Journal of Economics, Business, and Commerce (JEBC)
ISSN: 3007-9705 (Online)
Volume 2 Issue 2, (2025)
https://doi.org/10.69739/jebc.v2i2.1000
https://journals.stecab.com/jebc

Published by
Stecab Publishing

Research Article

Dark Data in Business Intelligence: A Systematic Review of Challenges, Opportunities, and Value Creation Potential

*[1]Ololade Funke Olaitan, [2]Temitope Anthony Adebanjo, [3]Loveth Itohan Obozokhai, [4]Ambrose Nwawuweneonye Iwerumoh, [3]Isaac Oluwaseyi Balogun, [5]Damilola Ayodele Ojo

About Article

Article History
Submission: September 01, 2025
Acceptance : October 06, 2025
Publication : October 10, 2025

Keywords
Artificial Intelligence (AI), Business Intelligence (BI), Dark Data, Data Governance, Ethical Data Management, Unstructured Data Analytics

About Author
[1] David Eccles School of Business, information Systems, University of Utah, USA

[2] Department of Marketing and Analytics, Regenesys Education, Johannesburg, South Africa

[3] Robinson College of Business, Georgia State University, Atlanta, Georgia, USA

[4] Department of Quality Management, Rushford Business School, Lucerne, Switzerland

[5] College of Business, Missouri State University, Springfield, Missouri, USA

Contact @ Ololade Funke Olaitan
ololadefunke@gmail.com

ABSTRACT

An increasingly important yet underexplored aspect of Business Intelligence (BI) is dark data the massive volume of data acquired by firms but left unanalyzed. Hidden assets include textual documents, IoT logs, photos, audio records, and hybrid datasets, which often constitute the majority of company data. This review examines the significance of dark data, the barriers businesses face in unlocking its value, and the technological advancements that are gradually revealing its potential. The paper highlights how organizations can use dark data to improve decision-making, enhance operational efficiency, and gain a competitive edge. Key challenges such as poor data quality, integration difficulties, and high processing costs currently hinder the systematic use of dark data. However, advancements in natural language processing, machine learning, computer vision, knowledge graphs, and synthetic data production offer promising avenues for overcoming these obstacles. Case studies in manufacturing, banking, and retail demonstrate how dark data can drive predictive maintenance, fraud detection, and personalized consumer engagement. The review also explores the ethical and legal implications surrounding the use of dark data, particularly in relation to privacy, bias, and regulatory compliance. The paper emphasizes the strategic imperative for organizations to adopt a proactive, ethically informed approach to dark data, integrating advanced technologies and robust governance frameworks to transform hidden information into actionable insights that drive sustainable business value.

Citation Style:
Olaitan, O. F., Adebanjo, T. A., Obozokhai, L. I., Iwerumoh, A. N., Balogun, I. O., & Ojo, D. A. (2025). Dark Data in Business Intelligence: A Systematic Review of Challenges, Opportunities, and Value Creation Potential. *Journal of Economics, Business, and Commerce, 2*(2), 135-142. https://doi.org/10.69739/jebc.v2i2.1000

1. INTRODUCTION

1.1. Contextualizing dark data

The digital revolution has brought about a time when data is created at an unprecedented rate. However, only a small amount of this data is systematically collected, evaluated, and used to make decisions. A lot of business data, which is often called "dark data," is still concealed in the systems, unstructured archives, and daily activities of the business (Bresciani *et al.*, 2021). Dark data is the information that is acquired, processed, and kept as part of normal corporate operations but is never looked at for strategic insights. According to industry assessments, about 80–90% of enterprise data is "dark," meaning it is not being used and is instead stored in log files, emails, multimedia records, sensor outputs, and archived documents (George *et al.*, 2023).

In the past, organizations primarily concentrated on structured data that was stored in relational databases. This made it simple for Business Intelligence (BI) systems to identify patterns and trends. However, the rapid proliferation of unstructured and semi-structured media, such as text, audio, and video, has altered the focus of data analytics. Traditional BI systems are becoming increasingly potent and are now utilizing artificial intelligence (AI) and natural language processing (NLP) to manage the vast quantity of complex, underutilized information (Sharma *et al.*, 2021). This alteration demonstrates the significance of identifying value in black data to support initiatives that are founded on evidence.

1.2. Research rationale

In today's competitive world, businesses see data as a valuable resource, much like money and people. However, ignoring dark data creates a paradox: even while the amount of data is growing at an exponential rate, the amount of useful information that can be derived from it is still quite modest. Because of this gap, important information about customer behavior, operational efficiency, risk management, and innovation is often missed (Shahid & Sheikh, 2021).

