPT product by OpenAI. This pivotal moment sparked a flurry of discussions that predominantly revolved around the role of data in shaping the trajectory of generative AI. As researchers and organizations alike delved into this innovative realm, a pronounced inclination toward investigating its potential applications emerged. Notably, organizations swiftly recognized the transformative potential of generative AI in bolstering productivity across various sectors.

At the heart of these deliberations lies the profound significance of data. With data as the focal point, a compelling exploration began to unfold, with researchers keenly scrutinizing the ramifications of integrating generative AI within the domain of data and analytics. This research initiative was driven by an intrinsic desire to uncover the ways in which generative AI could be harnessed to enhance and streamline analytical processes.

In this context, the present research undertook a comprehensive investigation, employing a multifaceted approach. Leveraging various social media platforms as a primary source of insights, the research embarked on a journey to discern the prevailing sentiments, concerns, and expectations surrounding generative AI tools. This was further complemented by the execution of proof-of-concept (POC) endeavors, which not only enabled hands-on experience with generative AI tools but also facilitated a nuanced comprehension of their practical implications.

The culmination of these efforts yielded a series of noteworthy findings. Principally, it was discerned that enterprises stand to gain substantial benefits from embracing the capabilities of generative AI within the domain of data and analytics. The integration of generative AI tools offers the potential to revolutionize productivity, propelling organizations toward novel insights and expediting analytical processes. Concurrently, a strategic partnership with generative AI entities emerged as a salient consideration for safeguarding intellectual properties. Collaborative engagements between companies and generative AI providers became imperative to navigate the evolving landscape of data-driven innovation.

In conclusion, the year 2023 ushered in a period marked by intense curiosity and apprehension surrounding generative AI, catalyzed by the introduction of ChatGPT and its ensuing discussions. The centrality of data within this discourse propelled researchers and organizations toward an exploration of generative AI's potential applications, notably in the realm of data and analytics. Through a comprehensive research endeavor encompassing social media insights, POC experimentation, and practical insights, it became evident that the integration of generative AI could usher in transformative enhancements to productivity and analytical processes. In parallel, collaborative endeavors with generative AI entities emerged as a strategic imperative, offering a dual advantage of innovation and intellectual property protection. This research underscores the compelling need for enterprises to harness generative AI's capabilities, thereby positioning themselves at the vanguard of data-driven progress.

Index Terms— advantages, ai, bi, chatgpt, data, etl, generative ai, genai, integration, middleware, pipeline, report, semantic tool, information technology

I. INTRODUCTION

The realm of artificial intelligence (AI) has long captivated the imagination of both researchers and the general public alike. Over the years, this fascination has steadily evolved from conceptual musings into tangible technological achievements. However, the recent surge in attention surrounding AI has been notably spurred by the confluence of two significant forces: the advancing frontiers of Data Science and the maturation of AI technologies. Central to this convergence lies the invaluable resource that propels the engines of innovation - data.

In this contemporary era, the significance of data has risen to unparalleled heights. This sentiment has been vividly underscored by the inception of OpenAI's ChatGPT, an exemplar of Natural Language Models (NLMs) birthed from the cradle of data-driven AI. As elucidated in our preceding paper, data has emerged as the modern-day "oil" that fuels innovation across diverse domains, from the automotive sector propelled by the achievements of companies like Tesla, to the dynamic landscapes of retail. Yet, beyond its application in specific industries, the symbiotic relationship between data and innovation resonates as a universal truth.

In the pursuit of understanding the intricate facets of this dynamic, the focus of this research paper is strategically set on exploring the nexus between Generative AI, Data, and Analytics in the contemporary milieu. Acknowledging that AI

itself is not a nascent concept, our inquiry homes in on the recent amplification of work in Data Science and AI, spotlighting data as the epicenter.

The advent of ChatGPT and similar endeavors serves as a compelling exemplar of data's centrality in AI. By harnessing voluminous datasets, OpenAI has pioneered the creation of models that can interact seamlessly with humans, bridging the gap between artificial and natural intelligence. However, the implications extend far beyond mere conversation, as the transformative power of AI is poised to recalibrate industries and domains that rely on data-driven insights. This paper seeks to delve into the depths of this transformation.

Furthermore, our investigation takes a step beyond the theoretical realm to discern practical applications. We endeavor to unravel the potential of Large Language Models (LLMs) like ChatGPT in revolutionizing the arena of Data and Analytics. In a world driven by technology, industries across the spectrum are awakening to the potential of these tools, ranging from ChatGPT's, BARD etc. Through this exploration, we aim to illuminate how the Information Technology (IT) industry can harness these novel toolsets to amplify productivity and catalyze innovation within the realm of Data and Analytics.

As we embark on this journey, our intent is to traverse the intricate intersections of Generative AI, Data, and Analytics, unearthing insights that offer both scholarly enrichment and pragmatic guidance. The subsequent sections of this research paper will traverse the historical trajectory of AI's evolution, magnify the symbiotic bond between data and innovation, and delve into the transformative potentials that LLMs harbor for the IT industry. Through this expedition, we endeavor to cast a light on the exciting prospects that await, forging a roadmap for stakeholders to harness the synergy between Generative AI, Data, and Analytics in the modern age.

II. DATA AND ANALYTICS

We need to familiarize ourselves with several terminologies, tools, and keywords commonly employed in the field of Data and Analytics.
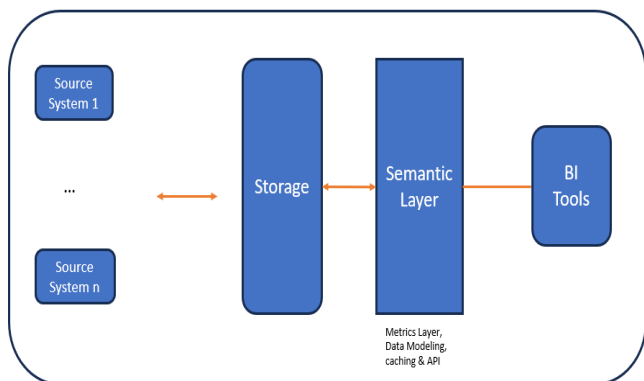

Fig 1. Data & Analytics, ingestion to Reporting.

Enterprise Applications Integration: In a diverse organizational landscape, a plethora of applications ranging from web-based and mainframe to standalone or desktop applications are utilized. These applications are developed using a wide array of programming languages, resulting in a significant variance in data generation formats. Some applications might exclusively process data in JSON format, while others utilize XML, fixed-line formats, or CSV. This diversity is extensive.

Within an organization, the imperative for applications to communicate with one another prevails; the isolated functioning of most applications is not feasible. Consequently, applications must possess the capability to comprehend the data formats of their counterparts. Analogously, consider an individual fluent in French attempting to communicate with someone who speaks Arabic. To facilitate their communication, an intermediary proficient in both languages is required for translation.

