# Integration of Text Analytics to Artificial Intelligence and Big Data for Data Processing

Manakshi Devi[1], Mr. Sourav Pindal[2]

*Department of Computer Applications, Chandigarh School of Business, Chandigarh Group of Colleges, Jhanjeri, Mohali, India*

[1]mishumishu561@gmail.com

[2]Sourav.j2085@cgc.ac.in

***Abstract: -*** **The continuously expanding textual data volume - from submissions on social media to reviews of customers - shows the need for an alternative approach to data processing that was used previously. Text Analytics, which is a branch of AI with roots in the NLP field, deserves a mention as it eliminates this gap. Such augmentation beyond traditional boundaries helps AI in revealing what was hidden from us before. The application of techniques like sentiment outputting and entity recognition makes the unstructured text into structured data which fosters AI functional intelligence. This paper is devoted to the interplay between text analytics, AI, and big data, which will help to solve difficult problems in different areas because in the future big data and AI will be basic analytical tools.**

***Keywords:*** **Artificial Intelligence, Big Data, Text Analytics, NLP, Text mining, sentiment analysis**.

## I. INTRODUCTION

The technique of employing computer systems to analyze and comprehend text authored by humans to extract business insights is known as text analysis. Text analysis software is capable of autonomously classifying, sorting, and extracting data from text to find correlations, trends, attitudes, and other useful information. Text analysis may be used to interpret many text-based sources, including emails, papers, social media posts, and product evaluations, in an accurate and fast manner, just like a human would analyze the attitude of blogs, forums, reviews, and other online media to find out if your consumers are satisfied with their purchases. With sentiment analysis, you may address PR concerns, identify emerging trends, and monitor shifts in sentiment. Through sentiment analysis and phrase identification, you may monitor shifts in consumer sentiment and pinpoint the issue's underlying cause.

### *A. Document administration*

Document management, classification, and searches become more effective with text analysis. This involves identifying insurance fraud, keeping an eye on brand references, and automating the administration of patient records. Text extraction, for instance, is used by LexisNexis Legal & Professional to locate certain entries among 200 million documents.

The most current moniker for Natural Language Understanding, Data, and Text Mining is Text Analytics. Big Data is a new term that has gained popularity in the last few years to refer primarily to unstructured text (or other information sources), more frequently in the commercial than in the academic domain. This is likely because unstructured free text, which includes tweets, blogs, wikis, and surveys, accounts for 90% of text in a business context.[1]

Prior methods in NLP research focused mostly on manually created rules or features, however, during data processing, they overlooked a lot of information. With the development of deep learning, more information can be gleaned from vast amounts of unprocessed data [2]. Natural language processing (NLP) techniques form the cornerstone of this integration, enabling the extraction of meaningful insights from unstructured textual data. Researchers have extensively investigated various NLP algorithms, such as sentiment analysis, named entity recognition, and topic modeling, to analyze text data efficiently.

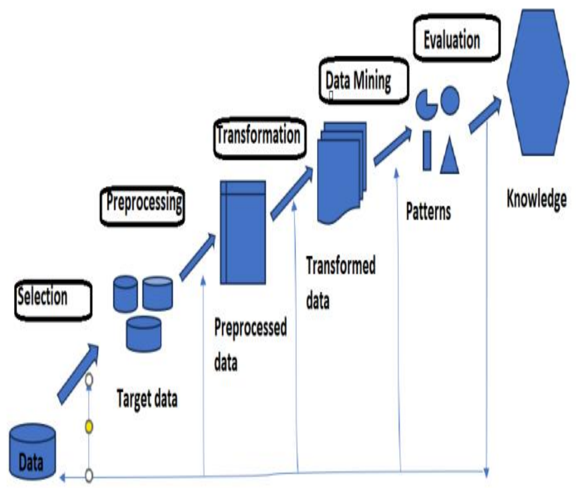


Fig. 1 Flowchart: from data to knowledge [2].