The study rationale is predicated on rectifying this imbalance. The notion of big data has garnered significant attention; yet, the concealed and underutilized aspect dark data remains relatively unexamined. The absence of systematic methodologies for assessing dark data has created a substantial research void, especially concerning how firms might convert inactive knowledge into a competitive edge. It's important to understand this area since using dark data could not only boost company results, but it could also change the way data-driven initiatives are thought of in the first place.

1.3. Research Questions and Objectives

This review seeks to examine dark data by establishing three fundamental inquiries:

i. What is dark data and where does it come from in businesses?

ii. What makes it hard to include it in mainstream analytics and BI systems?

iii. How can companies turn dark data into useful information that gives them an edge over their competitors?

The main goal of the paper is to give a methodical look at the problems, chances, and future paths that dark data in Business Intelligence presents. The review aims to underscore the dangers of neglecting this resource and the possible benefits of its appropriate use by critically synthesizing the existing information.

2. LITERATURE REVIEW

2.1. Defining dark data in business intelligence

Dark data has become a key topic in BI and enterprise analytics debates. The phrase is widely used, but its meaning differs by sector. Dark data is normal data that firms acquire but don't examine for strategic or operational goals. Gartner, which popularized the phrase, underlines that dark data includes underused information assets with potential value. These include customer contacts, transactional by-products, and archival documents, which are maintained but rarely analyzed for business insights (Schembera & Durán, 2020).

Different industries view dark data through their operational reality. Healthcare black data includes unstructured clinical notes, diagnostic pictures, and patient histories archived but not integrated into decision-making procedures. Financial transaction records and recorded discussions may contain risk indicators or compliance red flags. However, retailers detect dark data in customer comments, browser histories, and abandoned shopping carts. These interpretations show that dark data's definition is context-dependent (Kong, 2019; Olaitan *et al.*, 2025).

The unstructured nature of dark data is crucial. Unstructured data includes music, video, photos, and free text lacks regularity and resists typical BI procedures. Extraction and analysis are difficult, therefore such data is underutilized despite its richness. Dark data intrigues scholars and practitioners because of this paradox having potentially useful resources but not using them (George *et al.*, 2023).

2.2. Sources of dark data

Dark data sources reflect modern organizations' heterogeneity. Several main categories show the scope of this phenomenon:

- Logs from IoT sensors: With so many connected devices, companies generate massive logs of performance, environmental variables, and user interactions. Although vital for diagnosis, much of this data is stored without analysis (Olaitan *et al.*, 2025).
- Call transcripts: Even though contact centers record calls for quality purposes, only a small portion is reviewed. These transcripts reveal customer emotion, reoccurring difficulties, and product feedback (Płaza *et al.*, 2022).
- Medical records: Electronic health records, physician notes, and diagnostic imaging are significant dark data sources in healthcare. They include clinical detail but are sensitive and unstructured, limiting their analytical value (Fagbenle, 2025).
- Key sources of customer feedback include emails, online reviews, and social media posts. Due to relevance filtering and volume management issues, these sources are often missed despite their importance for customer behavior (Lawal *et al.*, 2025).
- Enterprise documents: Reports, PDFs, and Word files are often housed in silos. Lack of indexing and searchability makes past trends and compliance data unavailable.

Data silos isolate information within departments or private systems, compounding these issues. Without interoperability, valuable datasets cannot be cross-referenced or pooled for insights. Silos increase inefficiencies and the "darkness" of corporate data assets.

2.3. Challenges in utilizing dark data

Despite the potential of dark data, organizations encounter difficulties in unlocking its value.

Operational expenses constitute an essential impediment. The storage, maintenance, and cybersecurity expenses associated with maintaining an abundance of dormant data are substantial. Unnecessary logs, photographs, and documents are a waste of money due to the capacity-based pricing of cloud services (George *et al.*, 2023).

The situation is further complicated by legal and compliance requirements. Data collection, storage, and processing are subject to limitations under the General Data Protection Regulation (GDPR) of the European Union. Consent is necessary for the mining of sensitive datasets, such as health records or consumer interactions. Consequently, exploration is frequently impeded by the legal repercussions of misconduct, which include penalties and reputational harm, despite the potential value of data (Hoofnagle *et al.*, 2019).

