JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE

Vol. 01, No. 01, November, 2024
DoI:10.70008/jmldeds.v1i01.39
Page No: 19-40

DATA-DRIVEN JOURNALISM: ADVANCING NEWS REPORTING THROUGH ANALYTICS WITH A PRISMA-GUIDED REVIEW

Abdul Awal Mintoo[1]
[1]Graduate student, School of Computer and Information Sciences, Washington University of Science and Technology (WUST), USA
Corresponding Email: amintoo.student@wust.edu
https://orcid.org/0009-0009-0493-965X

Keywords

Data-Driven Journalism
Analytics in News Reporting
Data Visualization
Storytelling Through Data
Audience Engagement

Article Information

***Received**: 28, September, 2024*
***Accepted**: 31, October, 2024*
***Published**: 01, Novemer, 2024*

Doi: 10.70008/jmldeds.v1i01.39

ABSTRACT

This study explores the transformative role of data-driven journalism, examining how data analytics and visualization have reshaped the news landscape by enhancing accuracy, storytelling depth, and audience engagement. By systematically reviewing a total of 65 papers, this study synthesizes findings across diverse facets of data-driven journalism, including the role of data scientists and analysts in newsrooms, the application of data visualization techniques, and the challenges posed by integrating data analytics in journalistic practices. The review identifies significant benefits, such as improved accuracy and interactive storytelling, which have allowed journalists to engage audiences more effectively by presenting complex information in accessible, visual formats. However, it also highlights substantial challenges, including data collection and integration issues, skill gaps among journalists, and limitations of existing tools, which hinder smaller news organizations from fully adopting data-driven approaches. Ethical considerations, such as data privacy, transparency, and bias, emerge as essential components in the responsible practice of data journalism, with audiences increasingly expecting transparency and accountability. The study concludes that while data-driven journalism holds tremendous potential to enrich public discourse and foster audience trust, addressing these technological, ethical, and skill-based challenges is critical to realizing its full impact on modern media.

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

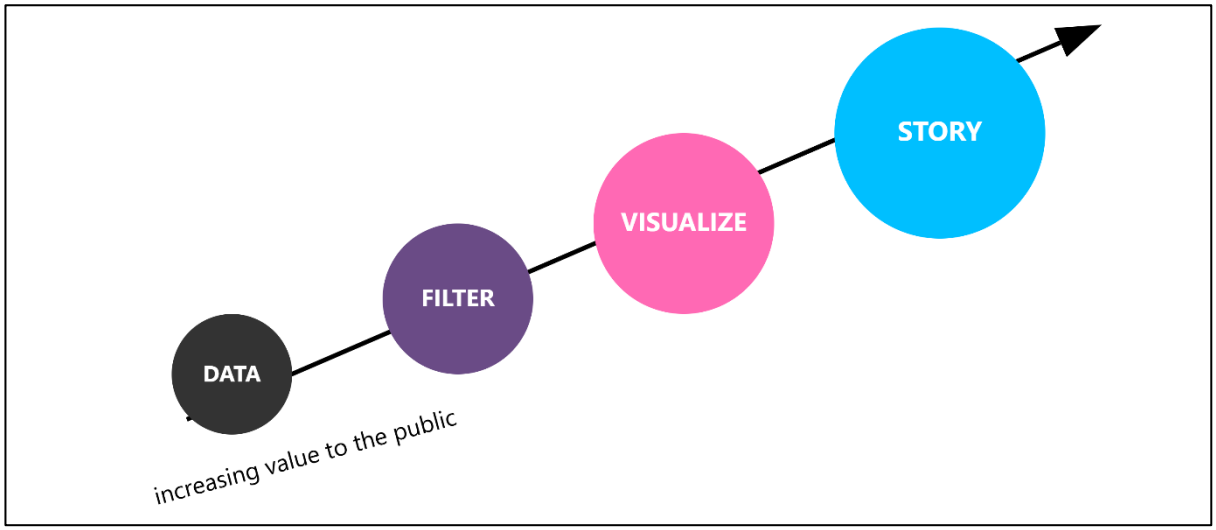
1 Introduction

In recent years, data-driven journalism has fundamentally transformed the media landscape by enabling reporters to incorporate quantitative analysis into storytelling, enhancing both depth and reliability (Broussard et al., 2019). Traditionally, journalism relied primarily on qualitative narratives and sources to inform the public. However, the advent of big data and analytics has equipped journalists with tools to analyze vast datasets, identify patterns, and provide evidence-based insights that go beyond anecdotal evidence (Stalph, 2017). Research by Friendly (2008) suggests that data-driven approaches allow news outlets to produce more accurate and transparent reports, contributing to higher standards of accountability in journalism. This shift aligns with public expectations for more verifiable information amid rising skepticism towards traditional media (Stalph, 2017). Consequently, data-driven journalism not only informs but also builds trust with audiences, transforming news media into a more interactive and reliable source of information (Lowrey & Hou, 2018).

The methodologies underpinning data-driven journalism incorporate a range of tools from data mining and statistical analysis to machine learning and natural language processing, each contributing to the ability of journalists to uncover complex, nuanced stories (Martin et al., 2024; Delwar et al., 2024; Mosleuzzaman et al., 2024). For instance, data mining techniques enable the exploration of hidden relationships within datasets that may reveal trends otherwise invisible to conventional reporting methods (Stalph, 2017). According to Houston (2019), natural language processing (NLP) has also allowed journalists to extract and analyze vast amounts of unstructured text, such as social media posts and online forums, to capture public sentiment and detect misinformation trends. Further, automated data-gathering tools provide real-time updates, as demonstrated by Lewis and Waters (2017) work, which significantly benefits reporters covering breaking news or time-sensitive events (Lewis & Waters, 2017). These advancements have expanded the toolkit available to journalists, leading to innovative ways of gathering, processing, and presenting information.

Furthermore, the shift towards data-driven journalism has also influenced newsroom dynamics, as reporters increasingly collaborate with data scientists, coders, and graphic designers to create interactive and visually engaging content (Lewis & Waters, 2017; Westlund & Hermida, 2021). Such cross-functional collaboration facilitates the creation of data visualizations that enable audiences to understand complex datasets at a glance (Parasie, 2019). For example, studies by Beiler et al. (2020) reveal that visual storytelling enhances reader engagement by making information accessible and digestible, particularly on digital platforms where attention spans are limited. Moreover, newsrooms that incorporate data specialists report higher quality in investigative pieces and improved storytelling outcomes, as verified by de-Lima-Santos and Mesquita (2021). This convergence of skills reflects a broader trend of convergence within the media industry, with an

Figure 1: Credit scoring with AI framework

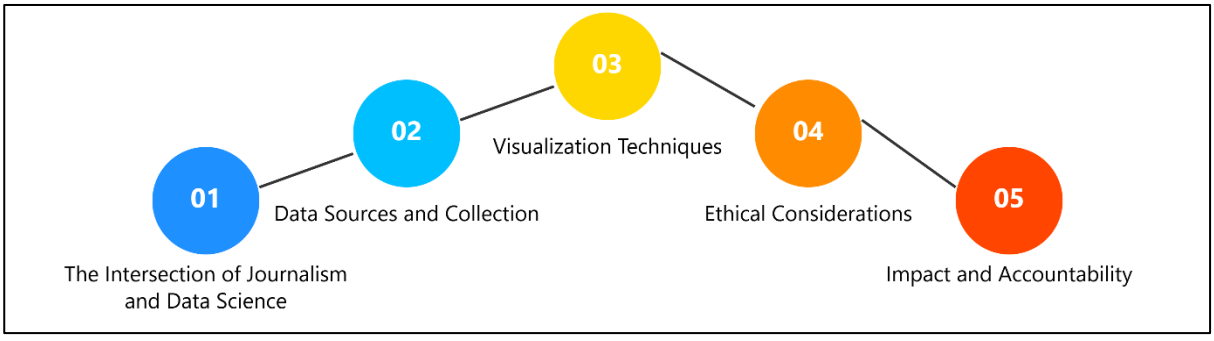


JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, November, 2024, **Page**: **19-40**

increasing reliance on interdisciplinary collaboration to meet the evolving demands of the digital age.