In the Big Data era, precise and scalable data-gathering approaches are imperative. For this reason, we are doing a thorough literature review of data collection from a data management perspective. For the most part, there are three ways to get data. First, datasets can be found, enhanced, or created using data capture techniques if the objective is to share and explore fresh datasets. Second, the individual instances may be labeled using a variety of data labeling approaches once the datasets are available. Lastly, it could be preferable to enhance already-existing data or train on top of previously learned models rather than labeling fresh datasets. These three approaches can be combined and are not inherently different from one another. For instance, one may while enhancing current datasets, search and label more ones [3]. The PLS regression approach served as the foundation for assessing the research hypotheses in the conceptual model. This approach benefited our study in several ways. It is more suited for exploratory research, which is pertinent to our investigation into a novel BDA technology, on the one hand. Hence, the PLS approach, which is less affected by model specification mistakes, might present an intriguing substitute to SEM (Smart PLS) for the estimation of a more generic model based on covariance [4].

With the advent of big data resources and the analysis of large amounts of financial data, text mining has become increasingly popular. Financial institutions may then extract useful data from social media postings, emails, phone logs, business papers, consumer comments, and more. They can also use this data to detect fraud or other hazards, client attrition, and fraud. This study seeks to address three research issues to illustrate the many facets of textual mining's application in banking and finance [5]:

1) What constitutes the field's intellectual center?
2) In the era of the Internet, big data, and social media, which text mining approaches are employed in the financial industry for textual mining?
3) Which data sources are most frequently utilized in the financial industry for text mining, and why?

The purpose of these cutting-edge analytical methods is to mine large datasets for information, unidentified relationships, and hidden patterns [6]. For example, a thorough examination of past patient data may help identify harmful diseases early on, opening the door to a cure or more effective course of therapy.

### *B. Definition of big data and its issues*

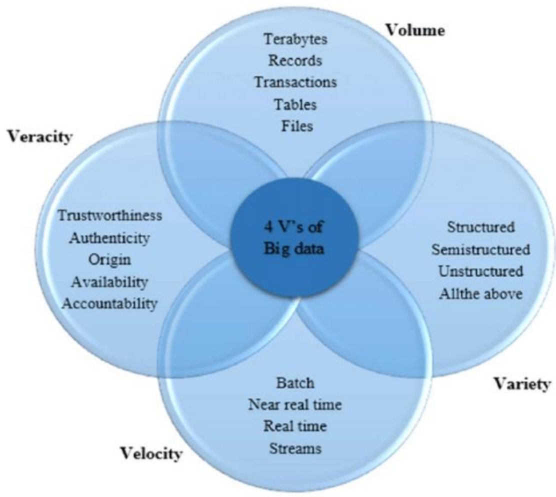


Fig. 2 Big data aspects [7].

Four main components make up big data [7]:
Variety uses many data formats and sources, such as written accounts of ongoing events, traffic datasets about traffic on public roadways, and scheduled events (such as athletic or musical performances).
Velocity: Data input and output speeds are displayed via velocity. It speaks about the dynamic nature of data, how frequently data is generated, and how important it is to produce results instantly.
Volume is concerned with the amount of data that can be stored; this can be megabytes, gigabytes, terabytes, or even petabytes.
Veracity is the degree to which the data can be relied upon, given the dependability of its source; for example, certain devices may be compromised when receiving data from sensors.

### C. Using healthcare as a big-data warehouse

The prevention, diagnosis, and treatment of health-related problems or impairments in humans is the only goal of the multifaceted healthcare system.
Healthcare facilities (clinics, hospitals for the delivery of medications and other diagnostic or treatment technologies), health professionals (physicians or nurses), and a financing institution supporting the former two are the main components of a healthcare system. These healthcare professionals work in a variety of fields, including medicine, dentistry, nursing, midwifery, physiotherapy, psychiatry, and many more. Depending on how urgent the problem is, different degrees of healthcare are needed [8].

### D. Health records in electronic form

EHRs have brought forth several benefits for managing data connected to modern healthcare. We go over a few of the main benefits of using EHRs below. The first benefit of electronic health records (EHRs) is enhanced access to a patient's complete medical history for medical providers. Medical diagnoses, medications, information about known allergies, demographics, clinical narratives, and the outcomes of numerous laboratory tests are all included in the data. Due to a decrease in the time lag of prior test findings, medical problem detection and treatment are now timelier [8].