The most significant challenges may be technological. Unstructured data necessitates deep cleansing, normalization, and format conversion. Many organizations are unable to scale such initiatives due to a lack of infrastructure or expertise. Accessibility concerns are frequently exacerbated by the dispersion of dark data among aging systems, cloud environments, and offline repositories. The processing of complex dark data is technically and resource-intensive as a result of fragmentation (Sedlakova *et al.*, 2023).

2.4. Technological approaches for unlocking dark data

Companies can now brighten their dark data vaults with technology.

- Emails, customer comments, and call transcripts can be interpreted using NLP. Sentiment analysis, entity recognition, and topic modeling can turn unstructured narratives into actionable intelligence. Customer complaints may be automatically categorized, showing product or service issues (Shah *et al.*, 2023).

AI and machine learning algorithms excel in pattern recognition in huge, complex datasets. Model training helps organizations uncover connections and prediction signals in unstructured data. While mining dark data, AI can automate repetitive analytical tasks, saving time and money (Olaitan *et al.*, 2025).

- Computer Vision: Healthcare, security, and retail need visual data. Computer vision can detect objects, anomalies, and behaviors in pictures and videos. BI goes beyond words and numbers to visualizations.

Structure unstructured data into entities and relationships for semantic analysis with knowledge graphs. Connecting datasets helps organizations avoid silos and boost contextual knowledge. Email-transaction history links boost consumer behavior insight (Afzal *et al.*, 2023).

- Synthetic data privacy considerations limit access to sensitive datasets. Creating synthetic data that mimics real datasets is promising. A risk-free model training environment with synthetic data lets organizations experiment with analytics while complying with data protection requirements (Patel, 2024).

New technologies make black data easier to use, as shown by these technical techniques.

3. METHODOLOGY

3.1. Approach to the review

This study follows a narrative-integrative review approach, which combines qualitative and quantitative methods to provide a comprehensive understanding of dark data in Business Intelligence (BI). The primary focus is on synthesizing existing literature to identify key themes, challenges, and opportunities related to the utilization of dark data in BI systems.

3.2. Data sources and collection

A systematic literature search was conducted using academic databases such as Scopus, Web of Science, IEEE Xplore, ScienceDirect, and Google Scholar. In addition to peer-reviewed publications, authoritative industry reports from leading organizations such as Gartner, IBM, McKinsey, and Deloitte were included to balance academic and applied perspectives on dark data. The inclusion criteria focused on publications from 2010–2025, emphasizing empirical studies, case studies, and theoretical frameworks. Initial searches yielded approximately 300 records, which were refined using a two-stage screening process. After evaluating the relevance and methodological quality of the studies, 85 key articles were selected for final analysis.

3.3. Thematic analysis

Themes were derived using thematic analysis to identify recurring patterns and concepts in the selected literature. The process was structured as follows:

i. Data Familiarization: The full-text articles were read and re-read to ensure a deep understanding of the content.

ii. Initial Coding: Key concepts and findings were highlighted and coded manually, focusing on recurring terms such as "dark data," "challenges," "technological solutions," and "ethical concerns."

iii. Theme Identification: Codes were grouped into categories to form preliminary themes, which were refined through iterative review and discussion. Themes included dark data definitions, challenges in utilization, technological enablers, and ethical/legal concerns.

iv. Review and Refinement: The themes were reviewed for overlap and coherence, and further refined to ensure they captured the full spectrum of findings in the literature.

v. Software Used: NVivo 12 software was employed to assist in the coding process, allowing for a more systematic and organized extraction of themes. This software facilitated the identification of patterns across a large volume of qualitative data and provided a visual map of the themes and their relationships.

3.4. Ethical and legal considerations

Technology makes dark data exploitation more possible, but ethical and legal challenges remain.

Privacy, consent, and monitoring are major ethical issues. Using consumer or employee data without their consent can damage trust and morality. Unstructured data-trained algorithms may increase biases and penalize vulnerable people. Thus, ethical dark data stewardship involves technical safeguards and transparent governance mechanisms (Ayaz *et al.*, 2025).
Legal issues matter too. GDPR, HIPAA, and other rules restrict the usage of personal and sensitive data. Financial and reputational repercussions might come from noncompliance. Digital data can be stored and examined in many jurisdictions, making enforcement difficult. Thus, organizations attempting to open dark data must balance innovation with shifting standards in a complex legal environment (Theodos & Sittig, 2020).

4. RESULTS AND DISCUSSION

This section presents a synthesis of the key findings from the literature on dark data in Business Intelligence (BI). The evidence highlights the potential of dark data, the challenges businesses face in utilizing it, and the technological advancements that are aiding its integration into BI systems (Herhausen *et al.*, 2025).