Similarly, within an organizational context, when an application operates on JSON data, it might need to interact with a publishing application that generates data in XML format. To bridge this disparity, an intermediary middleware tool or application is essential, capable of seamlessly translating JSON data into XML.

Enter middleware technology—an integral player in this scenario. It not only handles translations but also facilitates field-level transformations when necessary. The role of middleware technology becomes particularly pronounced in the domain of enterprise application integration [1].

In conclusion, the realm of enterprise application integration serves as the linchpin in harmonizing diverse applications, enabling them to communicate effectively and operate cohesively.

Middleware Tools: Amidst the myriad of available middleware [2] tools, we have chosen to spotlight a specific set in this research paper to underscore their particular relevance to our study:

1. IBM MQ (Message Queue): IBM MQ [3][4] is a messaging middleware that enables applications to communicate with each other asynchronously, providing reliable and secure message delivery. It supports various messaging patterns, ensuring that messages are delivered even if the sender or receiver is temporarily unavailable.

2. IBM Broker and IBM Integration Bus (IIB): These are integration platforms that facilitate the seamless exchange of data between various applications. IBM Broker [5] is used for message transformation and routing, while IBM IIB [6] provides a broader integration framework with features for routing, transformation, and protocol conversion.

3. Axway SFTP: Axway Secure File Transfer Protocol (SFTP) [7][8] is a tool used for secure and encrypted file transfers over a network. It ensures the confidentiality and integrity of data being transferred between systems.

4. TIBCO: TIBCO [9] offers a wide range of integration and analytics tools, including messaging, API management, and event processing. These tools enable businesses to connect applications, devices, and data sources in real-time.

5. Apache Kafka: Apache Kafka [10] is a distributed event streaming platform that is widely used for building real-time

data pipelines and streaming applications. It can handle high volumes of data and supports data processing in real-time.

6. MuleSoft: MuleSoft [11] provides an integration platform that allows organizations to connect various applications, data sources, and APIs. It simplifies the process of building application networks by enabling seamless communication between different systems.

7. IBM DataPower: IBM DataPower Gateway [12] is an appliance that helps secure, integrate, and optimize access to web, mobile, and API workloads. It provides features such as security enforcement, data transformation, and protocol mediation.

9. Legacy Applications:

 - *NEON:* NEON is likely a reference to Neon Systems, which used to offer mainframe integration solutions. These solutions would help modernize and integrate mainframe applications with contemporary systems.

 - *DATAGATE:* DATAGATE could be a proprietary integration solution or software used in a specific context. The specifics would depend on the organization's unique setup.

 - *SI, XI:* SI and XI could be acronyms for specific legacy integration tools or platforms. The exact interpretation would again depend on the context and organization.

Pipeline & ETL: A pipeline, within the context of data integration and enterprise application integration, can be conceptualized as a structured mechanism designed to establish a connection between distinct software systems, databases, or applications. Its fundamental purpose is to enable the orchestrated and streamlined transfer of data, messages, or information from one point to another within an organizational framework.

This process involves a series of well-defined steps that encompass data extraction, transformation, and loading (ETL), as well as potential data enrichment or manipulation. The pipeline acts as a conduit through which data moves seamlessly, often undergoing modifications to ensure compatibility and relevance in the target system.

Pipelines play a pivotal role in harmonizing heterogeneous systems by offering a standardized approach to data interchange. They encapsulate the necessary logic for data transformation, format conversion, and routing, thereby shielding the involved applications from intricacies associated with data translation and system-specific nuances.

By connecting systems through pipelines, organizations achieve a myriad of benefits. Data consistency and accuracy are maintained as the pipeline enforces standardized data formats and structures. Moreover, pipelines facilitate real-time or batch-based data movement, supporting both immediate insights and comprehensive data synchronization.

Pipelines can encompass various integration patterns, ranging from point-to-point connections to more intricate hub-and-spoke architectures. The former involves a direct link between two systems, while the latter employs a central integration hub to manage interactions among multiple systems. Such architectural decisions depend on factors like data complexity, system interdependencies, and scalability requirements.

In conclusion, a pipeline serves as the critical infrastructure that bridges the gap between disparate systems, enabling them to exchange data and insights cohesively. Its role extends beyond mere connectivity; it encompasses data transformation, enrichment, and orchestration, contributing significantly to the efficiency, accuracy, and agility of modern enterprise operations.

Semantic Layer: Certainly, here's a more detailed explanation of the sentence for your research paper:

In the realm of data management and analytics, the "semantic layer [13]" stands as a vital construct that facilitates effective communication between technical data structures and the needs of business users. It serves as a bridge, translating complex and technical data into a comprehensible business representation. This translation ensures that end users, who might not possess technical expertise, can access and analyze corporate data without grappling with intricate database structures or query languages.

The semantic layer encapsulates a set of business-focused abstractions, definitions, and relationships that align with the specific requirements and context of an organization. This layer acts as an intermediary that shields users from the underlying complexities of data storage, retrieval, and manipulation. Instead of directly querying databases or data warehouses, end users interact with the semantic layer, which offers a familiar business terminology and logical structure.

In the broader data architecture, the semantic layer typically resides between the data storage components (databases, data warehouses, etc.) and the tools used by end users for data consumption and analysis (such as reporting tools, dashboards, and visualization platforms). This positioning is strategic as it decouples the technical aspects of data storage and retrieval from the user-facing tools, promoting ease of use and flexibility.

The semantic layer plays several pivotal roles within an organization's data ecosystem:

1. Data Abstraction: It presents data in a way that aligns with how the business operates, masking the complexities of data storage and retrieval mechanisms.

2. Consistency and Reusability: By centralizing definitions and logic, the semantic layer ensures consistency and allows for the reuse of business rules and calculations across various reporting and analysis tools.

3. Data Governance: The layer can enforce data governance policies, ensuring that data is accessed and used in accordance with regulations and best practices.

4. Performance Optimization: It can optimize query performance by pre-calculating and storing results for commonly used queries, reducing the load on the underlying data storage systems.

In essence, the semantic layer empowers organizations to democratize data access, enabling business users to gain insights and make informed decisions without necessitating a deep understanding of the technical intricacies underlying the data infrastructure.

BI Reporting: In the realm of data-driven decision-making, the process of "visualizing data" emerges as a critical methodology that transforms raw information into meaningful insights. This transformation involves the presentation of data through visual elements like reports and dashboards [14], which are tailored to aid individuals in comprehending complex datasets and making more informed choices.

Data visualization employs a diverse range of graphical representations, such as charts, graphs, tables, and interactive visualizations, to convey patterns, trends, and relationships within data. This visual representation serves as a powerful means of communication, as humans are inherently adept at processing visual information. By translating data into visual formats, organizations enhance their ability to understand the significance of the data and identify actionable insights.