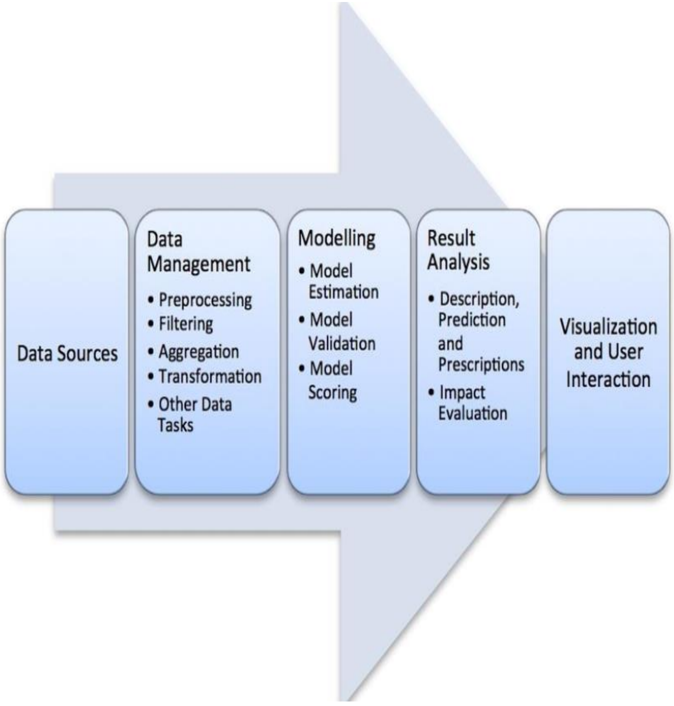


Fig. 3 Workflow of Big Data Analytics [8].

### E. Creation of a Big Data-Based Artificial Intelligence Education System

At both the content and function levels, the artificial intelligence education system is very goal-oriented. Thus, the system as a whole must have strong stability, a quick reaction time, consistent and dependable user data, easy system operation, and an intuitive user interface and primary functionalities [10]. The part on system needs examines and arranges the various demands of many disciplines. Eight components comprise the system's primary functions, categorized based on the prior demand study. They include data reporting, interactive learning, interactive exercises, authority management, course management, knowledge point management, and subject management [10].

#### 1) Techniques for sentiment analysis

Data retrieval, data extraction and selection, data pre-processing, feature extraction, topic detection, e) and f) are the steps involved in the multi-step process of sentiment analysis.
Finding and defining the data source—such as a social media platform or the portal of a commercial service provider—is necessary for data retrieval. A particular web crawling method is required to retrieve the review data from these sources, store it in a database, and take into account the data format [34].

The following gives these five categories a thorough explanation:

*2) Method of supervised machine learning*

A supervised machine learning-based approach to sentiment analysis entails building a model from annotated data or poorly labeled corpora. "What a wonderful holiday!!!" is labeled as a sentence with "positive" sentiment polarity in the manual annotation process. Data with weak labels are those for which the machine used a heuristic to derive the class labels. For instance, user-generated material on review systems frequently includes data with poor labels when reviewers provide their review categories and ratings (like stars).

*3) Unsupervised Machine Learning Methodology*

Image analysis, pattern recognition, and data mining have all made use of cluster analysis as an unsupervised machine-learning technique. The process of organizing a collection of data into clusters so that the elements within a cluster are more alike than the elements outside of it is known as clustering. In the literature, clustering methods for sentiment analysis of brief text data were used, including k-means (Xiang et al., 2015b) and statistical models based on the probability distribution of reviews in sentiment space (Rossetti et al., 2015). Furthermore, unsupervised versions of Naïve Bayes models were also modified for sentiment analysis.

*4) Dictionary-based methodology*

This review also utilizes the words as synonyms because dictionary-, lexicon-, and rule-based techniques were used interchangeably in the literature. Figure 2 depicts the entire architecture of a typical dictionary/rule-based sentiment analysis approach to give an overview of dictionary-based techniques. This method allows for the identification of subjectivity vs objectivity to be either handled by the sentiment polarity detection process itself or may be integrated into the framework. Depending on the particular requirements of the application, aspect or topic detection may also be incorporated into the architecture. Dictionary-based systems rely on sets of optimized rules and extensive sentiment lexicons. A sentiment dictionary can be generated semi-automatically by humans, automatically by machines, or by both.