4.1. Categorization of dark data

The review identified five major categories of dark data:

i. Textual Data: Includes emails, customer reviews, operational reports, and survey responses. These data are rich with insights about customer behavior, employee communication, and organizational processes but are often underutilized due to their unstructured nature (Le *et al.*, 2025).

ii. Numerical/Log Files: IoT sensor data, server logs, and transactional records are examples of numerical/log data. These datasets are organized at the source but become dark when not integrated into analytical pipelines(Al Kuwaiti *et al.*, 2023).

iii. Image-Based Data: Includes diagnostic images, product photos, and surveillance footage. Computer vision technologies have enabled the extraction of meaningful patterns from these datasets, but they remain underused due to challenges in analysis.

iv. Audio/Voice Data: Call center transcripts, conference recordings, and voice assistant logs often go unexamined. However, they hold potential for improving customer service and compliance monitoring through advanced audio analysis techniques.

v. Hybrid Datasets: These datasets combine multiple data types, such as social media posts that include text, images, and videos. The integration of these varied data sources offers significant potential for generating actionable insights.

Figure 1 shows a classification scheme of dark data categories and examples to show this variability.
Dark data comes in five main categories: textual, numerical/ log files, image-based, audio/voice, and hybrid, each with typical examples. The core node emphasizes “Dark Data” and its numerous sources and potential value in BI environments.

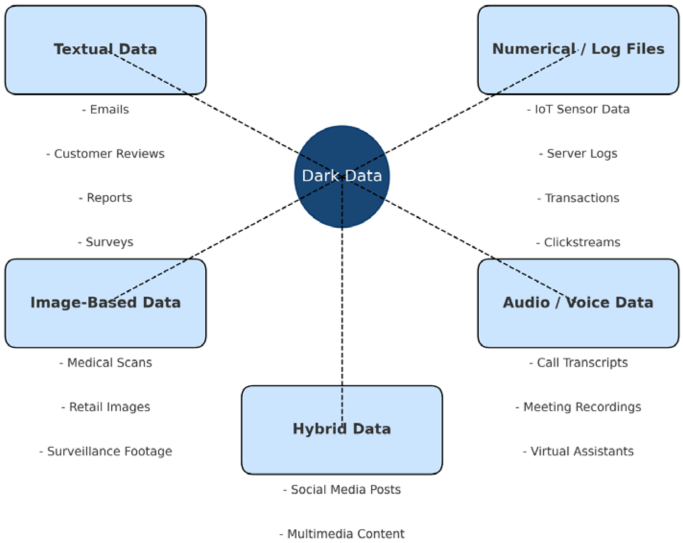

Figure 1. Business intelligence dark data classification.

4.2. Challenges in utilizing dark data

The review identifies several significant challenges to the effective use of dark data:

- *Data Quality Issues:* Dark data often suffers from issues like incompleteness, inconsistency, redundancy, and noise. For instance, IoT sensor logs may have duplicate or irrelevant entries, and customer reviews may contain unstructured text or slang that complicates sentiment analysis (Li *et al.*, 2024).
- *Integration Difficulties:* Many traditional BI systems are optimized for structured data, making it difficult to incorporate unstructured or semi-structured dark data. The presence of data silos within organizations further exacerbates this problem, limiting the ability to create a holistic view of data across departments (Mansouri *et al.*, 2021).
- *High Analysis Costs:* Storing, processing, and analyzing large volumes of unstructured data can be expensive, requiring significant computational resources and infrastructure. Cloud-based analytics platforms are helping to reduce these costs, but many businesses still face challenges related to financial and technical resources (Stefanovic *et al.*, 2025).

Figure 2. Problems with operations vs. solutions with technology in dark data analytics.

This diagram shows how to use dark data by showing the main operational problems, such as low data quality, hard integration, high analysis costs, and compliance risks, against the right technological solutions, such as advanced preprocessing, cloud-based BI, machine learning pipelines, and privacy-preserving frameworks. The diagram shows how certain innovations directly address each obstacle, giving enterprises a useful way to find hidden data value.

4.3. Technological advancements in dark data analysis

Despite the challenges, technological advancements are enabling businesses to harness the value of dark data:

• *Natural Language Processing (NLP):* NLP techniques are increasingly used to extract meaning from text-based dark data such as emails, customer reviews, and social media posts. Sentiment analysis, entity recognition, and topic modeling are commonly employed to gain insights from unstructured text data (Aminzadeh *et al.*, 2025).