Reports and dashboards are two primary forms of data visualization:

1. Reports: Reports encapsulate summarized information derived from raw data. They often include key performance indicators (KPIs), metrics, and textual descriptions. Reports present data in a structured format, allowing decision-makers to quickly assess the current state of affairs and identify any anomalies or areas requiring attention. Reports are particularly valuable for regular updates, trend analysis, and compliance reporting.

2. Dashboard: Dashboards provide a more interactive and dynamic way of visualizing data. They consist of a collection of visual elements, such as charts, graphs, gauges, and filters, presented on a single screen. Dashboards enable users to explore data from multiple perspectives, drill down into specific details, and uncover insights in real time. They cater to different user roles and allow customization, empowering individuals to focus on the information that's most relevant to their roles.

The ultimate goal of visualizing data through reports and dashboards is to facilitate better decision-making. Here's how:

1. **Accessibility:** Data visualization makes complex data accessible to a wider range of individuals, irrespective of their technical expertise. This inclusivity ensures that insights are available to decision-makers across the organization.

2. **Insight Discovery:** Visual representations expose patterns, trends, and correlations that might be obscured within raw data. Decision-makers can identify opportunities and challenges more readily.

3. **Communication:** Visualizations simplify the communication of insights. They enable teams to collaborate effectively, align their understanding, and collectively decide on appropriate actions.

4. **Quick Assessment:** Visualizations allow for rapid assessment of data, making it easier to grasp the overall situation and make swift decisions.

5. **Prediction and Planning:** By visualizing historical data, decision-makers can extrapolate trends and use these insights to inform future planning and strategies.

In conclusion, visualizing data through reports and dashboards is an indispensable tool for enhancing decision-making processes. It empowers organizations to extract insights, communicate effectively, and navigate the complexities of data-driven strategies with greater confidence.

Data Science: In accordance with the definition provided by Wikipedia, data science [15] emerges as a dynamic interdisciplinary domain that seamlessly integrates various academic facets. Encompassing statistics, scientific computing, rigorous methodologies, advanced algorithms, and intricate systems, data science serves as a conduit to unveil profound knowledge and insights from diverse datasets—ranging from the well-structured to the enigmatically unstructured. By diligently navigating through the realm of noisy and intricate data landscapes, data science stands poised to illuminate the hidden narratives and correlations that underlie complex real-world phenomena.

When discussing Data and Analytics, the scope extends from collecting data from systems of record to organizing it into appropriate model tables at a centralized location, ultimately leveraging it for reporting and forecasting purposes.

The process of acquiring data encompasses various modes such as real-time, near-real-time, or batch processing. System of record diversity ranges from legacy systems to advanced platforms, necessitating an array of connectors to seamlessly bring data into the designated queue.

If we reflect back to the early 2000s, an assortment of tools like NEON, DATAGATE, SI, XI, and even custom Python or shell script-based frameworks were prevalent for extracting data from systems of record. The landscape saw an influx of Enterprise Integration tools like IBM MQ, Broker, TIBCO, alongside proprietary solutions from SAP and Oracle, contributing to an expanding toolkit.

Reporting also presented numerous solutions, including Power BI and MicroStrategy. The advent of cloud computing further diversified options, with each cloud provider offering analogous capabilities.

Looking ahead to the Generative AI era, it is anticipated that these companies will augment their tool capabilities by integrating them with AI functionalities, shaping the future of data and analytics.

Returning to the central focus, the following activities are crucial within the realm of Data and Analytics:

Acquiring Data:
1. Data Mapping: Understanding and defining data sources and destinations.
2. Ingestion: Gathering data using APIs or in real/near-real-time or batch modes.

ETL (Extract, Transform, Load):
1. Cleaning and Transformation: Refining and restructuring data for analysis.

Aggregating Data for Reporting:
1. Joining and Filtering: Creating unified datasets by combining or refining data from various sources.

Integrating BI Tools:

1. Utilizing Business Intelligence tools for generating insights and reports.

Projecting Data for Future Forecasting (Data Science):

Ascertaining trends and patterns in data to predict future outcomes.

As we explore the potential of Generative AI, it's pertinent to consider how it could revolutionize these areas, enhancing data acquisition, transformation, and analysis for improved decision-making and predictive capabilities.

We can go step by step and explain in each area how it is going to be helpful, but again it is not necessary it is going to replace expert systems but will have certainly impact on productivity. But possibly small organization can still use for their respective use cases.

III. WHAT IS ARTIFICIAL INTELLIGENCE (AI)?

This section is reference from the research paper [16].

Artificial intelligence (AI) is intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by humans or by other animals.

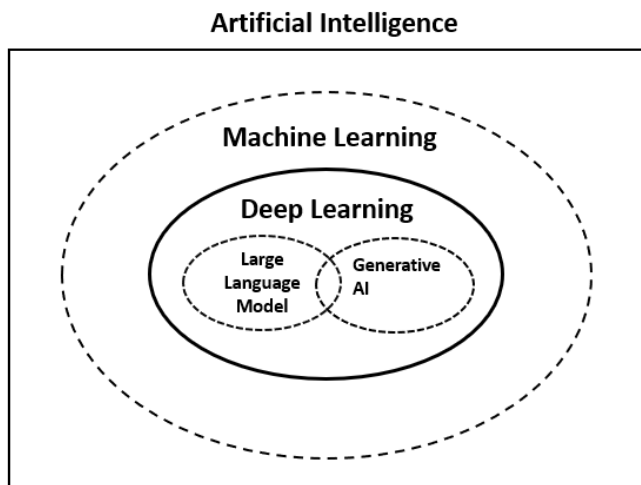**Artificial Intelligence**

Fig. 2. The relationship between AI, Machine Learning, and deep learning

Machine learning is a subset of AI, and deep learning (Natural Language processing (NLP), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN)) is a subset of machine learning (See Fig 1). Deep and machine learning have two important types: supervised learning (labelled data) and unsupervised learning.

IV. ARTIFICIAL INTELLIGENCE (AI) BRIEF HISTORY

This section is reference from the research paper [16].

70 years ago, the first seeds of AI were planted when Alan Turing, in his 1950 paper titled "Computer Machinery and Intelligence," posed the fundamental question, "Can machines think? [17]".

The term "artificial intelligence" was coined by John McCarthy in 1956 during the Dartmouth Summer research project [18].

In 1958, John McCarthy developed the programming language LISP, which became instrumental in AI research and remains influential in the field to this day [19].

In the 1960s, Joseph Weizenbaum created ELIZA, a computer program designed to simulate conversation and demonstrate the potential of natural language processing.

In 1965, Ivakhnenko and Lapa published the groundbreaking paper "Group Method of Data Handling - A Rival Method of Stochastic Approximation," which presented the first functional neural network.