*5) Semantic methodology*

Semantic-based analysis was introduced to improve the dictionary-based technique. The semantic method primarily uses a language model based on rules to determine the polarity of each text fragment. This method requires the use of a dictionary with words unique to the domain and the polarity values that go along with them.

*6) Hybrid methodology*

In hybrid systems, two sentiment polarities can be computed simultaneously using machine learning and dictionary-based methods. To produce a final sentiment polarity, the outcomes from machine learning-based techniques and the vocabulary are integrated. Additionally, a sentiment analysis model may be created by combining machine learning and dictionary-based techniques at various phases of the model's development.

## II. LITERATURE REVIEW

A. The document entitled Text Mining for Big Data Analysis in Financial Sector by authors Mirjana Peji´c Bach, *, Živko Krsti´c, Sanja Seljan, and Lejla Turulja defines that The report emphasizes how big data technologies are becoming more and more important in the financial industry, especially when it comes to using textual data to inform decision-making. The results provide insight into how the field of study is developing and emphasize the need for more investigation and use of text-mining methods to fully utilize the abundance of data found in unstructured data sources [5].

B. The document entitled Text Analytics: The Convergence of Big Data and Artificial Intelligence by authors Antonio Moreno1, and Teófilo Redondo2 defines that text analytics holds immense potential, driven by the exponential growth of textual data and the demand for extracting actionable insights. As communication continues to expand, text analytics will play an indispensable role in navigating and harnessing the wealth of information available, contributing to further advancements in AI and cognitive applications [1].

C. The document entitled Challenges and Opportunities: From Big Data to Knowledge in AI 2.0 by authors Yue-ting ZHUANG, Fei WU‡, Chun CHEN, and Yun-he PAN the paper underscores the transformative potential of AI 2.0 and the importance of integrating human knowledge with data-driven approaches to advance AI capabilities and applications [2].

D. The paper titled Concepts, Opportunities, and Challenges of Re-Thinking Data Strategy and Integration for Artificial Intelligence Authors Abdulaziz Aldoseri, Khalifa N. Al-Khalifa, and Abdel Magid Hamouda * state that this work offers a thorough analysis of the difficulties involved in using data for AI applications. It highlights how crucial it is to deal with these issues to properly utilize AI's potential across a range of businesses. The following crucial aspects of the difficulties are emphasized. Data quality: guaranteeing timely, accurate, comprehensive, and consistent data collection, pre-processing, administration, and governance Data Volume: Managing the enormous volumes of data produced, including heterogeneity and access challenges, storage, processing, administration, privacy, and security. Data Security and Privacy: Preserving data from Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges, Abdulaziz Aldoseri, Khalifa N. Al-Khalifa, and Abdel Magid Hamouda *, the authors outline how to ensure compliance with privacy regulations through techniques like differential privacy and robustness training, as well as inference attacks and adversarial attacks. Bias and Fairness: To guarantee fairness and representation in AI systems, biases in data collecting, labeling, sampling, and aggregation must be addressed [12].

E. The document entitled Sentiment analysis in tourism: Capitalising on Big Data by authors Alireza Alaei1, 2, Susanne Becken2, and Bela Stantic1 defines that the way information is made and consumed has changed dramatically as a result of technology's revolutionary effects on tourism. Online tourism content has become increasingly influential as travelers use a wider range of information sources and actively participate in social media content creation. However, the sheer amount of data available online has rendered human processing impracticable, calling for the creation of automated analytical techniques.
Sentiment analysis is one such technique that is becoming more and more popular. It automates the investigation of semantic linkages and meaning in reviews. In the context of tourism, this work has studied and evaluated many sentiment analysis techniques, taking into account the datasets employed and performance on important assessment criteria. Notwithstanding progress, problems still exist, especially in aspect-oriented sentiment analysis. Finding implicit elements in reviews is still quite difficult [34].