• *Machine Learning (ML):* ML algorithms are utilized for pattern recognition, anomaly detection, and predictive analytics, especially in numerical/log-based datasets such as IoT sensor data and transaction logs. These techniques allow businesses to uncover hidden trends and forecast future events (Dichev *et al.*, 2025).

• *Computer Vision:* In industries such as healthcare and retail, computer vision technologies are being used to analyze image-based dark data. For example, diagnostic images can be analyzed to detect early signs of disease, while retail product photos can be used to predict trends in consumer preferences (Madanchian, 2024).

• *Knowledge Graphs:* These are used to structure and link disparate data sources, enabling a more comprehensive understanding of the relationships between different data points. This technology helps to break down data silos and create a more interconnected view of organizational data.

Table 1 highlights how businesses in several fields have used black data.

Table 1. Case studies of dark data applications in business intelligence

Industry	Dark Data Source	Analytical Approach	Application	Outcomes
Manufacturing	IoT sensor logs	Machine Learning	Predictive maintenance	Cost savings, reduced downtime, extended equipment lifespan (Olaitan *et al.*, 2025)
Finance	Call transcripts	Natural Language Processing (NLP)	Fraud detection	Risk reduction, improved compliance, enhanced trust (Coecke *et al.*, 2020)
Retail	Social media, customer feedback	Sentiment Analysis & Knowledge Graphs	Customer analytics	Higher customer satisfaction, improved targeting, retention gains (Ayaz *et al.*, 2025)
Healthcare	Medical records & diagnostic images	Computer Vision & AI-driven analytics	Clinical decision support	Improved care quality, faster diagnosis, better patient outcomes (Ahuja, 2019)

4.5. Discussion

This section addresses the research questions posed in the introduction and synthesizes the evidence from the literature to provide a comprehensive understanding of dark data in BI.

4.5.1 Research question 1: what is dark data and where does it come from in businesses?

Dark data refers to the vast amounts of information that businesses collect but fail to analyze. It includes unstructured data such as text (e.g., emails, customer feedback), numerical/ log data (e.g., IoT sensor data, transaction logs), image-based data (e.g., diagnostic images, product photos), and audio/ voice data (e.g., call center transcripts, voice assistant logs). These data are often stored in isolated systems or silos across the organization and are not incorporated into regular BI or analytics processes.

Industries vary in the sources of their dark data. For instance, healthcare providers deal with unstructured clinical notes and diagnostic images, while financial institutions generate transaction logs and call center recordings. Retailers may collect dark data in the form of customer interactions across multiple platforms, including social media, emails, and abandoned shopping carts. Despite the potential value of this data, it remains underutilized in most organizations (Siddique *et al.*, 2024).

4.5.2. Research question 2: what makes it hard to include dark data in mainstream analytics and bi systems?

There are several barriers to integrating dark data into mainstream BI systems:

• *Data quality:* Dark data often suffers from quality issues such as incompleteness, noise, and redundancy, which complicate efforts to process and analyze it effectively. For example, IoT sensor logs may contain erroneous or duplicate entries, and customer reviews may include informal language or irrelevant information.

• *Integration issues:* Traditional BI systems are designed to handle structured data, making it difficult to integrate dark data, which is often unstructured or semi-structured. The lack of interoperability between different systems and the presence of data silos within organizations further complicates the integration of dark data across departments.

• *High processing costs:* Storing and processing large volumes of unstructured data require significant computational resources. Cloud-based solutions are helping reduce these costs, but many organizations still struggle to afford the infrastructure and expertise needed for large-scale dark data analysis (Chiruvella & Guddati, 2021).

4.5.3 Research question 3: how can companies turn dark data into useful information that gives them an edge over their competitors?

To unlock the value of dark data, companies must overcome the challenges of data quality, integration, and processing costs. The review highlights several strategies that businesses can adopt:

• *Leveraging Advanced Technologies:* The use of NLP, machine learning, and computer vision allows businesses to extract actionable insights from unstructured data. For example, NLP can be used to analyze customer feedback and identify emerging trends, while machine learning can predict equipment failures based on IoT sensor data (Houssein *et al.*, 2025).

- *Data Governance and Compliance:* Organizations must ensure that their use of dark data adheres to legal and ethical standards, including compliance with regulations like GDPR. Implementing privacy-preserving AI techniques and ensuring proper data anonymization and consent mechanisms are critical for mitigating risks associated with dark data exploitation.
- *Cross-Functional Collaboration:* Breaking down data silos and fostering cross-functional collaboration is essential for unlocking the value of dark data. Knowledge graphs can help integrate disparate data sources, providing a more comprehensive view of organizational activities and enabling better decision-making.