The 1970s witnessed advancements in knowledge representation with the development of expert systems. The MYCIN system, developed by Edward Shortliffe [20], showcased the ability of computers to diagnose diseases based on expert-level knowledge.

In the 1980s, the field of AI experienced a shift towards knowledge-based systems and symbolic reasoning. The development of rule-based expert systems like XCON and R1 showcased the potential of AI in various domains.

In the 1990s, AI research saw a resurgence with the introduction of machine learning techniques. The field witnessed breakthroughs in areas such as neural networks, reinforcement learning, and genetic algorithms.

In recent years, advancements in deep learning, fuelled by the availability of big data and improvements in computational power, have revolutionized AI applications. Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have achieved remarkable success in image recognition, natural language processing, and other complex tasks.

Today, AI is integrated into our daily lives, powering virtual assistants, autonomous vehicles, recommendation systems, and much more. Ongoing research and development continue to push the boundaries of AI, aiming to create systems that can reason, understand, and learn like humans.

The history of AI is marked by a series of milestones, breakthroughs, and paradigm shifts, laying the foundation for the ever-evolving field that continues to shape our world.

V. NEURAL NETWORK

This section is reference from the research paper [16].

In our day-to-day lives, inspiration can come from various sources, including other humans, nature, animals, and birds, leading to greater innovation. The quest for innovation has led to the development of intelligent machines, which has given rise to the concept of Artificial Neural Networks (ANNs).

The exponential growth in computing power over the past few decades, particularly since the 1990s, has facilitated the training of large-scale neural networks within a feasible timeframe. This remarkable progress can be attributed not

only to the advancements dictated by Moore's Law but also to the contributions of the industry, which has manufactured and distributed millions of powerful GPU cards, further enhancing computational capabilities.

Before going into the ANNs, lets discuss biological neuron.

Biological Neurons: According to the Oxford dictionary, a neuron is defined as a specialized cell responsible for transmitting nerve impulses within a nervous system.

Wikipedia describes a neuron as an electrically excitable cell within a neural network that generates and propagates electric signals known as action potentials.

Below reference is taken from the [21] chapter "Neural network and deep learning".

A biological neuron (see Fig 2) consists of a cell body that houses the nucleus and most of the cell's complex components. It also has branching extensions called dendrites and a long extension known as the axon. The axon can be several times longer than the cell body or even tens of thousands of times longer. Towards the end of the axon, it splits into multiple branches called telodendria, which terminate in tiny structures called synaptic terminals or synapses. These synapses are connected to the dendrites or directly to the cell body of other neurons. Biological neurons receive short electrical signals known as impulses from other neurons through these synapses. When a neuron receives a sufficient number of signals from other neurons within a few milliseconds, it generates its own signals.
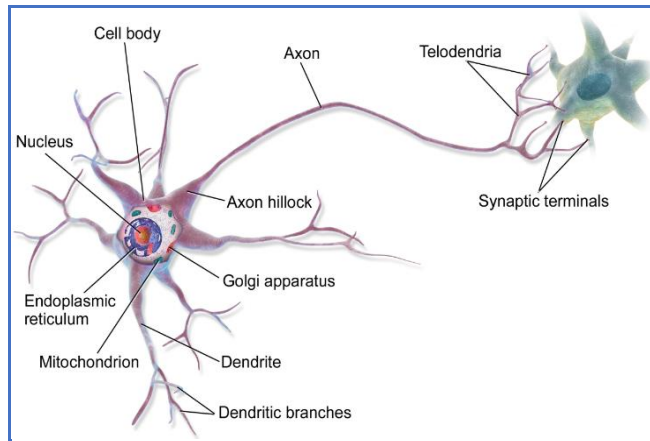

Fig. 3. Biological Neurons

Without diverting the topic, lets discuss briefly on ANNs.

In a neural network, there are multiple layers that work together to process and analyse data. The first layer is the input layer, where the data is initially fed into the network. From the input layer, the data is passed through a complex hidden layer. The hidden layer consists of nodes or neurons that apply mathematical operations such as adding weights and biases to the input data. These weights and biases help the network adjust the importance of different features in the data. Additionally, a threshold is applied to determine whether a neuron should activate or not based on the weighted sum of inputs.

During the process known as forward propagation, the data is passed through each layer of the neural network, with calculations performed at each neuron to generate output values. These output values are then passed to the next layer as inputs, and the process continues until the data reaches the output layer.

As the data passes through the hidden layers, the neural network learns by adjusting the weights and biases through a process called backpropagation. Backpropagation involves comparing the network's predictions with the expected outputs, calculating the errors, and then propagating these errors backward through the network. The network updates the weights and biases based on the magnitude of the errors, aiming to minimize them and improve the overall performance of the network.

The training process involves iterative steps of forward propagation and backpropagation, with the network making predictions, comparing them to the expected outputs, and adjusting the internal parameters (weights and biases) accordingly. This iterative process helps the network learn to recognize patterns and make accurate predictions based on the provided training data.

Overall, a neural network consists of an input layer, one or more hidden layers, and an output layer. Through the process of forward propagation and backpropagation, the network learns to adjust its internal parameters to make accurate predictions or classifications based on the provided input data.

VI. GENERATIVE AI

This section is reference from the research paper [16].

Generative AI [22][23] has the capability to learn from existing artifacts and generate new, realistic creations that capture the characteristics of the training data without merely duplicating it. This technology can produce diverse forms of original content, including images, videos, music, speech, text, software code, and product designs.
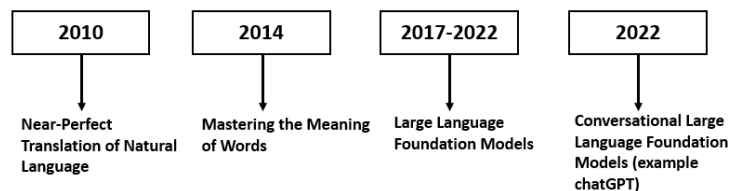

Fig. 4. Generative AI Timeline [2]

See GenAI Timeline (Fig 3).

As GenAI has been widely used, it is more likely that significant ongoing data on the internet will be generated using chatbots.

Generative models in AI, such as Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), utilize neural networks to generate new data samples. The core mathematical formulas for these models are as follows:

Variational Autoencoder (VAE):
The VAE consists of an encoder network that maps the input data to a latent space and a decoder network that generates output data from points in the latent space. The objective is to learn the parameters of the encoder and decoder networks to maximize the evidence lower bound (ELBO), which is equivalent to minimizing the reconstruction loss and the Kullback-Leibler (KL) divergence.