F. The document entitled Big Data in Healthcare: management, analysis, and Prospects by authors Sabyasachi Dash1†, Sushil Kumar Shakyawar2,3†, Mohit Sharma4,5, and Sandeep Kaushik6* defines the emergence of big data has revolutionized the healthcare industry, offering immense potential to improve services and drive financial advantages. Over the past two decades, big data has garnered significant interest due to its vast possibilities. In healthcare, various sources such as hospital records, medical examinations, and biomedical research contribute to the generation of substantial datasets.
However, effectively managing and analyzing this data presents numerous challenges, akin to finding a needle in a haystack without proper tools and techniques. High-end computing solutions are essential for navigating these challenges and deriving meaningful insights from big data. To fully harness the potential of big data in healthcare, organizations must invest in appropriate infrastructure for systematic data generation and analysis.
Integration of biomedical and healthcare data holds the promise of revolutionizing medical therapies and personalized medicine. With tools such as genomics, mobile biometric sensors, and smartphone apps, vast amounts of data are generated, necessitating thorough assessment and analysis to unlock their potential [8].

## III. SOME USE CASES

Text analytics has led to the development of practical, daily-use apps in healthcare and industry too.
We're merely showcasing a few of these here:

### A. *Applications in Healthcare*:

1) *Better Diagnosis and Treatment:* Help Unfold the Fundamental Textual Analytics. The branch of AI, called text analysis, which utilizes NLP (Natural Language Processing), is dragging healthcare to a new era by teasing out patterns and hidden trends among the long texts. This data encompasses:

2) *Patient Medical Records:* Electronic Health Records (or EHRs for short) have a lot of data

in them - for instance, doctors' notes, diagnoses, and medications, along with lab results ad other test results. Text analytics can dig up these logs to detect slight links comprised of symptoms, medications, and disease progression stages.

3) *Clinical Research Articles:* Clinical research forwards the frontiers of science by regularly unveiling discoveries. Text analytics helps through its ability to sift through huge lengths of literature and come up with summaries of key points, identify any irreconcilable drug interactions, and pinpoint alternative treatment possibilities for certain patient profiles.

4) *Clinical Trial Data:* Text analytics is capable of breaking through the barriers related to the volume and chaos of unstructured text among clinical trials, exemplified through patient narratives and the observations of physicians. Unexpectedly, this will allow investigation the new unknown, differentiate benefactors of the drugs within the demographic groups and hasten the discovery of the life-prolonging drugs.

B. *Risk Management and Patient Care*

Preventive Measures of Text Analytics. In predictive risk management and active patient care, text analytics has a critical role. Through fostering analytics of a tremendous volume of textual data (EMRs and discharge summaries) it helps healthcare professionals to precisely find patients who are at a high risk of certain diseases. This brings the crisis to the attention of the authorities early and the implementation of preventive actions, which leads to better patient results and reduced healthcare costs.

1) *Pattern Recognition:* Text analytics can reveal trends in medical records such as patient history, demographics, laboratory outcomes, and prescription drugs present in EMRs. These patterns may reveal people with more chances of developing specific diseases such as heart disease, diabetes, or even some cancers.

2) *Predictive Modeling:* Sophisticated text analysis methods can be employed to segment patient data and build machine-learning models by using historical data to predict the likelihood of an individual contracting a certain disease in the future. This way enables healthcare providers to focus on pre-health care for patients with high-risk factors.

3) *Social Determinants of Health Integration:* The analysis of the text can outline the social factors that are mentioned in the medical records and the patients ' situation, like housing, income, and access to healthy food. These parameters substantially influence health outcomes, and we can identify at-risk patients as per eh social determinants to devise specific interventions for addressing the root causes of the problem.

C. *Fraud Detection and Claims Processing*

Since text analytics is instrumental in understanding the customers, it is crucial to have a platform that collects, analyses, and visually interprets consumer feedback. Analysis of text has a significant part in executing the fight against fraud and in optimizing claims processing in the health insurance industry. The processing of the large number of textual findings in medical records and insurance claims, therefore provides insurers with the tools necessary to identify fraud outright while at the same time speeding up otherwise legitimate claims. Here's a deeper dive into its impact: Here's a deeper dive into its impact:

*D. Enhanced Fraud Detection*

1) *Identifying Red Flags:* Using text analytics, medical records, and claims data can be scrutinized for an integrity check to catch the gaming of these systems which often results in the detection of fraudulent activities. Sorting out these factors involves studying the diagnosis, treatment procedures, prescribed drugs that appear to be common, and billing practices.