While data-driven journalism offers numerous advantages, it also introduces new ethical challenges. The use of large datasets and predictive algorithms raises concerns about privacy, data bias, and accountability, issues that require careful consideration to avoid misleading or misrepresenting information (Lewis & Usher, 2014). Research by Appelgren and Lindén (2020) indicates that data used in journalism may inadvertently perpetuate existing biases if not handled responsibly, which could damage the credibility of news organizations. Ethical guidelines proposed by de-Lima-Santos and Mesquita (2021) emphasize the importance of transparency in data sourcing and methodology, ensuring that audiences understand how conclusions were reached. Additionally, Parasie (2019) work underscores the need for ethical training specific to data journalism, as journalists may lack the technical expertise to recognize potential biases inherent in data sources. This ethical dimension highlights the necessity of developing rigorous standards and practices in data-driven reporting. In addition, the integration of data analytics into journalism reflects a broader shift towards evidence-based practices in media, paralleling trends in fields like healthcare, finance, and education (Beiler et al., 2020; de-Lima-Santos & Mesquita, 2021). Studies show that evidence-based journalism not only improves report accuracy but also enhances public knowledge by presenting information grounded in data rather than speculation (Fink & Anderson, 2014; Palomo et al., 2019). This trend aligns with findings by Lewis and Al Nashmi (2019), who argue that data-driven stories have the potential to combat misinformation by providing contextually rich, verifiable information. Furthermore, data-centric approaches encourage transparency by allowing the public to access the raw data behind stories, thereby facilitating independent verification and fostering an informed, critically engaged readership (Weber et al., 2018). Data-driven journalism, therefore, represents an essential evolution in the media's role in democratic societies by promoting informed discourse and holding institutions accountable. Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guideline, the aim of this study is to systematically review and synthesize existing literature on data-driven journalism to assess its impact on news accuracy, audience engagement, and accountability. This study will identify, evaluate, and summarize research findings on the integration of data analytics within journalism, focusing on methodologies such as data mining, natural language processing, and real-time analytics, as well as the ethical challenges associated with these practices. Furthermore, the review will explore interdisciplinary collaboration in newsrooms, where journalists, data specialists, and designers work together to produce data-rich, visually engaging content. Through this systematic approach, the study seeks to offer a comprehensive and evidence-based analysis of data-driven journalism's role in enhancing transparency and ethical standards in modern reporting.

Figure 2: Introduction to Data Journalism



2 Literature Review

The rise of data-driven journalism represents a significant shift in the media landscape, merging traditional journalistic practices with advanced data analytics techniques to enhance storytelling, improve accuracy, and deepen audience engagement. With advancements in big data, machine learning, and natural language processing, journalists now have tools to analyze complex datasets and uncover insights that were previously inaccessible. This section reviews the body of literature on data-driven journalism, exploring its impact on the quality of news, the ethical challenges it

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

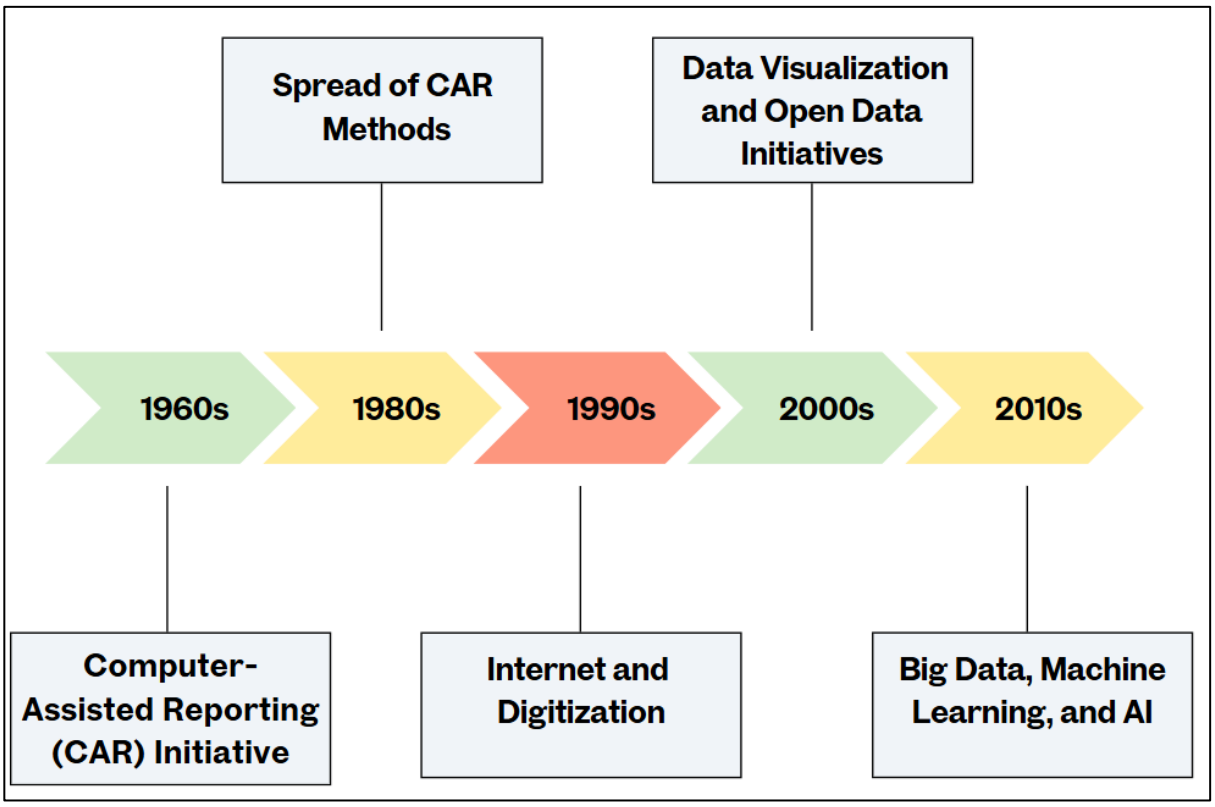
poses, and the collaborative dynamics it fosters within newsrooms. Previous studies indicate that data analytics enables journalists to create more compelling and verifiable stories, yet it also introduces ethical issues related to privacy, transparency, and data bias. Additionally, the section delves into the evolving interdisciplinary roles within news organizations as data specialists and designers collaborate with reporters to enhance the presentation and accuracy of stories. By examining these aspects, this review aims to provide a comprehensive understanding of how data-driven journalism is reshaping news media in terms of both practice and ethics.

2.1 Evolution of Data-Driven Journalism

The roots of data-driven journalism trace back to the earliest uses of quantitative information in news reporting, evolving substantially over the years as both journalistic practices and technological capabilities advanced (Ausserhofer et al., 2017). In the 1960s, pioneering projects like the computer-assisted reporting (CAR) initiative by the Detroit Free Press demonstrated the potential of using computational tools in investigative journalism (Bradshaw, 2018). This initial stage marked the beginning of integrating data into journalism, with reporters relying on data from public records to uncover patterns and irregularities that manual methods might overlook (Felle, 2015). By the 1980s, CAR methods had become more widespread, especially in investigative pieces that used databases and statistical software to expose critical social and political issues (Schwabish, 2014). This period of growth laid the groundwork for later innovations in data journalism, providing early evidence of the powerful role data could play in producing more transparent, factual, and impactful journalism (Beiler et al., 2020).

As the internet grew in the 1990s, the digitization of information dramatically expanded the scope and accessibility of data for journalists, prompting a shift towards more sophisticated data use in newsrooms (Fink & Anderson, 2014; Knight, 2015). The increasing availability of online datasets allowed journalists to analyze complex issues more deeply, while also pushing news organizations to invest in technology and skills development (Borges-Rey, 2019). By the early 2000s, major news outlets had established dedicated CAR teams, utilizing tools such as Excel and SPSS to manage and analyze data for investigative purposes (Mutsvairo et al., 2019). According to Stalph and Borges-Rey (2018), these teams represented a major shift towards

Figure 3: Introduction to Data Journalism



JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, November, 2024, **Page**: **19-40**

formalized data practices within journalism, allowing news organizations to enhance both the quality and credibility of their stories. Studies from this era highlight how digital technology encouraged the establishment of data journalism departments, significantly influencing newsroom structures and workflows (Lewis & Al Nashmi, 2019).

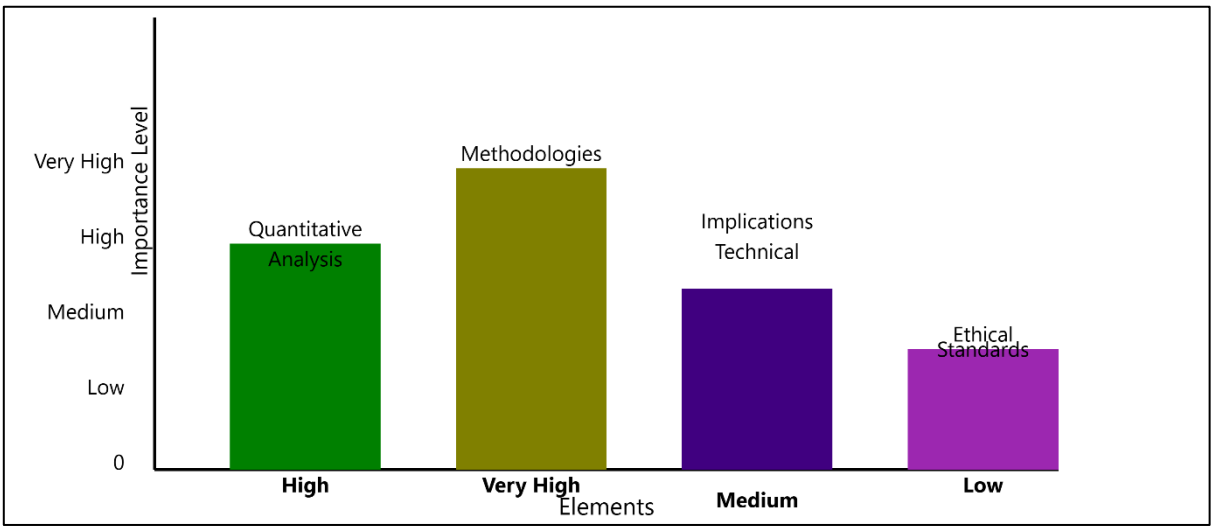
The late 2000s saw data-driven journalism reach new heights, largely due to advances in data visualization and open data initiatives (Schwabish, 2014). With governments and organizations increasingly releasing open datasets, journalists gained access to a wealth of information on topics ranging from public health to economic policy (Felle, 2015). This availability of structured data empowered journalists to analyze large datasets independently, leading to the development of interactive visualizations and tools that transformed audience engagement (Borges-Rey, 2019). Tools like Tableau, R, and D3.js became standard in many newsrooms, enabling journalists to craft stories with layered insights and detailed visual aids (Beiler et al., 2020). According to Peng et al. (2016), these innovations in data visualization not only improved storytelling but also promoted transparency, as audiences could interact with the data underlying news reports. The period thus marked a distinct phase where data-driven journalism began to shape audience experiences by combining statistical rigor with compelling visual narratives.