Reconstruction Loss:
$L\_recon = -\sum(x \log(x\_hat) + (1 - x) \log(1 - x\_hat))$

KL Divergence:
$L\_KL = 0.5 * \sum(1 + \log(var) - mean^2 - var)$

Overall Loss:
$L\_total = L\_recon + L\_KL$

Generative Adversarial Network (GAN):
The GAN consists of a generator network that generates synthetic samples and a discriminator network that tries to distinguish between real and synthetic samples. The objective is to find a Nash equilibrium between the generator and discriminator, which is achieved through an adversarial training process.

Generator Loss:
$L\_G = -\log(D(G(z)))$

Discriminator Loss:
$L\_D = -\log(D(x)) - \log(1 - D(G(z)))$

Overall Loss:
$L\_total = L\_G + L\_D$

In both VAEs and GANs, the parameters of the neural networks (weights and biases) are optimized using gradient-based optimization algorithms such as stochastic gradient descent (SGD) or its variants, by computing the gradients of the loss function with respect to the network parameters and updating them iteratively.

VII. METHODOLOGY

Few methods have been taken from research paper [16][24][25]

Method 1- Conducting social media research has entailed a comprehensive exploration across diverse platforms, ranging from industry giants like Google, YouTube, and LinkedIn to the innovative realm of ChatGPT. This multifaceted approach has enabled me to glean insights from a spectrum of sources, each contributing to a holistic understanding of the subject matter.

The array of perspectives I've examined spans the entire corporate spectrum, encompassing insights from data heads of both prominent enterprises and emerging small-scale ventures.

This gamut of expertise provides a nuanced tapestry of opinions, strategies, and experiences, offering a comprehensive view that transcends organizational sizes.

Engaging with these diverse viewpoints has not only enriched the depth of my research but has also fortified its relevance. The amalgamation of insights from distinguished experts affiliated with large enterprises to the intricate narratives woven by data leaders in smaller companies has bestowed my research with a mosaic of perspectives, fostering a well-rounded analysis.

Through this collective exploration, my research emerges as a synthesis of thoughts that echoes both the resounding voices of established industry leaders and the intricate harmonies of nimble, emerging players. This synthesis encapsulates not only the breadth of the social media landscape but also the intricate interplay of challenges, strategies, and opportunities faced by a spectrum of organizations.

In essence, my research paper stands as a testament to the diverse dimensions of social media research, underpinned by the chorus of insights gleaned from experts who collectively shape the narrative of this dynamic domain.

Method 2- I embarked on a preliminary proof-of-concept (POC) exploration centered around the capabilities of generative AI. This endeavor encompassed the creation of a chatbot, realized through a multifaceted approach employing technologies such as NodeJS, Python, and Databricks. Moreover, a web application was developed using function app architecture, while an integration pattern was strategically deployed to facilitate the dynamic re-generation of images.

In parallel, a comprehensive infrastructure for data management was constructed. This entailed the creation of data models and tables, providing a structured foundation for subsequent analysis. The analytical facet was robustly addressed through the utilization of ChatGPT, affording the generation of reports and facilitating data analysis.

The amalgamation of these undertakings culminated in an intricate yet synergistic POC ecosystem. This ecosystem not only showcased the versatile potential of generative AI but also underscored the diverse technologies, platforms, and paradigms necessary to harness its capabilities effectively. This POC venture emerges as a microcosm of the broader landscape of generative AI integration, encapsulating the intricacies of development, integration, and data-driven insights.

Method 3- In crafting this research paper, I have harnessed the cumulative wisdom amassed over two decades of immersion in the dynamic landscape of computer technology. This extended period has allowed me to traverse an intricate journey, marked by exposure to an eclectic array of integration and analytics tools, spanning the spectrum from antiquated legacy systems to the cutting-edge marvels of the modern era.

In this journey, the roster of encountered tools reads like a historical mosaic of technological evolution. Noteworthy names such as NEON, DATAGATE, SI, and XI stand as testament to the diverse palette of solutions that were instrumental in the earlier epochs of integration. This lineage segued seamlessly into the domain of IBM's trailblazing integration suite, replete with stalwarts like IBM MQ, IBM Broker, Integration Bus (IIB), and IBM DataPower. Parallel to these proprietary offerings, my journey also intersected with

the vibrant realm of open-source tools like Kafka and Spark, effectively contributing to the diverse toolkit at my disposal. This veritable cavalcade of tools, though far from exhaustive, underlines the substantial array with which I've engaged.

Yet, my journey wasn't confined to toolsets alone. My exploration encompassed a dynamic spectrum of technological milieus. Transitioning between cloud-based paradigms and traditional on-premise environments, I've grappled with the nuances and implications inherent to both. This fluidity in adaptation allowed me to fathom the ever-shifting tapestry of technological landscapes.

The linguistic gamut was not overlooked. My journey traversed the landscapes of programming languages, oscillating adeptly from foundational languages like Java and .Net to the more contemporary scripting prowess of Python. This cross-linguistic competency ensured that I could decipher and engage with diverse ecosystems of technology.

In this research endeavor, the culmination of my experiences and knowledge is harnessed to fathom the enigma that is generative AI. This multifaceted exploration probes beyond the surface, dissecting its implications with a discerning eye. How might generative AI transmute the paradigms of Data and Analytics? This question, embedded within the core of my investigation, forms the fulcrum upon which my research pivots. My extensive foray into integration, analytics, and diverse technological domains serves as the scaffolding upon which I ascend towards the anticipated dawn of generative AI's potential.

In summary, this research paper is not merely a presentation of findings; it is a testament to a lifelong odyssey in the realm of computer technology. It exemplifies how each twist and turn, every encounter with a tool, a language, or a technological trend, has meticulously crafted the foundation for understanding and embracing the transformative power of generative AI within the landscape of Data and Analytics.

VIII.   RESULTS

A. Ingestion Area
Through our comprehensive research, we've unearthed numerous compelling use cases that exemplify the potential of harnessing generative AI for data ingestion. Take, for instance, the scenario where a PDF file needs to be ingested, and the content extracted or synthesized based on images or maps. By supplying latitude and longitude coordinates, generative AI can facilitate the enhancement of photos and videos, accompanied by accurate descriptions that elevate the informative value.

Another intriguing use case involves leveraging ChatGPT to generate code in various programming languages for the purpose of ingesting data from systems of records. This innovative approach offers the convenience of a code-generation assistant, expediting the process of integrating data from diverse sources. An example query could be as follows: "Provide me with code to perform real-time data ingestion from systems.

Furthermore, the realm of API integration presents yet another avenue where generative AI can excel. By crafting APIs or harnessing existing ones, data ingestion from source systems can be made seamless and efficient.

A notable trend observed is the remarkable evolution of data pipeline development. The landscape has transitioned from labor-intensive, verbose codebases to lean, concise, and even no-code (mapping-based) pipelines. An array of integration tools has surfaced to cater to a spectrum of real-time, near real-time, and batch data movement needs. In the midst of this transformative landscape, the emergence of generative AI has been particularly impactful. Companies, in their pursuit of staying relevant and efficient, are increasingly investing in natural language-based capabilities to streamline the pipeline construction process.