2) *Predictive Modeling:* A breakthrough in complex techniques allows machine learning to formulate models based on prior seats of fraudulent claims. Thus, these systems can evaluate naturally arriving cases and flag the ones with the most probability of being fraudulent, which would enable a deeper investigation.

3) *Network Analysis:* Through tracking the complex relationships in various reports, text analytics may show that certain claims share a deep correlation without the slightest mention of them being related. This comes in very handy in cases of complex fraud rings; when various parties cooperate in sending fake claims to the insurers.

E. *Applications in the Industry:* Disclosing Customer Insight through Text Analytics

1) *Customer Support and Feedback Analysis:* Text analytics is a key factor nowadays in the industry, helping many companies make insightful decisions through a better understanding of customers undefined.

i. *Sentiment Analysis:* Text analytics can extract information from emails, social media threads, and online reviews to investigate users' attitudes. With this service, the business can distinguish between happy and sad customers by knowing the reasons for them to experience either positive or negative situations, and hence focus on ways to make improvements.

ii. *Identifying Recurring Issues:* Through the examination of the huge quantities of customer reviews, text analytics can find out the problems that often occur as well as the areas of discomfort. The result is this: companies can anticipate these issues in advance and take countermeasures against future client irritation.

iii. *Improving Customer Support:* Through text analytics, you can categorize and prioritize customer inquiries related to their urgency and topic. This makes us able them to respond in a more efficient and effective, in this way enhancing the overall customer experience.

iv. *Extracting Consumer Preferences:* Text analytics have the capacity not only to analyze enormous amounts of data but also news articles, social media conversations, and market studies. This way companies are allowed to surface nuanced trends in consumer tastes, reveal new emerging needs, and acknowledge the competitive environment.

v. *Social Listening*: Through the tracking of social media talk, text analytics can sense public sentiment by disclosing what people say about a brand, its products, and its competitors. This social listening empowers businesses to remain on top during emerging trends, solve issues regarding brand reputation, and customize their marketing approaches as well.

vi. *Product Development and Innovation:* Identifying Feature Requests: Text analytics allows for analysis of customer reviews, social media posts, and product forums to detect the most often desired features and suggestions for implementation. It enables them to concentrate their new product introduction efforts on those that customers consider worthwhile and relevant.

## IV. CONCLUSION

More potent text analytics tools will result from developments in artificial intelligence (AI) and natural language processing (NLP).
Further in-depth understanding from several data sources will be possible with integration with big data.
There will be a growing specialization of text analytics for the financial and healthcare sectors.
The following factors are also examined in the paper:
Various methods are applied in sentiment analysis, a popular text analytics application.
The role of big data in text analytics and how it's affecting different industries.
Real-world applications of text analytics in the medical field and business.
Overall, the study emphasizes how text analytics is becoming more and more significant in the big data era and how it has the potential to transform several industries.

With its lengthy and distinguished history, text analytics is a field that is always changing. It is located at the core of the Variety vector of Big Data, which is unstructured information. This is particularly true with social communications, where the material is created by millions of people and frequently consists of written comments or entire articles in addition to photographs. Textual expressions of information include a wealth of knowledge about the world, its entities, and the relationships between them. Global knowledge has already been put to use in the development of cognitive apps, such as iSOFT's Amelia and IBM's Watson, which will interact with people to enhance their capacities and improve their performance.

## V. FUTURE SCOPE

The practice of examining unstructured textual data to glean insightful information is known as text analytics.
Effective analytic approaches are required because of the exponential rise of textual data, including documents, emails, and social media.
There is enormous potential for text analytics across several industries, including customer service, banking, and healthcare.
Utilizations in the Medical Field
Improved Diagnosis and Treatment: Patterns and trends may be found by analyzing large volumes of patient data, research articles, and medical records. This can help with individualized treatment plans, more precise diagnoses, and improved patient outcome prediction.
Risk management and patient care: Through the analysis of medical data, text analytics may identify

patients who are at high risk for particular diseases, facilitating early intervention and preventative measures.