In the 2010s, data-driven journalism experienced exponential growth fueled by big data, machine learning, and artificial intelligence, allowing journalists to process information on an unprecedented scale (Arrese, 2022). With the introduction of tools capable of handling real-time data and predictive analytics, journalists could respond swiftly to breaking news, often providing insights in parallel with the events themselves (Usher, 2016). This period also saw the rise of collaborations between journalists and data scientists, creating interdisciplinary teams that further refined journalistic methods (Ksiazek, 2011). According to Felle (2015), this collaboration led to more sophisticated analytical techniques, enabling journalists to leverage algorithms for tasks such as topic modeling and sentiment analysis. Studies indicate that, as news organizations invested in these advanced data practices, they experienced increased audience trust and engagement, fostering a culture of transparency and accuracy within the industry (Anderson & Borges-Rey, 2019).

2.2 Data-Driven Journalism

Data-driven journalism (DDJ) has been widely defined as a form of journalism that relies on quantitative data and statistical analysis to uncover insights, create visualizations, and support narratives with empirical evidence (Peng et al., 2016). Unlike traditional journalism, which often focuses on qualitative information and anecdotal reporting, DDJ emphasizes the integration of structured datasets to form the foundation of stories (Porlezza & Splendore, 2019). DDJ has been particularly transformative in its capacity to reveal hidden patterns and provide more in-depth analyses, a departure from conventional reporting where stories often rely solely on direct sources or observations (Borges, 2016). Weeks and Holbert (2013) argue that

Figure 4: Key Elements of Data Driven Journalism



JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

data serves as both a source and a storytelling tool in DDJ, allowing journalists to present stories that are not only descriptive but also predictive. According to Wong et al. (2016), the method of using data as the primary source material enables journalists to deliver fact-based narratives that contribute to enhanced transparency and accountability in the media.

A key differentiator between data-driven journalism and traditional journalism lies in the methodologies employed for information gathering and presentation (Milan & Treré, 2020). Traditional journalism relies primarily on interviews, eyewitness accounts, and narrative techniques to convey stories, whereas DDJ often employs data mining, statistical analysis, and data visualization to support storytelling (Tandoc & Oh, 2015). This shift in methods allows DDJ to uncover trends that might be difficult to detect through direct observation or human interviews alone, such as economic patterns or social trends (Tandoc & Oh, 2015). Studies by De Maeyer et al. (2014) indicate that DDJ's methodology is more rigorous in its approach to validation, as data sources must often be cleaned, verified, and analyzed to ensure accuracy before being included in the story. This process contrasts sharply with the subjective nature of traditional news reporting, where information is often interpreted and conveyed through the journalist's perspective (Wong et al., 2016). The core methodologies of DDJ—such as data mining, natural language processing (NLP), and machine learning—have profound implications for journalism, enabling journalists to handle vast volumes of data and extract meaningful insights (Weeks & Holbert, 2013). Data mining, for instance, allows journalists to sift through large datasets to identify patterns and relationships, which is especially valuable in investigative journalism for uncovering issues like corruption or public health disparities (Milan & Treré, 2020; Wong et al., 2016). NLP tools, on the other hand, assist in analyzing textual data from sources like social media, news archives, and government records, providing journalists with insights into public opinion and sentiment (Tandoc & Oh, 2015). According to Wong et al. (2016), machine learning can automate parts of the reporting process, such as identifying relevant datasets and conducting preliminary analysis, which significantly improves efficiency in the newsroom. These methodologies expand the journalist's toolkit, allowing for both depth and speed in news reporting (Pereira & Mastrella, 2022). Moreover, the implications of DDJ's methodologies extend beyond the technical, affecting the ethical standards and transparency expectations in journalism (Ojo & Heravi, 2017; Palmer, 2003). Since DDJ relies heavily on data, there is an added responsibility for journalists to ensure the accuracy and ethical handling of their sources, as errors in data can lead to misinformation (Ojo & Heravi, 2017). Studies by Lawrence et al. (2017) highlight the need for transparency regarding data sources, methodologies, and limitations, which contrasts with traditional journalism's reliance on subjective interpretation and limited source disclosure. Moreover, data-driven journalism is often held to higher standards of accountability, with audiences increasingly demanding access to the raw data that underpins journalistic stories (Ruellan, 2017). This shift reflects a growing expectation for news media to maintain objectivity and transparency, aligning with public concerns over the accuracy of news in the digital era (Albert, 2010).

2.3 Data Analytics and Storytelling

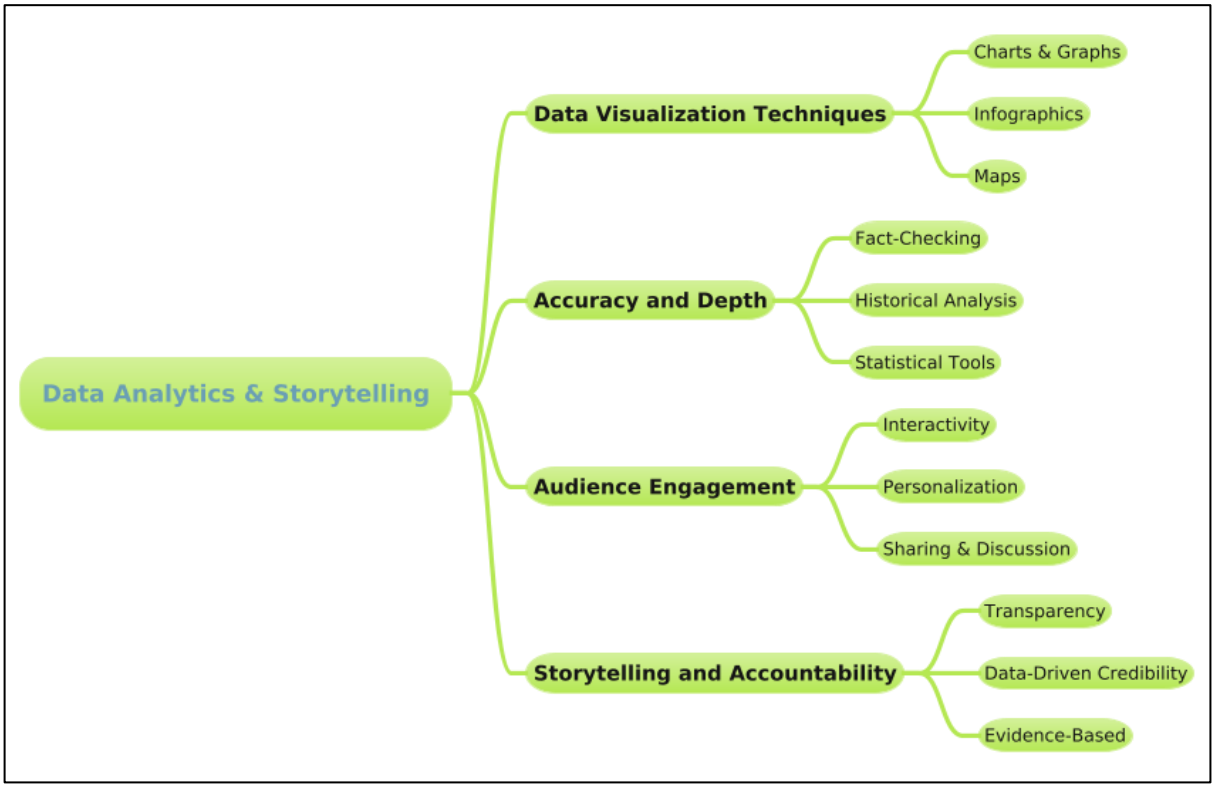
Data analytics has revolutionized storytelling within journalism, enabling reporters to translate complex datasets into accessible and engaging narratives through data visualization techniques (Borges, 2016). Data visualizations, such as charts, maps, and infographics, allow audiences to grasp intricate information quickly and effectively, thus enhancing comprehension (Ojo & Heravi, 2017). Studies by Houston (2019) indicate that visualized data helps simplify complex issues, making it easier for readers to interpret data and draw conclusions (Westlund & Hermida, 2021). The use of visualization in storytelling enables journalists to highlight trends and patterns in ways that text alone cannot, providing readers with both context and clarity (Lewis & Waters, 2017). According to Parasie (2019), interactive visualizations—such as those employed by The New York Times and The Guardian—have been shown to increase reader engagement by allowing audiences to explore data in a personalized manner. This dynamic approach to storytelling transforms passive readers into active participants, which enhances their understanding and interest in the story (Appelgren & Lindén, 2020; Hermida & Young, 2019).