Case studies in industries such as manufacturing, banking, and retail demonstrate how organizations that strategically invest in the right technologies and governance frameworks can turn dark data into a competitive advantage. For example, predictive maintenance in manufacturing can reduce downtime and repair costs, while fraud detection in banking can minimize financial losses and reputational risks (Williamson & Prybutok, 2024).

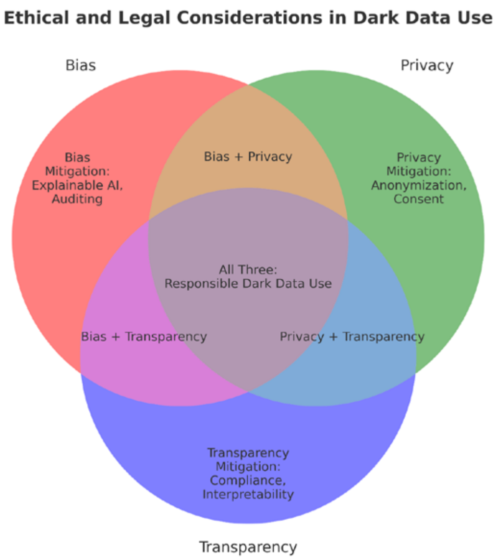

Figure 3. Dark Data Ethics and Law. A Venn diagram showing how bias, privacy, and transparency three main ethical and legal hazards in dark data use interact with mitigation techniques including explainable AI, anonymization, and regulatory compliance. Innovation and data governance must be balanced, as the intersections shown (Kgakatsi *et al.*, 2024).

5. CONCLUSION

Dark data holds significant untapped potential in Business Intelligence. However, businesses face several challenges in utilizing it, including poor data quality, integration difficulties, and high analysis costs. Technological advancements in NLP, machine learning, computer vision, and knowledge graphs provide promising solutions to these challenges (Nwosu *et al.*, 2024; Pancić *et al.*, 2023). By adopting advanced technologies, implementing robust data governance frameworks, and fostering cross-functional collaboration, organizations can unlock the value of dark data and gain a competitive edge. Future research should focus on developing industry-specific frameworks and investigating emerging technologies to further enhance dark data utilization while ensuring privacy and compliance (Eke & Stahl, 2024).

REFERENCES

Afzal, S., Ghani, S., Hittawe, M. M., Rashid, S. F., Knio, O. M., Hadwiger, M., & Hoteit, I. (2023). Visualization and visual analytics approaches for image and video datasets: A survey. *ACM Transactions on Interactive Intelligent Systems, 13*(1), 1-41. https://doi.org/10.1145/3576935

Ahuja, A. S. (2019). The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ, 7*, e7702. https://doi.org/10.7717/peerj.7702

Al Kuwaiti, A., Nazer, K., Al-Reedy, A., Al-Shehri, S., Al-Muhanna, A., Subbarayalu, A. V., Al Muhanna, D., & Al-Muhanna, F. A. (2023). A Review of the Role of Artificial Intelligence in Healthcare. *Journal of Personalized Medicine, 13*(6), 951. https://doi.org/10.3390/jpm13060951

Aminzadeh, A., Sattarpanah Karganroudi, S., Majidi, S., Dabompre, C., Azaiez, K., Mitride, C., & Sénéchal, E. (2025). A Machine Learning Implementation to Predictive Maintenance and Monitoring of Industrial Compressors. *Sensors, 25*(4), 1006. https://doi.org/10.3390/s25041006

Ayaz, O., Tabaghdehi, S. A. H., Rosli, A., & Tambay, P. (2025). Ethical implications of employee and customer digital footprint: SMEs perspective. *Journal of Business Research, 188*, 115088. https://doi.org/10.1016/j.jbusres.2024.115088

Bresciani, S., Ciampi, F., Meli, F., & Ferraris, A. (2021). Using big data for co-innovation processes: Mapping the field of data-driven innovation, proposing theoretical developments and providing a research agenda. *International Journal of Information Management, 60*, 102347. https://doi.org/10.1016/j.ijinfomgt.2021.102347

Chiruvella, V., & Guddati, A. K. (2021). Ethical Issues in Patient Data Ownership. *Interactive Journal of Medical Research, 10*(2), e22269. https://doi.org/10.2196/22269