Imagine a scenario where a developer requests the creation of a data pipeline to transfer a CSV file from System A to System B. In response, generative AI constructs an automated pipeline, intelligently suggesting mapping schemas and seamlessly linking systems. Notably, SnapGPT stands as a prime example of this paradigm, as evidenced by its employment of the ChatGPT API to empower such capabilities. This transformative approach is elaborated further in a detailed article [26].

While the potential of generative AI is undeniable, it's crucial to acknowledge that it's still in its nascent stages. It might not yet encompass an exhaustive array of connectors to seamlessly integrate with the wide gamut of systems. However, the promise it holds cannot be dismissed, and as the technology matures, its scope and capabilities are anticipated to expand exponentially.

B. For Testing Team
The paradigm of Generative AI introduces a novel avenue for data generation through the creation of synthetic or fabricated data. This capability holds particular relevance in the domain of testing and performance testing. Often, the necessity for data to facilitate testing procedures arises, necessitating the creation of sample datasets. Conventionally, this process involves the manual authoring of scripts—a time-consuming endeavor. However, the emergence of generative AI offers a transformative alternative. By harnessing generative AI tools, organizations can expedite the generation of sample data. This synthetic data, meticulously crafted by AI, serves as a valuable resource for testing purposes, negating the need for manual scriptwriting and expediting the testing lifecycle.

Moreover, the impact of Generative AI is not limited to data generation; it extends to the realm of generating test cases. Traditional test case development can be an intricate and time-intensive task. Generative AI offers a solution to this challenge by enabling the rapid creation of test cases. This efficiency is particularly noteworthy, as it translates into reduced testing timelines and enhanced agility in software development cycles. By leveraging Generative AI, organizations can swiftly formulate a comprehensive suite of test cases, ensuring rigorous testing while maintaining an accelerated pace of development.

In summary, Generative AI presents a transformative opportunity to revolutionize the data generation process, yielding synthetic datasets that prove invaluable for testing

scenarios. This innovation extends to the expedited creation of test cases, thereby enhancing the efficiency and effectiveness of testing procedures within the software development lifecycle.

C. Data Analyst Team

Within the domain of data analysis, the integration of generative AI tools introduces a paradigm shift in how data analysts approach their tasks. These tools offer a gamut of capabilities that significantly augment the productivity and effectiveness of data analyst teams.

One compelling facet lies in the realm of SQL query generation. Crafting precise and efficient SQL queries is a cornerstone of data analysis, but it often demands meticulous syntax and logical structuring. Generative AI tools serve as invaluable aids, automatically generating SQL queries based on the specified criteria. This not only expedites the query formulation process but also minimizes the likelihood of human errors, ultimately enhancing the quality of analysis.

Moreover, the proficiency of generative AI extends beyond mere query generation. These tools empower data analysts to delve deeper into data analysis by producing comprehensive reports and visualizations. The automated creation of sample reports, accompanied by insightful data analyses and meticulously organized tables, streamlines the analytical process. By integrating these AI-driven capabilities, data analysts can allocate more time to interpretative tasks and strategic insights, rather than getting bogged down by the intricacies of report creation.

In essence, generative AI becomes a transformative partner for data analyst teams. By automating SQL query generation, report creation, and data analysis, these tools unlock new dimensions of efficiency, accuracy, and agility within the data analysis workflow. This symbiotic relationship between data analysts and AI fosters a more dynamic and impactful approach to harnessing data for informed decision-making.

D. Data Engineering Team

Within the landscape of data engineering, the composition of SQL or Python-based code stands as an essential and recurrent task. The advent of generative AI tools heralds a transformative shift in this paradigm, offering a profound boon to data engineering teams.

One of the primary mandates of data engineering involves crafting intricate code to manipulate, transform, and structure data effectively. This often demands meticulous attention to syntax, logic, and optimization. Generative AI tools enter this scenario as instrumental allies, capable of automatically generating the desired SQL or Python code based on specified parameters and objectives. This augmentation of data engineering capabilities serves as a catalyst for heightened productivity and precision.

By leveraging these AI-driven tools, data engineering teams are liberated from the burdensome intricacies of manual code creation. Time and effort are redirected towards more strategic endeavors, such as devising innovative data pipelines, optimizing data flows, and enhancing system performance. The resultant productivity surge not only expedites project timelines but also fosters an environment where creativity and problem-solving take precedence over routine code-writing tasks.

In conclusion, the integration of generative AI tools within the realm of data engineering marks a pivotal advancement. This symbiotic relationship between human expertise and AI-driven automation cultivates a landscape where data engineering teams can dedicate their talents to higher-order challenges, secure in the knowledge that the meticulous and routine aspects of code generation are efficiently handled by intelligent systems.

E. BI Area

Generative AI emerges as a pivotal tool to extract enhanced business intelligence from data, fundamentally transforming how business partners engage with data-driven decision-making processes. This technology empowers business partners to swiftly generate timely mock-up data, effectively streamlining the process of gaining sign-off from stakeholders. This pre-emptive approach enables the presentation of high-level outcomes in advance, providing a tangible understanding of projected outcomes to the business.

In this collaborative landscape, analysts and developers stand to benefit significantly. These timely and insightful mock-ups can be harnessed as a foundation for both the analysis and development stages. Analysts can adeptly translate these mock-ups into sophisticated reports, foretelling the trajectory of insights before their formalization. Similarly, developers can draw from these previews to refine their integration processes, ensuring that the resultant data streams are seamlessly aligned with business objectives.

Furthermore, the potential synergy between Generative AI and platforms like Power BI unfolds a new realm of possibilities. Direct integration between ChatGPT and Power BI can empower developers to leverage the prowess of the renowned OpenAI model. ChatGPT's capabilities extend to assisting in the construction of intricate calculations, formulating advanced queries for Power BI models, conducting error diagnostics, and optimizing the process of report generation.

In sum, Generative AI's impact spans the entire spectrum of data-driven decision-making. From providing tangible mock-ups to expediting analysis and development processes, this technology is poised to redefine the efficiency and efficacy of business intelligence endeavors, enhancing collaboration and elevating the insights derived from data.

F. Code optimization

An enduring challenge pervading industry, regardless of whether they operate in the cloud or on-premises, revolves around the optimization of code. The fundamental principle at play is that the efficiency of your code directly impacts the consumption of CPU cycles. Code that is streamlined and optimized translates into minimized resource usage, a factor that invariably contributes to significant cost savings. Herein lies one of the core advantages offered by generative AI—it possesses the capacity to aid in code optimization, effectively addressing a pivotal cost-related concern.