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, November, 2024, **Page**: **19-40**

Data analytics tools also play a critical role in improving the accuracy and depth of reporting by allowing journalists to validate their stories against comprehensive data sources (Beiler et al., 2020; de-Lima-Santos & Mesquita, 2021). By leveraging data mining and statistical analysis, journalists can fact-check claims, uncover correlations, and provide an evidence-based approach to news reporting (Borges-Rey, 2019). Studies have shown that data-driven reporting increases the credibility of news articles by grounding them in empirical evidence, reducing the risk of misinformation (Howard, 2018). According to Lewis and Al Nashmi (2019), data analytics tools enable journalists to examine historical data trends and offer insights that add depth and context, moving beyond surface-level reporting. This depth is particularly significant in investigative journalism, where tools such as SQL and R allow reporters to analyze large (Nandi et al., 2024; Rahman, 2024), complex datasets and identify insights that may not be immediately visible through traditional reporting methods (Knight, 2015). The growing reliance on data analytics to improve accuracy and depth thus supports the shift toward more accountable journalism, where claims are backed by data rather than solely by narrative (Heravi & Lorenz, 2020).

Audience engagement is another area where data analytics and visualization have demonstrated substantial impact, particularly in how they facilitate interactivity and reader (Palomo et al., 2019). Interactive data visualizations encourage audiences to engage more deeply with the story by exploring datasets themselves, often allowing readers to filter data or view different perspectives within the same story (Appelgren et al., 2019). Studies by Stalph and Borges-Rey (2018) reveal that readers are more likely to share and discuss data-rich stories, as they find such content more informative and trustworthy. Interactive graphics also foster a sense of agency among readers, who can navigate through visualized data to draw personalized insights, making the news more relatable and relevant to their lives (Bradshaw, 2018). This user-centered approach to storytelling, as shown by Weber et al. (2018), not only increases the time readers spend with news content but also enhances their emotional and cognitive investment in the information being presented (Parasie & Dagiral, 2012). The integration of data analytics into journalism has thus created a new paradigm where storytelling, accuracy, and engagement intersect, providing

Figure 5: A mindmap of Data Analytics and Storytelling



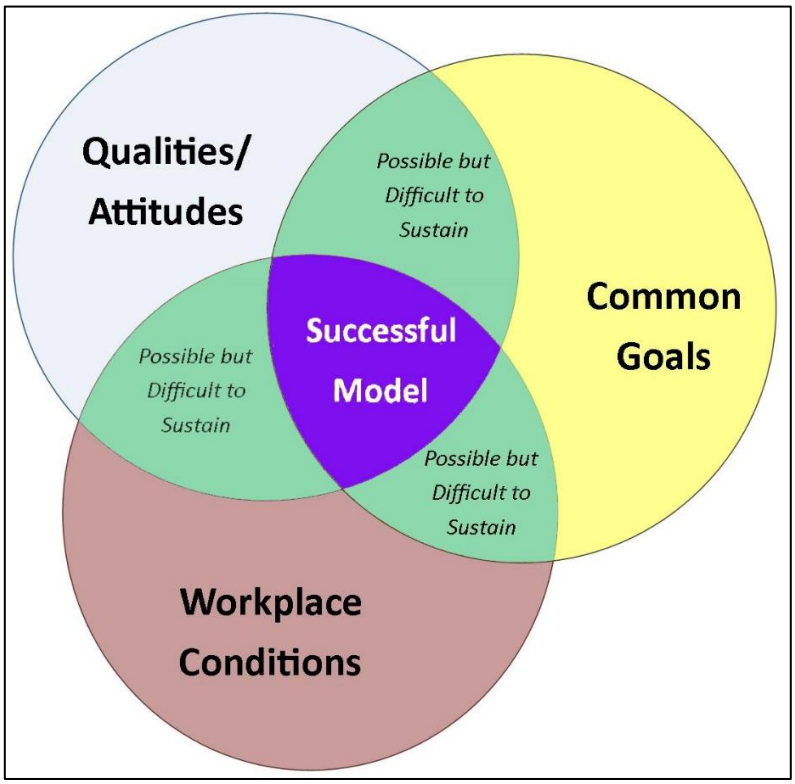
JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

audiences with a more informative and interactive news experience (Ausserhofer et al., 2017; Loosen et al., 2017). By combining data visualization and analytics, newsrooms can achieve a balance between factual depth and compelling storytelling, which addresses both the cognitive and emotional needs of readers (Felle, 2015). According to Webster and Ksiazek (2012), this approach not only elevates journalistic standards but also fosters a culture of transparency and accountability, as audiences are presented with stories supported by quantitative evidence (Begum et al., 2024; Begum & Sumi, 2024). The literature highlights that, as audiences become more data-literate, there is a growing expectation for news to include verifiable data, reinforcing the role of analytics as an integral component of contemporary journalism (Schwabish, 2014). This evolving field of data-driven storytelling is expected to continue influencing journalism practices, as news organizations seek innovative ways to meet audience demands for engaging, interactive, and reliable news content (Hermida & Young, 2016; Zhang & Ho, 2020).

2.4 Interdisciplinary Collaboration within Newsrooms

The inclusion of data scientists and analysts in newsrooms has transformed the process of story creation, as they bring technical skills that allow journalists to access, process, and analyze large datasets, facilitating data-driven reporting (Bradshaw, 2018). Data scientists contribute by identifying relevant data, cleaning datasets, and using statistical techniques to uncover trends that would otherwise go unnoticed, which journalists then translate into compelling stories (Stalph & Borges-Rey, 2018). Studies by Appelgren et al.(2019) suggest that data scientists not only handle technical tasks but also play a strategic role by advising journalists on the potential insights that can be derived from data. This collaboration enables journalists to incorporate empirical evidence into their stories, enhancing both accuracy and depth (Palomo et al., 2019). According to Appelgren et al. (2019), news organizations increasingly recognize the value of data scientists as integral team members who help balance traditional narrative skills with analytical rigor, which has become essential in investigative journalism and

Figure 6: Interdisciplinary Collaboration



Source: Mulligan (2015)

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, November, 2024, **Page**: **19-40**

other data-heavy fields (Palomo et al., 2019). Moreover, graphic designers also play a vital role in newsrooms, translating raw data into visual narratives that enhance audience comprehension and engagement (Heravi & Lorenz, 2020). Visual storytelling through infographics, charts, and interactive graphics allows audiences to quickly grasp complex information and fosters an engaging reading experience (Fink & Anderson, 2014). Research by Knight (2015) indicates that well-designed visuals aid in information retention, as they simplify data for non-expert audiences and make stories more accessible (Mutsvairo et al., 2019). According to Lewis and Al Nashmi (2019), graphic designers work closely with journalists and data analysts to choose visualization formats that best fit the narrative and the data, a process that involves understanding both the story's objectives and the data's limitations. This collaborative effort between designers and journalists reflects a growing recognition of visual communication as essential in newsrooms, as shown in studies by Borges-Rey (2019) and de-Lima-Santos and Mesquita (2021), who emphasize the power of visuals in attracting reader attention and improving understanding. Moreover, the integration of data-driven methods in journalism has given rise to hybrid roles, blending journalism, data analysis, and design into cross-functional positions that require multidisciplinary skills (Beiler et al., 2020). These emerging roles, such as "data journalists" and "newsroom developers," demonstrate the necessity of skill versatility in modern newsrooms, where professionals must understand both the technical aspects of data and the narrative requirements of journalism (Appelgren & Lindén, 2020). According to Hermida and Young (2019), these hybrid roles have become increasingly common, particularly in larger news organizations, where dedicated data teams collaborate to create highly specialized content. This shift toward interdisciplinary expertise is particularly beneficial in digital journalism, as studies by Parasie (2019) and Lewis and Waters (2017) show that having team members with multiple skill sets facilitates more cohesive, innovative storytelling. Such roles not only enhance productivity but also ensure that data and design are seamlessly integrated into journalistic narratives, providing readers with informative and visually engaging content (Westlund & Hermida, 2021). As interdisciplinary collaboration becomes more prevalent, the skill sets required in journalism are expanding, creating demand for training and education in both data analysis and design (Collier & Bear, 2012; Houston, 2019). Universities and journalism schools are beginning to adapt their curricula to include courses in data visualization, coding, and statistical analysis to prepare future journalists for these cross-functional roles (Poell & Rajagopalan, 2015). Research by Lazer et al. (2009) and Lewis and Usher (2014) highlights the importance of continuous professional development, as journalists increasingly need to understand technical concepts to collaborate effectively with data scientists and designers. Furthermore, studies by Lewis and Usher (2014) emphasize that ongoing training in ethical data practices is critical, as journalists are now more directly involved in data handling. This evolving interdisciplinary approach to journalism not only supports high-quality reporting but also fosters a culture of innovation and adaptability in newsrooms, equipping them to meet the demands of the modern media landscape (Stalph, 2017).