Coecke, B., Felice, G. de, Meichanetzidis, K., & Toumi, A. (2020). *Foundations for Near-Term Quantum Natural Language Processing* (No. arXiv:2012.03755). arXiv. https://doi.org/10.48550/arXiv.2012.03755

Dichev, A., Zarkova, S., & Angelov, P. (2025). Machine Learning as a Tool for Assessment and Management of Fraud Risk in Banking Transactions. *Journal of Risk and Financial Management, 18*(3), 130. https://doi.org/10.3390/jrfm18030130

Eke, D., & Stahl, B. (2024). Ethics in the Governance of Data and Digital Technology: An Analysis of European Data Regulations and Policies. *Digital Society, 3*(1), 11. https://doi.

org/10.1007/s44206-024-00101-6

Fagbenle, E. (2025). Leveraging predictive analytics to optimize healthcare delivery, resource allocation, and patient outcome forecasting systems. *International Journal of Research Publication and Reviews, 6*(4), 6224–6239. https://doi.org/10.55248/gengpi.6.0425.14143

George, D. A. S., Dr.V.Sujatha, George, A. S. H., & Dr.T.Baskar. (2023). Bringing Light to Dark Data: A Framework for Unlocking Hidden Business Value. *Partners Universal International Innovation Journal (PUIIJ), 01*(04), 35–60. https://doi.org/10.5281/zenodo.8262384

Herhausen, D., Ludwig, S., Abedin, E., Haque, N. U., & de Jong, D. (2025). From words to insights: Text analysis in business research. *Journal of Business Research, 198*, 115491. https://doi.org/10.1016/j.jbusres.2025.115491

Hoofnagle, C. J., Sloot, B. van der, & Borgesius, F. Z. (2019). *The European Union general data protection regulation: What it is and what it means**. Information & Communications Technology Law. https://www.tandfonline.com/doi/abs/10.1080/13600834.2019.1573501

Houssein, E. H., Gamal, A. M., Younis, E. M. G., & Mohamed, E. (2025). Explainable artificial intelligence for medical imaging systems using deep learning: A comprehensive review. *Cluster Computing, 28*(7), 469. https://doi.org/10.1007/s10586-025-05281-5

Kgakatsi, M., Galeboe, O. P., Molelekwa, K. K., & Thango, B. A. (2024). The Impact of Big Data on SME Performance: A Systematic Review. *Businesses, 4*(4), 632–695. https://doi.org/10.3390/businesses4040038

Kong, H.-J. (2019). Managing Unstructured Big Data in Healthcare System. *Healthcare Informatics Research, 25*(1), 1–2. https://doi.org/10.4258/hir.2019.25.1.1

Lawal, O., Oyebamiji, H. O., Kelenna, I. J., Chioma, F. J., Oyefeso, E., Adeyemi, B. I., Foster-Pagaebi, E., & Moses, E. F. (2025). A Review on Usage of Digital Health Literacy to Combat Antibiotic Misuse and Misinformation in Nigeria: Review Article. *Journal of Pharma Insights and Research, 3*(2), Article 2. https://doi.org/10.69613/dja1jc18

Le, T. D., Le-Dinh, T., & Uwizeyemungu, S. (2025). Cybersecurity Analytics for the Enterprise Environment: A Systematic Literature Review. *Electronics, 14*(11), 2252. https://doi.org/10.3390/electronics14112252

Li, H., Lu, Y., & Zhu, H. (2024). Multi-Modal Sentiment Analysis Based on Image and Text Fusion Based on Cross-Attention Mechanism. *Electronics, 13*(11), 2069. https://doi.org/10.3390/electronics13112069

Madanchian, M. (2024). The Impact of Artificial Intelligence Marketing on E-Commerce Sales. *Systems, 12*(10), 429. https://doi.org/10.3390/systems12100429

Mansouri, T., Moghadam, M. R. S., Monshizadeh, F., & Zareravasan, A. (2021). *IoT Data Quality Issues and Potential Solutions: A Literature Review* (No. arXiv:2103.13303). arXiv. https://doi.org/10.48550/arXiv.2103.13303

Nwosu, N. T., Babatunde, S. O., & Ijomah, T. (2024). Enhancing customer experience and market penetration through advanced data analytics in the health industry. *World Journal of Advanced Research and Reviews, 22*(3), 1157–1170. https://doi.org/10.30574/wjarr.2024.22.3.1810