The identification and processing of fraudulent claims can be aided by the examination of medical data and insurance claims.

Currently, several sectors are using text analytics tools. For example, sentiment and opinion analysis in media, banking, medical, branding, or consumer markets.

Since the general population is becoming the greatest producer of text content (just think of online messaging services like WhatsApp or Telegram), insights are derived not just from traditional corporate data sources but also from online and social media.

Although text analytics is doing quite well right now, there is still an opportunity for improvement in areas like social listening and customer experience.

This has great potential for both technological innovation and scientific experimentation: Machine learning (ML) and developments in machine translation, along with customer experience, market research, and other factors, enable multilingual analytics.

## REFERENCES


[1] Moreno, A., & Redondo, T. (2016). *Text analytics: the convergence of big data and artificial intelligence*. *IJIMAI*, *3*(6), 57-64.

[2] Zhuang, Y. T., Wu, F., Chen, C., & Pan, Y. H. (2017). *Challenges and opportunities: from big data to knowledge in AI 2.0. Frontiers of Information Technology & Electronic Engineering*, *18*, 3-14.

[3] Roh, Y., Heo, G., & Whang, S. E. (2019). *A survey on data collection for machine learning: a big data-ai integration perspective. IEEE Transactions on Knowledge and Data Engineering*, *33*(4), 1328-1347.

[4] Benzidia, S., Makaoui, N., & Bentahar, O. (2021). *The impact of big data analytics and artificial intelligence on green supply chain process integration and hospital environmental performance. Technological forecasting and social change*, *165*, 120557.

[5] Pejić Bach, M., Krstić, Ž., Seljan, S., & Turulja, L. (2019). *Text mining for big data analysis in financial sector: A literature review. Sustainability*, *11*(5), 1277.

[6] Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). *Uncertainty in big data analytics: survey, opportunities, and challenges*. *Journal of Big Data*, *6*(1), 1-16.

[7] Mohamed, A., Najafabadi, M. K., Wah, Y. B., Zaman, E. A. K., & Maskat, R. (2020). *The state of the art and taxonomy of big data analytics: view from new big data framework. Artificial Intelligence Review*, *53*, 989-1037.

[8] Fillinger, S., de la Garza, L., Peltzer, A., Kohlbacher, O., & Nahnsen, S. (2019). *Challenges of big data integration in the life sciences. Analytical and bioanalytical chemistry*, *411*, 6791-6800.

[9] Chung, Y., Kraska, T., Polyzotis, N., Tae, K. H., & Whang, S. E. (2019). *Automated data slicing for model validation: A big data-ai integration approach. IEEE Transactions on Knowledge and Data Engineering*, *32*(12), 2284-2296.

[10] Xu, B. (2021). *Artificial intelligence teaching system and data processing method based on big data. Complexity*, *2021*, 1-11.

[11] Costa, J. P., Grobelnik, M., Fuart, F., Stopar, L., Epelde, G., Fischaber, S., ... & Davis, P. (2020). *Meaningful big data integration for a global COVID-19 strategy. IEEE Computational Intelligence Magazine*, *15*(4), 51-61.

[12] Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). *A survey of machine learning for big data processing. EURASIP Journal on Advances in Signal Processing*, *2016*, 1-16.

[13] Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). *Big data in healthcare: management, analysis, and prospects*. *Journal of big data*, *6*(1), 1-25.

[14] Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). *Artificial intelligence for decision making in the era of Big Data–evolution, challenges, and research agenda*. *International journal of information management*, *48*, 63-71.

[15] Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). *Artificial intelligence for decision making in the era of Big Data–evolution, challenges, and research agenda. International journal of information management*, *48*, 63-71.

[16] Munawar, H. S., Qayyum, S., Ullah, F., & Sepasgozar, S. (2020). *Big data and its applications in smart real estate and the disaster management life cycle:* A systematic analysis. *Big Data and Cognitive Computing*, *4*(2), 4.

[17] Munawar, H. S., Qayyum, S., Ullah, F., & Sepasgozar, S. (2020). *Big data and its applications in smart real estate and the disaster management life cycle:* A systematic analysis. *Big Data and Cognitive Computing*, *4*(2), 4.