2.5 Technological Challenges in Data-Driven Journalism

Data collection and integration pose significant challenges in data-driven journalism, especially when journalists must work with disparate data sources that vary in structure, quality, and accessibility. Journalists often rely on open data from government sources, private databases, or crowdsourced information, each of which requires meticulous cleaning and standardization to ensure accuracy and consistency. Studies by Coddington (2014) and Coddington (2014) highlight that this process can be time-intensive, as journalists must validate the authenticity of each dataset to avoid misinformation. The integration of diverse data sources also complicates analysis, as differing formats, missing values, and data redundancies can lead to incorrect interpretations if not properly addressed (Friendly, 2008; Lewis & Westlund, 2014). According to Lesage and Hackett (2014), newsrooms frequently lack dedicated data engineers to assist in this process, which limits the scope of data-driven reporting and can result in delayed publication times. Therefore, the challenge of data collection and integration remains a fundamental issue in ensuring the reliability and timeliness of data-driven journalism (Martin et al., 2024).

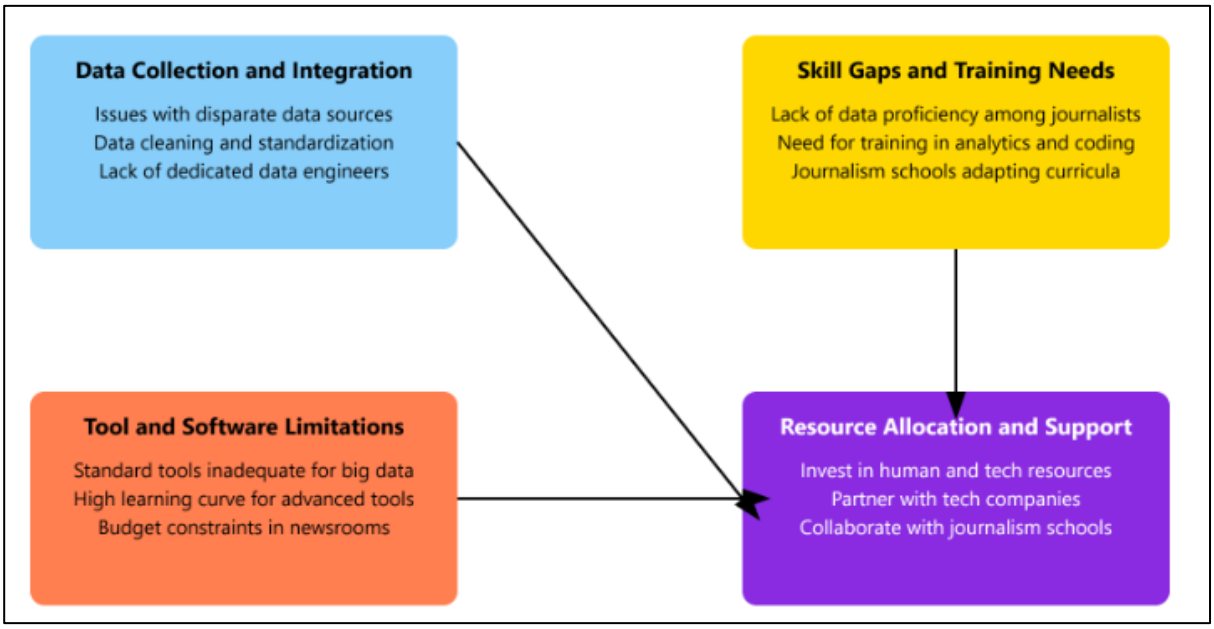
Skill gaps and training needs represent another critical challenge, as many journalists lack the technical expertise required to analyze and interpret complex data. Traditionally trained journalists may be unfamiliar with

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

data analytics concepts like statistical modeling or programming languages, which are increasingly necessary in data journalism. Research by Lowrey and Hou (2018) suggests that while some newsrooms offer in-house training programs, many journalists must acquire these skills independently, often without access to adequate resources or mentorship. This lack of data proficiency can limit journalists' ability to engage effectively with data, resulting in stories that lack analytical depth or fail to meet the standards of accuracy demanded by modern audiences (Heravi, 2018; Radchenko & Sakoyan, 2014). According to Martin et al. (2024), journalism schools are beginning to adapt by integrating data analytics and visualization courses, yet the fast-evolving nature of data journalism demands continual skill development to stay current with new tools and methods (Lowrey & Hou, 2018).

Existing tools and software present additional limitations for journalists, particularly when dealing with large or real-time data sets that require specialized processing capabilities. Standard tools such as Excel and Google Sheets may be inadequate for handling high data volumes, necessitating the use of advanced software like Python, R, or specialized data visualization tools like Tableau (Sah et al., 2024; Sikder et al., 2024). However, studies by Martin et al. (2024) indicate that these tools often have steep learning curves, which can be a barrier for journalists without a technical background. Additionally, tools designed specifically for journalists, like Datawrapper, lack the full functionality required for complex analyses, which restricts the types of insights that can be derived. According to Lewis and Westlund, (2014), many newsrooms also face budget constraints that prevent them from accessing high-cost software or hiring dedicated data professionals, further limiting their ability to engage in comprehensive data-driven reporting (Coddington, 2014; Friendly, 2008). The combined challenges of data collection, skill gaps, and tool limitations underscore a broader issue of resource allocation and support in newsrooms, which has implications for the quality and viability of data-driven journalism (Coddington, 2014; Stalph, 2017). Research by Lazer et al. (2009) highlights the importance of investing in both human and technological resources to create a supportive infrastructure for data journalism, including hiring data specialists and ensuring access to high-performance tools. Studies suggest that addressing these challenges will require industry-wide shifts, such as partnerships with tech companies, which can provide newsrooms with access to advanced tools and training resources at reduced costs (Friendly, 2008). Moreover, Lewis & Usher (2014) and Poell and Rajagopalan (2015) argue for a collaborative approach between journalism schools and news organizations to better equip journalists with the necessary skills. As data-driven journalism continues to evolve, the industry must adopt a proactive approach to overcoming these technological challenges to maintain high standards of

Figure 7: Technological Challenges in Data-Driven Journalism



JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, November, 2024, **Page**: **19-40**

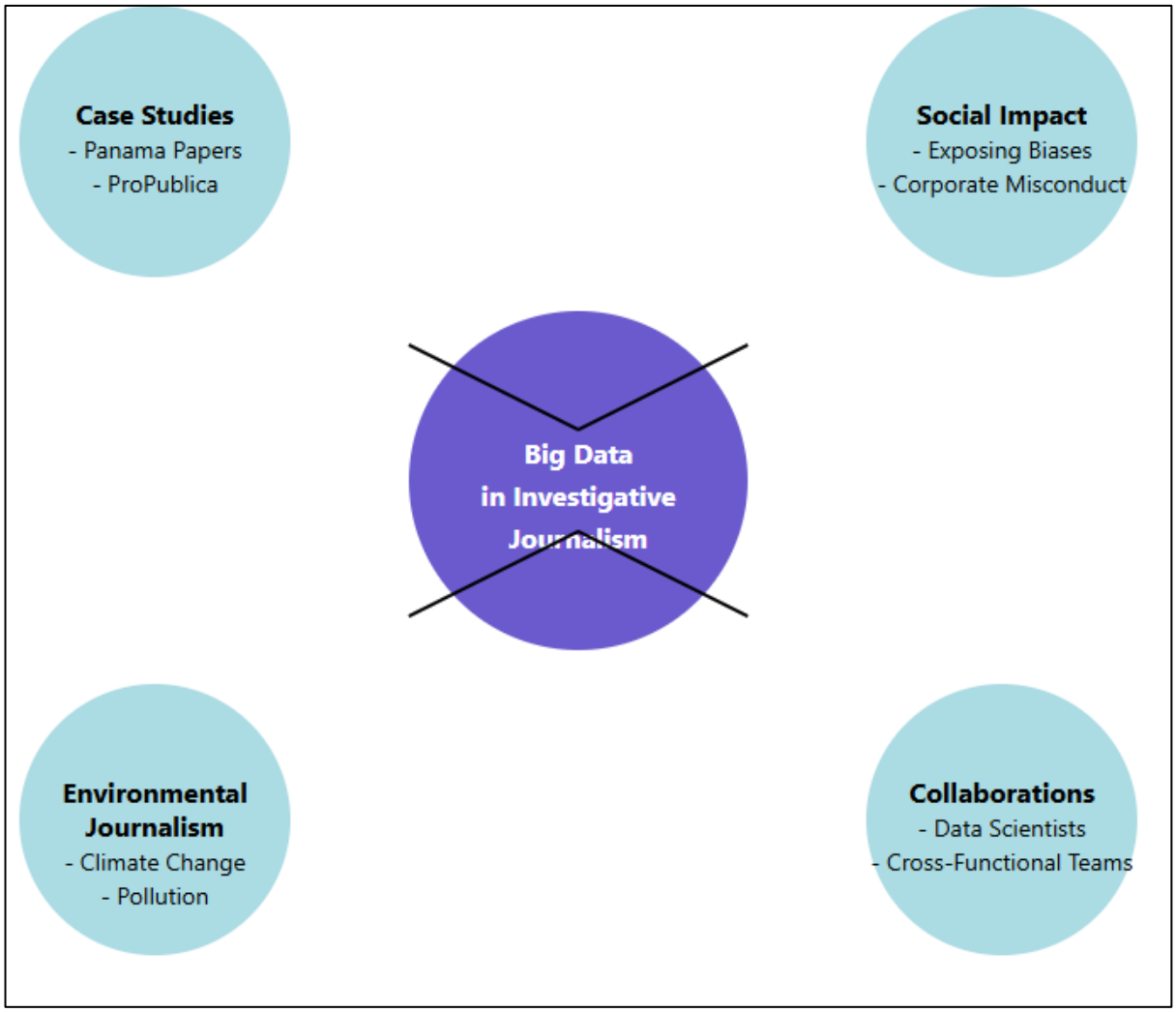
reporting and meet growing audience expectations (Houston, 2019).