The potential to harness generative AI for code optimization introduces a transformative avenue for businesses aiming to curtail expenses. By leveraging AI-driven techniques to fine-tune code, organizations can reduce the strain on computational resources, thereby diminishing operational costs. This advantage is especially pertinent in the

context of modern computing paradigms, where scalability and resource efficiency are paramount.

However, while the prospects are promising, it's imperative to approach this avenue with a discerning eye. The application of generative AI for code optimization demands a meticulous validation process. Results need to be rigorously tested and cross-verified, ensuring that the optimizations achieved align with the desired outcomes. Furthermore, any alterations introduced through AI-driven optimization should not compromise the integrity or functionality of the codebase. It's an arena where vigilance is paramount to ensure that potential benefits are realized without introducing unforeseen complexities or issues.

In essence, while the promise of generative AI in code optimization holds immense potential, it must be approached with due diligence. Its impact can be revolutionary, leading to substantial cost savings, yet the veracity of its optimizations should be affirmed through careful testing and validation protocols.

G. For Business partner

In the current landscape, chatbots have ascended to the zenith of popularity, particularly as products of various generative AI companies. However, an alternative approach lies in the realm of prompt engineering—an avenue that empowers organizations to cultivate similar capabilities in-house. By embracing prompt engineering, companies can construct tailored chatbot solutions that cater to their specific needs, enabling them to bestow English language proficiency upon their business partners. This, in turn, equips partners with the capacity to receive accurate and swift responses to their queries, further enhancing operational efficiency.

The allure of this strategy lies in its potential to foster self-sufficiency and versatility. In-house development enables organizations to craft chatbots that are attuned to their distinct operational paradigms and linguistic nuances. This degree of customization empowers business partners with the ability to swiftly and effectively retrieve information and insights. The result is an enriched collaboration that thrives on seamless communication and prompt access to essential data.

Nevertheless, while the prospect is promising, it's paramount to tread carefully in the implementation of this approach. Rigorous validation and verification processes are necessary to ensure the accuracy, reliability, and consistency of the responses generated by these in-house chatbots. As these chatbots play a crucial role in communication, any inaccuracies or misinterpretations could potentially yield undesirable outcomes. Thus, while the approach holds immense potential, its deployment should be accompanied by a stringent verification regimen.

In summation, the realm of chatbots and generative AI is evolving rapidly, offering multiple avenues for organizations to harness these capabilities. The choice between leveraging established chatbot solutions or cultivating in-house alternatives necessitates careful consideration, validation, and continuous assessment. While the allure of efficiency and autonomy is compelling, the veracity and precision of the responses should be the guiding force behind its deployment.

H. Reviewer

The integration of Generative AI into the realm of code review and documentation holds immense potential for augmenting efficiency and productivity. When integrated into an organization's workflow following due approval, Generative AI tools can significantly streamline the review process and documentation tasks.

In the context of code review, Generative AI offers the capability to automatically analyze and assess code snippets. This automation can expedite the identification of potential errors, inefficiencies, or deviations from best practices. As a result, the time spent on manual code inspection is reduced, allowing developers to focus more on refining the code's logic and functionality. This integration has the potential to considerably reduce the review time, enhancing productivity while maintaining code quality.

Similarly, Generative AI's involvement in documentation tasks has transformative implications. The process of drafting comprehensive documentation can be time-consuming, often necessitating meticulous attention to detail. Generative AI can alleviate this burden by automatically generating well-structured and coherent documentation based on the provided inputs. The result is not only a reduction in documentation creation time but also an assurance of consistency and accuracy across the documentation suite.

For instance, let's consider an organization following an agile process. Tasks are often assigned a certain point value in tools like Jira to indicate their complexity and estimated effort. By harnessing Generative AI capabilities, a task that is assigned, for example, 2 Jira points might be streamlined to completion within a matter of hours rather than days. This acceleration stems from the optimization of code review and documentation processes, translating directly into enhanced agility and time-to-market.

In summary, the integration of Generative AI into code review and documentation workflows presents a compelling opportunity to revolutionize conventional processes. By automating aspects of code review and documentation creation, organizations stand to benefit from reduced review times, heightened productivity, and the capacity to navigate agile processes with greater efficiency.

I. Cloud cost optimization

Generative AI has the potential to revolutionize cloud cost-saving and optimization strategies through its capacity to automate and enhance various aspects of cloud resource management. Here's how generative AI can be utilized for cloud cost-saving and optimization:

1. Resource Allocation and Scaling: Generative AI can analyze historical usage patterns, workload demands, and performance metrics to predict optimal resource allocation and scaling strategies. By identifying when and how to scale resources, organizations can prevent overprovisioning, ensuring that they only pay for the resources they truly need.

2. Automated Rightsizing: Generative AI can assess the utilization of virtual machines and containers and recommend appropriate resizing based on actual usage. This helps avoid underutilization of resources, leading to significant cost savings.

3. Serverless Workload Optimization: For serverless computing, generative AI can analyze functions' resource consumption and execution times to suggest optimizations

that reduce runtime and resource consumption, directly impacting costs.

4. Spot Instance Utilization: Generative AI can predict when cloud providers' spot instances will be available and recommend utilizing these cost-effective instances for non-time-sensitive workloads.

5. Cost-Aware Architecture: Generative AI can assist in designing cost-effective cloud architectures by suggesting the optimal distribution of workloads across various cloud services based on their cost-performance trade-offs.

6. Data Storage and Retrieval: AI can analyze data access patterns and suggest data storage tiers (e.g., hot, cool, archive) to optimize data storage costs.

7. Workload Scheduling: By understanding workload priorities and resource availability, generative AI can optimize workload scheduling, minimizing contention for resources and potentially reducing the need for excess capacity.

8. Anomaly Detection and Cost Alerts: AI-powered anomaly detection can identify unexpected spikes in resource consumption, helping organizations detect and address cost-related issues promptly.

9. Optimized Data Transfer: Generative AI can recommend efficient data transfer strategies, such as selecting appropriate regions for data storage to reduce transfer costs and latency.

10. Predictive Cost Analysis: By analyzing historical data and cloud usage patterns, AI can predict future costs and help organizations proactively manage and budget for cloud expenses.

11. Energy-Efficient Computing: AI can guide organizations in selecting energy-efficient data centers and regions, aligning with sustainability goals while potentially reducing energy-related costs.

In summary, generative AI's data-driven insights, predictive capabilities, and automation potential make it a powerful ally in the pursuit of cloud cost-saving and optimization. By leveraging AI-driven recommendations and strategies, organizations can achieve more efficient resource utilization, enhanced performance, and substantial reductions in cloud-related expenditures.