[18] Jagatheesaperumal, S. K., Rahouti, M., Ahmad, K., Al-Fuqaha, A., & Guizani, M. (2021). *The duo of artificial intelligence and big data for industry 4.0: Applications, techniques, challenges, and future research directions. IEEE Internet of Things Journal*, *9*(15), 12861-12885.

[19] Jimenez-Marquez, J. L., Gonzalez-Carrasco, I., Lopez-Cuadrado, J. L., & Ruiz-Mezcua, B. (2019). *Towards a big data framework for analyzing social media content. International Journal of Information Management*, *44*, 1-12.

[20] Elgendy, N., & Elragal, A. (2016). *Big data analytics in support of the decision-making process. Procedia Computer Science*, *100*, 1071-1084.

[21] Yoo, H., & Chung, K. (2020). *Deep learning-based evolutionary recommendation model for heterogeneous big data integration. KSII Transactions on Internet and Information Systems (TIIS)*, *14*(9), 3730-3744.

[22] Wang, J., Zheng, P., Lv, Y., Bao, J., & Zhang, J. (2019). Fog-IBDIS: *Industrial big data integration and sharing with fog computing for manufacturing systems*. *Engineering*, *5*(4), 662-670.

[23] Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. (2020). *IoT, big data, and artificial intelligence in the agriculture and food industry. IEEE Internet of Things Journal*, *9*(9), 6305-6324.

[24] Ahmadi, S. (2024). *A Comprehensive Study on the Integration of Big Data and AI in the Financial Industry and its Effect on Present and Future Opportunities*. *International Journal of Current Science Research and Review*, *7*(01), 66-74.

[25] Xu, Y., Zhang, X., Li, H., Zheng, H., Zhang, J., Olsen, M. S., ... & Qian, Q. (2022). *Smart breeding is driven by big data, artificial intelligence, and integrated genomic-environs prediction. Molecular Plant*, *15*(11), 1664-1695.

[26] Chen, H., Chiang, R. H., & Storey, V. C. (2012). *Business intelligence and analytics: From big data to big impact*. *MIS Quarterly*, 1165-1188.

[27] Bag, S., Gupta, S., Kumar, A., & Sivarajah, U. (2021). *An integrated artificial intelligence framework for knowledge creation and B2B marketing rational decision-making for improving firm performance*. *Industrial marketing management*, *92*, 178-189.

[28] Vassakis, K., Petrakis, E., & Kopanakis, I. (2018*). Big data analytics: applications, prospects, and challenges. Mobile big data: A roadmap from models to technologies*, 3-20.

[29] Farzindar, A., Inkpen, D., & Hirst, G. (2015). *Natural language processing for social media*. San Rafael: Morgan & Claypool.

[30] Lin, X., Zhang, Y., & Wang, J. (2022). *Digital library information integration system based on big data and deep learning*. *Journal of Sensors*, *2022*.

[31] Gandomi, A., & Haider, M. (2015). Beyond the hype*: Big data concepts, methods, and analytics. International journal of information management*, *35*(2), 137-144.

[32] Alsunaidi, S. J., Almuhaideb, A. M., Ibrahim, N. M., Shaikh, F. S., Alqudaihi, K. S., Alhaidari, F. A., ... & Alshahrani, M. S. (2021). *Applications of big data analytics to control the COVID-19 pandemic*. *Sensors*, *21*(7), 2282.

[33] Demirkan, H., & Delen, D. (2013). *Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in the cloud*. *Decision Support Systems*, *55*(1), 412-421.

[34] Alaei, A. R., Becken, S., & Stantic, B. (2019). *Sentiment analysis in tourism: capitalizing on big data. Journal of travel research*, *58*(2), 175-191.

[35] Sagiroglu, S., & Sinanc, D. (2013, May). *Big data: A review*. In *2013 international conference on collaboration technologies and systems (CTS)* (pp. 42-47). IEEE.

[36] Karanam, S. D., Kamath, R. S., Kulkarni, R. V. R., & Pai, B. H. S. K. (2021). Big data integration solutions in organizations: A domain-specific analysis. *Data Integrity and Quality*, 1-31.