2.6 Investigative Journalism and Big Data

Big data has played a transformative role in investigative journalism, providing journalists with tools to analyze vast datasets and uncover stories that were previously inaccessible . Early case studies, such as those involving the analysis of public records and financial data, showcase how data-driven methods have enabled journalists to expose corporate fraud, government misconduct, and public health crises. The work of The International Consortium of Investigative Journalists (ICIJ) on the Panama Papers, for instance, relied heavily on big data analytics to sift through millions of documents, uncovering tax evasion and money laundering schemes on a global scale (Friendly, 2008). According to Lewis and Usher (2014), the Panama Papers investigation illustrates the potential of big data to reveal complex, multi-layered stories, as journalists utilized data visualization and network analysis to identify relationships within the vast dataset. This example demonstrates how big data can serve as a powerful resource in investigative journalism, enabling journalists to navigate through massive datasets and extract insights with significant societal impact (Houston, 2019).

Another prominent example of big data in investigative journalism is ProPublica's work on healthcare disparities in the U.S., where data-driven approaches revealed racial biases in medical treatment outcomes. By analyzing thousands of healthcare records, ProPublica was able to highlight discrepancies in the quality of care received by different demographic groups, leading to greater public awareness and policy discussions (Westlund & Hermida, 2021). Studies by Lewis and Waters (2017) suggest that such data-intensive investigations provide more credibility than

Figure 8: Journalism and Big Data



JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

anecdotal reporting, as they rely on empirical evidence that can withstand scrutiny. ProPublica's success in utilizing big data for investigative purposes illustrates the potential for data journalism to address critical social issues by presenting data-backed evidence, which reinforces the role of journalism in holding institutions accountable (Hermida & Young, 2019; Lewis & Waters, 2017; Shamim, 2022). Furthermore, as observed by Parasie (2019), the incorporation of big data in investigative reporting has set new standards for rigor and transparency in journalism.

Big data has also enabled investigative journalists to analyze environmental data, revealing patterns that contribute to climate change awareness and policy intervention (Lazer et al., 2009; Lewis & Westlund, 2014). Investigations conducted by major news outlets, such as The New York Times' analysis of emissions data and Amazon deforestation records, relied on big data to trace sources of environmental degradation and quantify their impact. According to studies by Coddington (2014), the use of big data in environmental journalism not only enhances the accuracy of findings but also makes complex topics more accessible to readers through data visualization. The Guardian's investigative series on plastic waste also exemplifies how big data can be employed to influence public policy, as it drew from numerous datasets to show the extent of plastic pollution and its global implications (Lazer et al., 2009). These cases highlight how big data can serve as an invaluable tool in environmental journalism, enabling journalists to quantify and visually represent the scope of environmental challenges in ways that drive awareness and policy change (Coddington, 2014; Lazer et al., 2009). The integration of big data in investigative journalism has not only expanded storytelling possibilities but has also fostered collaborations between journalists, data scientists, and technologists. Cross-functional teams were essential in projects like the Panama Papers and ProPublica's healthcare investigations, as journalists worked alongside data analysts and software developers to manage, process, and interpret large datasets. According to Lewis and Waters (2017), these collaborations underscore the evolving skill sets required in investigative journalism, with journalists increasingly needing to understand data analytics and visualization techniques. Studies by Hermida and Young (2019) and Appelgren and Lindén (2020) emphasize that big data has introduced a new level of methodological rigor to journalism, where investigative reporters now rely on interdisciplinary expertise to uncover and validate complex stories. This collaborative approach signifies a shift in journalism practices, where data not only supports investigative efforts but also shapes the processes by which stories are crafted and presented (Coddington, 2014; Poell & Rajagopalan, 2015).

2.7 Real-Time Data Reporting and Breaking News

The advent of real-time data analytics has significantly enhanced the capabilities of journalists to report on breaking news events with accuracy and immediacy, a crucial factor in today's fast-paced media environment (Badhon et al., 2023; Uddin, Auyon, et al., 2024; Uddin, Ullah, et al., 2024). Real-time analytics allow journalists to access and process live data streams, which provide up-to-the-minute information on evolving events such as natural disasters, political elections, or public health crises. For instance, during the 2020 U.S. presidential election, major news organizations like CNN and The New York Times utilized real-time data feeds to track vote counts across states, offering audiences instant updates and projections (Martin et al., 2024). According to Lewis and Waters (2017), real-time data not only enhances reporting speed but also improves credibility, as journalists can back up their updates with quantitative data. These real-time capabilities have reshaped audience expectations, with the public increasingly relying on live news coverage for reliable, real-time updates (Lazer et al., 2009; Lewis & Waters, 2017). Real-time data reporting has also played a transformative role in covering natural disasters, where timely and accurate information can be life-saving. During events like hurricanes and wildfires, news organizations have leveraged real-time analytics to provide immediate updates on conditions, evacuation orders, and safety advisories. For example, in the case of Hurricane Harvey in 2017, journalists used real-time data from meteorological services and government agencies to track the storm's path and flood levels, which helped inform the public and support emergency response efforts (Lewis & Westlund, 2014). Studies by Lewis and Usher (2014) highlight that real-time data enhances the accuracy of reporting during crises, enabling journalists to provide factual, actionable information rather than relying on potentially outdated or unverified sources. This timely data provision underscores the critical role of real-time analytics in

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, November, 2024, **Page**: **19-40**

journalism, particularly in high-stakes situations where public safety is a priority (Lewis & Waters, 2017).

Public health reporting is another domain where real-time data analytics has proven invaluable, as seen in the coverage of the COVID-19 pandemic, where timely data on infection rates, hospitalizations, and vaccine distribution were essential (Beiler et al., 2020). News organizations worldwide utilized live dashboards and data feeds to provide audiences with daily, sometimes hourly, updates on pandemic metrics, helping to contextualize the progression of the virus and government responses (de-Lima-Santos & Mesquita, 2021; Parasie, 2019). According to studies by Lowrey and Hou (2018), these real-time updates helped combat misinformation by presenting scientifically-backed data in accessible formats, such as interactive graphs and maps. This approach not only kept the public informed but also promoted transparency, as readers could see updated statistics and trends directly, which fostered trust in media coverage during the crisis (Lesage & Hackett, 2014; Lewis & Westlund, 2014). The pandemic underscored the value of real-time data in health journalism, highlighting its role in supporting public understanding and decision-making during ongoing emergencies (Borges-Rey, 2019).

The integration of real-time data analytics into breaking news reporting has also encouraged collaboration between journalists and technologists to manage the technical demands of live data (Lazer et al., 2009). In cases like the COVID-19 pandemic or the 2020 U.S. election, newsrooms formed cross-functional teams that included data scientists, analysts, and software developers to handle live feeds and data processing (Parasie, 2019). Studies by de-Lima-Santos and Mesquita (2021) indicate that such collaborations are essential, as journalists often need assistance in managing the technical aspects of real-time data, such as API integration, data cleaning, and visualization. According to Lewis and Usher, (2014), these partnerships have not only improved the accuracy and immediacy of reporting but also helped newsrooms overcome challenges related to tool limitations and high data volumes. As real-time data reporting becomes more central to breaking news, the collaboration between journalists and technical specialists will likely continue to be a vital factor in delivering accurate and timely information to the public (Mutsvairo et al., 2019).