IX. CONCLUSION

As we are well aware, the prevailing market dynamics drive many aspects of business decisions. Notably, companies are progressively allocating greater financial resources toward the development and deployment of generative AI technologies. This investment imparts a twofold impact—firstly, it resonates within the marketing sphere, enhancing a company's positioning and outreach. Simultaneously, it ushers in benefits for customers by facilitating more advanced and refined products and services.

Following our comprehensive research, we have distilled the essence of generative AI's potential within the realms of Data and Analytics into the following overarching summaries:

Building upon our earlier discussion, we foresee a maturation of middleware tools and Business Intelligence (BI) platforms to encompass AI capabilities. This evolution will offer customers the opportunity to seamlessly integrate AI functionalities, facilitating a stepwise and incremental approach toward desired outcomes. Importantly, this does not render middleware tools obsolete; rather, it underscores their adaptability and ongoing relevance within an evolving landscape.

A second focal point revolves around the substantial enhancement of productivity through the integration of generative AI. This amplification can span a spectrum of functions, encompassing data ingestion, analysis, testing, and reporting. By automating these processes, generative AI bolsters the efficiency of data-centric tasks, contributing to swift and agile decision-making.

A pivotal consideration pertains to the economic implications of adopting generative AI. While the technology bears immense potential, its operational framework often necessitates the use of high-performance GPU machines, entailing associated costs. Thus, organizational stakeholders must meticulously assess the balance between investment in generative AI and the resultant returns on investment.

A fourth significant facet entails the refinement of requirements and the corresponding alignment of businesses with these evolved prerequisites. Simultaneously, companies need to recalibrate their estimation models, as the integration of generative AI capabilities invariably shapes and elevates overall productivity benchmarks.

Anticipating the trajectory ahead, the market will witness the proliferation of a plethora of tools empowered by generative AI. Noteworthy among these is the emergence of snapLogic (SnapGPT), RATH etc., which employs ChatGPT for data analysis endeavors.

In summation, generative AI's pervasive influence transcends multiple dimensions. It propels marketing strategies and customer experiences, aligns with existing tools, enhances productivity across various stages, prompts cost-benefit analysis, redefines business requirements, and heralds a future where AI-infused tools abound.

academic pursuit. Lastly, we extend our sincere appreciation to the creators of ChatGPT, whose assistance in refining grammar and sentence structure has been invaluable in enhancing the quality of our paper.

REFERENCES

[1] 02:04, 21 October 2022
https://en.wikipedia.org/w/index.php?title=Enterprise_application_integration&oldid=1117316924
[2] https://www.ibm.com/topics/middleware
[3] Davies, Saida, et al. WebSphere MQ v6 Fundamentals. IBM Redbooks, 15 Dec. 2005.
[4] WebSphere MQ V7.0 Features and Enhancements. IBM Redbooks, 26 Jan. 2009
[5] https://www.ibm.com/topics/message-brokers
[6] https://www.ibm.com/docs/en/integration-bus/10.0?topic=overview-integration-bus-introduction
[7] https://docs.axway.com/bundle/Gateway_6161_UsersGuide_allOS_en_HTML5/page/Content/protocols/sftp/sftp_about.htm
[8] https://www.axway.com/en/products/managed-file-transfer
[9] https://www.tibco.com/
[10] Shapira, Gwen, et al. Kafka: The Definitive Guide Real-Time Data and Stream Processing at Scale. O'Reilly, 2022.
[11] https://www.mulesoft.com/
[12] https://www.ibm.com/docs/en/cloud-private/3.1.2?topic=services-datapower-gateway
[13] 15:47, 13 July 2023
https://en.wikipedia.org/w/index.php?title=Semantic_layer&oldid=1165191065
[14] https://powerbi.microsoft.com/en-us/
[15] 12:39, 21 August 2023
https://en.wikipedia.org/w/index.php?title=Data_science&oldid=1171493672
[16] Dhoni, Pan. Unleashing the Potential: Overcoming Hurdles and Embracing Generative AI in It Workplaces: Advantages, Guidelines, and Policies, 2023, https://doi.org/10.36227/techrxiv.23696709.v1.
[17] Koul, Anirudh, "Chapter 1: Exploring the Landscape of Artificial Intelligence, What Is AI," Practical Deep Learning for Cloud, Mobile, and Edge: Real-World AI; Computer-Vision Projects Using Python, Keras; Tensorflow, O'Reilly Media, Sebastopol, CA, 2019, pp. 6–7.
[18] https://en.wikipedia.org/, 5 June 2023, "Dartmouth workshop,"
https://en.wikipedia.org/w/index.php?title=Dartmouth_workshop&oldid=1158728511
[19] https://en.wikipedia.org/, Jan 22, 2023, "Logic Theorist,"
https://en.wikipedia.org/w/index.php?title=Logic_Theorist&oldid=1135115007
[20] Shortliffe, Edward, "Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project," Rule-Based Expert Systems: MYCIN, www.shortliffe.net/Buchanan-Shortliffe-1984/MYCIN%20Book.htm. Accessed 4 July 2023.
[21] Géron, Aurélien, "Neural Networks and Deep Learning," O'Reilly Online Learning, learning.oreilly.com/library/view/neural-networks-and/9781492037354/ch01.html#idm139624972225600, Accessed 4 July 2023.
[22] "What Is Generative Ai?" NVIDIA, www.nvidia.com/en-us/glossary/data-science/generative-ai/, Accessed 26 June 2023.
[23] Bell, Elysse, "Generative AI: How It Works, History, and Pros and Cons," Investopedia, 7 June 2023, www.investopedia.com/generative-ai-7497939#:~:text=Generative%20AI%20can%20benefit%20just,future%20AI%20models%20can%20train.
[24] Dhoni, Pan. A Cost-Effective It Approach to Rapidly Build a Data Platform and Integrate Retail Applications for Small and Mid-Size Companies., 2023, https://doi.org/10.36227/techrxiv.23538447.
[25] Dhoni, Pan, and Ravinder Kumar. Synergizing Generative AI and Cybersecurity: Roles of Generative AI Entities, Companies, Agencies, and Government in Enhancing Cybersecurity, 2023, https://doi.org/10.36227/techrxiv.23968809.v1.
[26] Courtois, Stephanie. "Generative AI: Taming Data Pipeline Sprawl." SnapLogic, 10 May 2023, www.snaplogic.com/blog/generative-ai-taming-data-pipeline-sprawl.

Pan Singh Dhoni: 1211 Denton Drive, Chester Spring, PA, 19425, USA
Masters in science (M. Sc) Mathematics 2000, Kumaun University, Nainital, Uttrakhand, India
Master Diploma (M. Tech) in Computer Science 2008, National Institute of Electronics, and Information Technology (NIELIT), Delhi, India
Technical Manager, Data Platform, Data Science & Analytics at Five Below Inc.
Close to two-decade experience in IT and computer related field.
Email: ps.dhoni@gmail.com
Work: pan.dhoni@fivebelow.com