Olaitan, O. F., Akatakpo, O. N., Victor, C. O., Emejulu, C. J., Ayoola, T. M., Olayiwola, D. E., & Ajibola, A. A. (2025). Secure and Resilient Industrial IoT Architectures for Smart Manufacturing: A Comprehensive Review. *Journal of Engineering Research and Reports, 27*(6), 331–344. https://doi.org/10.9734/jerr/2025/v27i61548

Olaitan, O. F., Ayeni, S. O., Olosunde, A., Okeke, F. C., Okonkwo, U. U., Ochieze, C. G., Chukwujama, O. V., Akatakpo, O. N., Olaitan, O. F., Ayeni, S. O., Olosunde, A., Okeke, F. C., Okonkwo, U. U., Ochieze, C. G., Chukwujama, O. V., & Akatakpo, O. N. (2025). Quantum Computing in Artificial Intelligence: A Review of Quantum Machine Learning Algorithms. *Path of Science, 11*(5), Article 5. https://doi.org/10.22178/pos.117-25

Olaitan, O. F., Okoh, I., Onuche, P. U. O., Alabi, J. O., Sone, P. E., Kalyanaraman, K. K., Ukanu, J., & Arthur, C. (2025). The Role of Organic Solar Cells in U.S. Energy Transition: Technical Advances, Deployment Challenges, and Policy Pathways. *Journal of Environment, Climate, and Ecology, 2*(1), 49–60. https://doi.org/10.69739/jece.v2i1.546

Pancić, M., Ćućić, D., & Serdarušić, H. (2023). Business Intelligence (BI) in Firm Performance: Role of Big Data Analytics and Blockchain Technology. *Economies, 11*(3), 99. https://doi.org/10.3390/economies11030099

Patel, P. (2024). *Synthetic data*. Business Information Review. https://doi.org/10.1177/02663821241231101

Płaza, M., Kazała, R., Koruba, Z., Kozłowski, M., Lucińska, M., Sitek, K., & Spyrka, J. (2022). Emotion Recognition Method for Call/Contact Centre Systems. *Applied Sciences, 12*(21), 10951. https://doi.org/10.3390/app122110951

Schembera, B., & Durán, J. M. (2020). Dark Data as the New Challenge for Big Data Science and the Introduction of the Scientific Data Officer. *Philosophy & Technology, 33*(1), 93–115. https://doi.org/10.1007/s13347-019-00346-x

Sedlakova, J., Daniore, P., Horn Wintsch, A., Wolf, M., Stanikic, M., Haag, C., Sieber, C., Schneider, G., Staub, K., Alois Ettlin, D., Grübner, O., Rinaldi, F., & von Wyl, V. (2023). Challenges and best practices for digital unstructured data enrichment in health research: A systematic narrative review. *PLOS Digital Health, 2*(10), e0000347. https://doi.org/10.1371/journal.pdig.0000347

Shah, S., Ghomeshi, H., Vakaj, E., Cooper, E., & Fouad, S. (2023). A review of natural language processing in contact centre automation. *Pattern Analysis and Applications, 26*(3), 823–846. https://doi.org/10.1007/s10044-023-01182-8

Shahid, N. U., & Sheikh, N. J. (2021). Impact of Big Data on Innovation, Competitive Advantage, Productivity, and Decision Making: Literature Review. *Open Journal of Business and Management, 9*(2), 586–617. https://doi.org/10.4236/ojbm.2021.92032

Sharma, K., Shetty, A., Jain, A., & Dhanare, R. K. (2021). A Comparative Analysis on Various Business Intelligence (BI), Data Science and Data Analytics Tools. *2021 International Conference on Computer Communication and Informatics (ICCCI),* 1–11. https://doi.org/10.1109/ICCCI50826.2021.9402226

Siddique, S., Haque, M. A., George, R., Gupta, K. D., Gupta, D., & Faruk, M. J. H. (2024). Survey on Machine Learning Biases and Mitigation Techniques. *Digital, 4*(1), 1–68. https://doi.org/10.3390/digital4010001

Stefanovic, N., Radenkovic, M., Bogdanovic, Z., Plasic, J., & Gaborovic, A. (2025). Adaptive Cloud-Based Big Data Analytics Model for Sustainable Supply Chain Management. *Sustainability, 17*(1), 354. https://doi.org/10.3390/su17010354

Theodos, K., & Sittig, S. (2020). Health Information Privacy Laws in the Digital Age: HIPAA Doesn't Apply. *Perspectives in Health Information Management, 18*(1), 1l.

Williamson, S. M., & Prybutok, V. (2024). Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare. *Applied Sciences, 14*(2), 675. https://doi.org/10.3390/app14020675