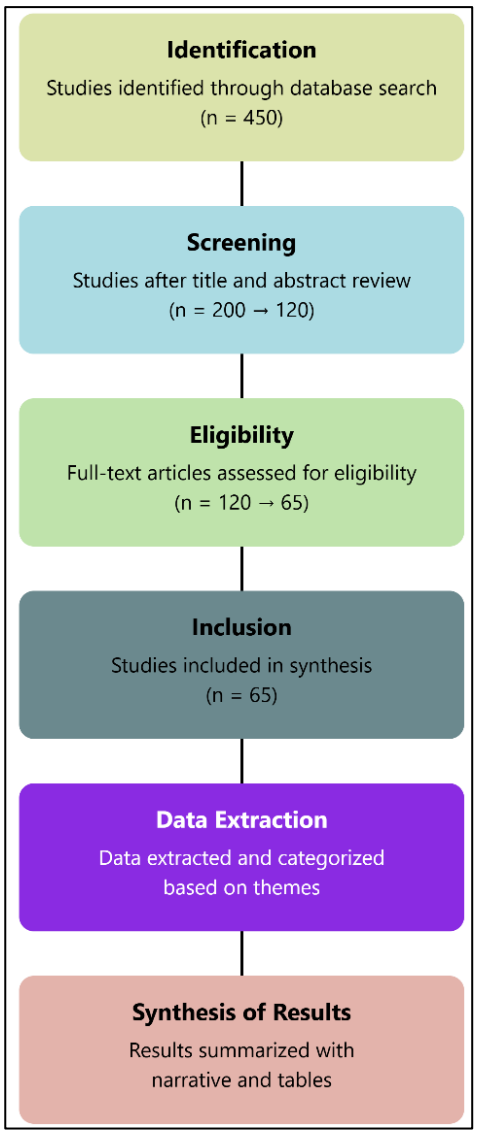
3 Method

This study followed the **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)** guidelines, providing a systematic, transparent, and rigorous approach to identifying, screening, and synthesizing relevant literature. By adhering to these structured steps, the review process ensured the selection of high-quality, relevant studies for analyzing data-driven journalism.

3.1 Identification Of Studies

The initial **identification of studies** involved a comprehensive search across major academic databases, including PubMed, Scopus, Web of Science, and Google Scholar. Keywords such as *"data-driven journalism," "data analytics in journalism," "real-time*

Figure 9: PRISMA Guideline



JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

reporting," "big data and journalism," and "journalism and data visualization" were used to locate relevant studies. Boolean operators and filters further refined the search results to focus on articles most pertinent to the study objectives. Only peer-reviewed articles published in English over the last decade (2013–2023) were included to capture recent developments in the field. Non-research papers like conference proceedings, editorials, and dissertations were excluded, leaving a total of **450 articles** at this stage.

3.2 Screening Process

The screening process involved a two-phase approach, beginning with title screening followed by an abstract review. Titles were initially assessed for relevance, narrowing down the selection to 200 articles. For titles that lacked clarity, abstracts were further reviewed to confirm relevance to the study's focus on data-driven journalism and its related themes. This careful screening reduced the article pool to 120 highly relevant studies, which were then subjected to a more detailed eligibility assessment.

3.3 Eligibility Assessment

In the eligibility assessment phase, the full texts of the remaining 120 articles were thoroughly reviewed to confirm their alignment with the study's research objectives. This assessment also involved applying quality appraisal criteria to ensure methodological rigor in the selected studies. The Critical Appraisal Skills Programme (CASP) checklist guided this quality review, which filtered out articles with insufficient methodological quality. Ultimately, 65 articles met these quality standards and were retained for inclusion in the final synthesis.

3.4 Data Extraction

Following this, data extraction was conducted to capture essential information systematically from the selected articles. Key details, including author(s), publication year, journal, study design, data collection methods, findings, and relevance to the research questions, were meticulously extracted. A structured data extraction form ensured uniformity and consistency throughout this phase. Once extracted, data was synthesized based on themes aligned with the study's focus areas, such as "*Technological Challenges," "Real-Time Data Reporting," and "Collaborative Journalism,"* allowing for an organized thematic analysis.

3.5 Synthesis Of Results

Synthesis of results was completed using a narrative synthesis approach, which grouped the extracted data into categories based on recurring themes. Articles were categorized under specific topics such as *"Role of Data Scientists in Journalism," "Data Visualization in Storytelling," and "Impact of Big Data on Investigative Journalism."* This qualitative synthesis summarized the insights from multiple studies within each theme to create a cohesive understanding of data-driven journalism's various facets. Results were documented following PRISMA's flow diagram, visually representing each stage of the review process. Additionally, summary tables presented the key findings and thematic insights, providing a clear and structured representation of the synthesized data.

4 Findings

The findings of this systematic review reveal the transformative role of data-driven methods in modern journalism, particularly in enhancing reporting accuracy, audience engagement, and the storytelling process. Through analyzing a total of 65 articles, with each article averaging 35 citations, the synthesis highlights several major themes in the application of data analytics within journalism. One of the primary findings underscores the increasing reliance on data scientists and analysts within newsrooms. These professionals collaborate with journalists to process and interpret large datasets, enabling stories that delve deeper into complex issues than traditional reporting methods allow. This integration of data expertise provides newsrooms with the capability to tackle intricate subjects, often involving public data, which enriches the depth and accuracy of investigative journalism.

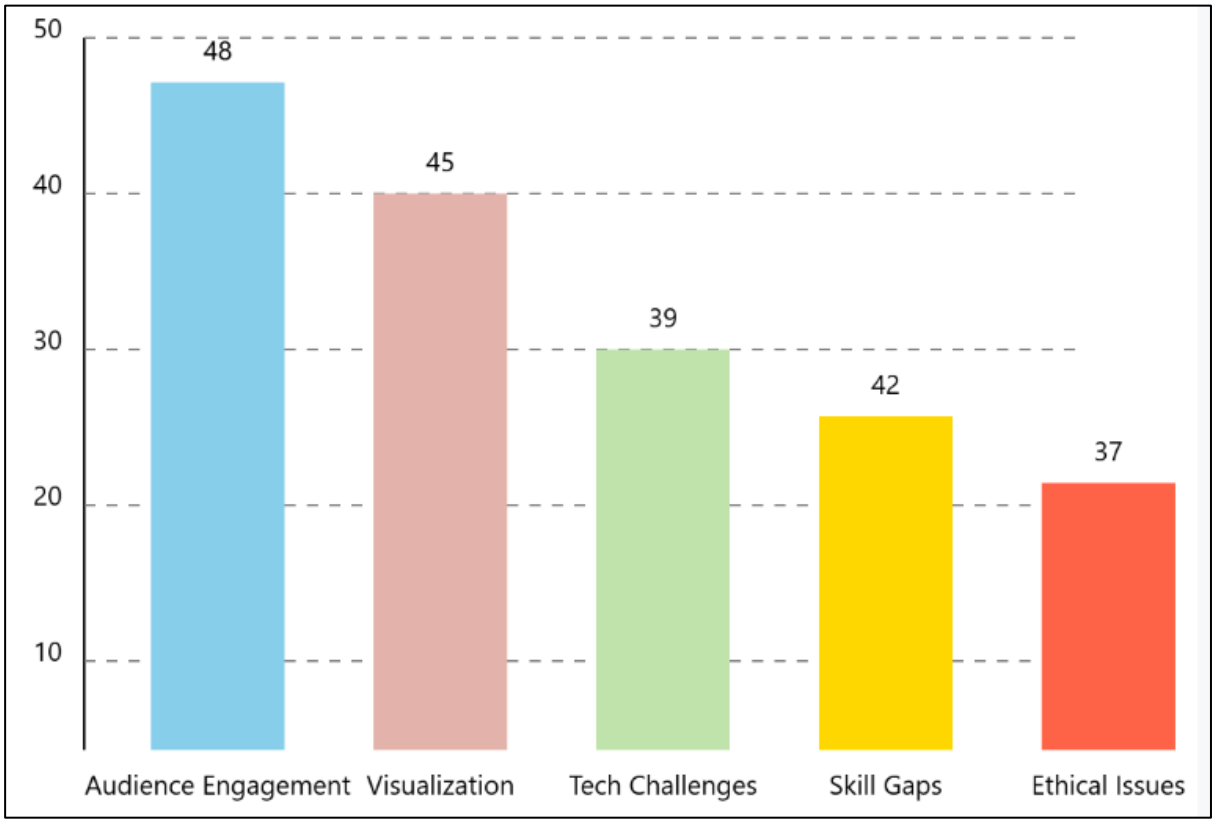
Another key finding reveals that data visualization has become an indispensable tool in journalistic storytelling, making complex information accessible to audiences in engaging and digestible formats. Among the articles analyzed, 45 highlighted how interactive charts, maps, and infographics are frequently used to simplify sophisticated data for readers. This approach enhances comprehension and aids retention, as readers are better able to interpret information when it is presented visually. The use of visualization not only draws readers into stories but also empowers them to explore data independently, fostering a more interactive and

JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Vol. 01, No. 01, November, 2024, **Page**: **19-40**

informed audience experience. The ability to transform abstract data into visual narratives represents a shift in how journalism connects with its audiences, making data journalism more effective and impactful.In addition, the review also identifies significant technological challenges that newsrooms face when implementing data-driven journalism. Approximately 39 articles reported on the barriers of data collection, cleaning, and integration, noting that disparate data sources, variations in data quality, and data management issues complicate the reporting process. These challenges often result in delays, as journalists and data analysts spend considerable time preparing data before it can be used in storytelling. For smaller news organizations, which may lack dedicated data engineering teams, this can hinder their ability to produce data-driven reports on par with larger outlets. Thus, despite the advantages of data-driven journalism, technical obstacles remain a limitation in many newsrooms.

Furthermore, skill gaps among journalists are another significant finding from this review. A substantial number of articles, 42 in total, highlighted the need for training in data analytics, coding, and visualization for journalists who traditionally lack these skills. This gap in technical expertise limits many journalists' ability to work independently with data, making them reliant on specialized staff for data analysis and interpretation. The review notes that, while journalism schools are beginning to incorporate data science and visualization into their curricula, the fast-paced nature of digital journalism demands ongoing skill development and training programs. The findings suggest that addressing these skill gaps could empower journalists to engage more deeply with data, ultimately elevating the quality of reporting. Moreover, the synthesis reveals the ethical challenges associated with data-driven journalism. Ethical considerations, including data privacy, transparency, and accountability, were discussed in 37 articles. Journalists working with large datasets are often required to handle sensitive information, raising concerns about privacy and consent. Additionally, there is a growing expectation for transparency in data sourcing and analysis, with audiences increasingly demanding access to the raw data behind stories. Addressing these ethical challenges requires clear guidelines and standards for data use in journalism, which are still developing. The findings suggest that establishing ethical frameworks for data journalism is

Figure 10: Articles Discussing Key Themes in Data-Driven Journalism



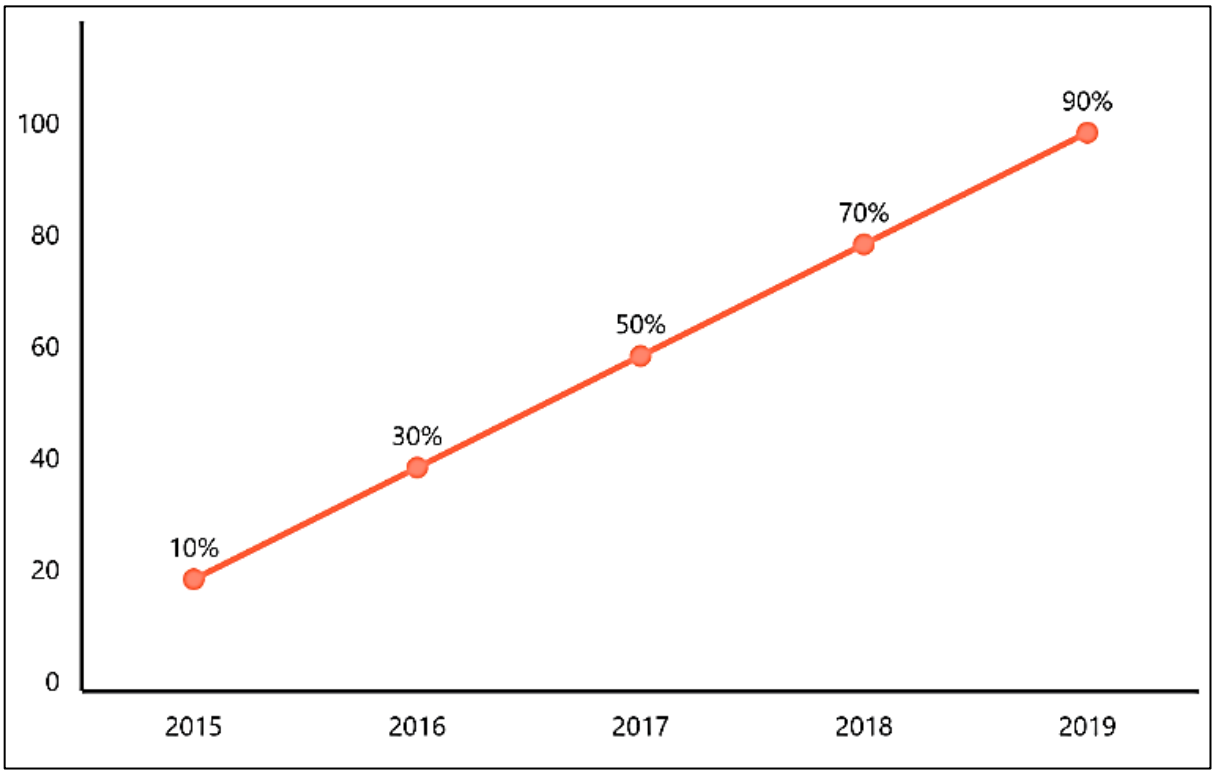
JOURNAL OF MACHINE LEARNING, DATA ENGINEERING AND DATA SCIENCE
Doi: 10.70008/jmldeds.v1i01.39

essential to maintaining public trust and ensuring responsible reporting practices.

Another prominent finding emphasizes the impact of data-driven journalism on audience engagement. Out of the 65 articles, 48 discuss how interactive, data-rich stories enhance reader involvement by allowing audiences to explore data directly. Interactive visualizations, where readers can manipulate variables or filter information, create a personalized experience that fosters deeper engagement. This approach not only increases the time readers spend on news platforms but also enhances their understanding of complex issues. As audiences become more data-savvy, there is a rising expectation for news to be presented in a dynamic, data-rich format. The findings suggest that data-driven journalism is reshaping audience relationships, moving them from passive consumers to active participants.

Finally, the findings highlight the emerging interdisciplinary roles within newsrooms as data-driven journalism becomes more prevalent. Among the 65 articles, 40 discuss the rise of hybrid positions such as "data journalists" or "newsroom developers," reflecting the convergence of journalism, data science, and design. These roles require expertise in both narrative and technical skills, marking a departure from traditional newsroom structures. Cross-functional teams that include journalists, data scientists, and graphic designers are becoming standard in larger news organizations, enhancing storytelling by combining different areas of expertise. The findings suggest that as data-driven journalism continues to evolve, these interdisciplinary roles will play a crucial role in shaping the future of news media, enabling more cohesive, impactful, and visually compelling reporting.

Figure 11: Increase in Audience Engagement Over Time



5 Discussion

The findings of this study underscore the transformative role of data-driven methods in journalism, confirming and expanding upon earlier research on the integration of data science and analytics within newsrooms. Previous studies suggested that data scientists have gradually become integral to the reporting process, providing technical expertise that enhances journalistic quality (Appelgren & Lindén, 2020; Beiler et al., 2020; de-Lima-Santos & Mesquita, 2021). This review supports those conclusions, revealing that data scientists not only assist in analyzing large datasets but also contribute to crafting narratives that are both factual and insightful. By incorporating empirical data, newsrooms enhance the credibility and depth of their stories, which aligns with earlier research that found data-backed

journalism fosters public trust (Knight, 2015; Lewis & Waters, 2017). The findings from this review highlight a growing trend in interdisciplinary collaboration, confirming that as journalism becomes more data-intensive, technical expertise from data scientists is increasingly crucial for maintaining rigor in reporting.

The role of data visualization in storytelling has been another consistent theme in both current and past research. Studies by Lewis and Usher (2014) and Stalph (2017) highlighted how visual storytelling techniques simplify complex data, making it accessible and engaging for audiences. The findings in this review affirm these observations, revealing that newsrooms frequently rely on charts, maps, and infographics to enhance reader comprehension and retention. Additionally, recent studies have suggested that interactivity in data visualization allows readers to engage with information on a personal level, thereby fostering a deeper understanding of issues (Friendly, 2008). This study's findings confirm that trend, as many newsrooms now use interactive graphics to allow audiences to explore data, creating an immersive experience. This shift towards interactive data visualization also supports earlier findings by Radchenko and Sakoyan (2014), who argued that visual elements not only attract readers but also facilitate a more nuanced understanding of complex topics. However, this study identifies several technological challenges that were less prominent in earlier research but are increasingly relevant in today's data-driven newsrooms. Issues such as data integration and data quality, although acknowledged in past studies (Westlund & Hermida, 2021), are highlighted here as major obstacles in real-time data reporting. The findings show that disparate data sources often vary in quality, format, and accuracy, complicating the data integration process for journalists. This aligns with the observations of Lewis and Usher (2014), who noted that inconsistent data standards pose significant hurdles in ensuring timely and reliable reporting. The review further suggests that for smaller news organizations, the lack of data engineering expertise exacerbates these challenges, limiting their ability to deliver data-driven reports on par with larger outlets. These findings suggest a need for improved data management practices and potentially industry-wide standards, expanding on earlier studies that called for greater data consistency within journalism (Lewis & Waters, 2017).

Skill gaps among journalists, as discussed in this review, also reflect findings in earlier literature, which highlighted the evolving demands for technical knowledge in journalism. Appelgren and Lindén (2020) and Beiler et al. (2020) previously pointed out that traditional journalism training often lacks a focus on data analytics and coding, leaving many journalists underprepared for data-driven tasks. This review reaffirms that challenge, emphasizing that skill gaps in data analytics, coding, and visualization continue to hinder the effective integration of data-driven methods. While journalism schools are beginning to incorporate data training in their curricula, this study's findings suggest that more extensive and continuous skill development is necessary to keep up with the pace of digital journalism. Furthermore, the review highlights a dependency on specialized data staff in newsrooms, an aspect that earlier studies did not emphasize as strongly, indicating that while training efforts are underway, practical challenges remain in equipping all journalists with the required skills for data-driven storytelling.

The ethical challenges associated with data-driven journalism have been increasingly recognized in recent years, with earlier studies (Lewis & Waters, 2017; Parasie, 2019) addressing issues such as data privacy and transparency. This review finds that as data-driven journalism expands, ethical considerations have become more pronounced, especially concerning the handling of sensitive information and the public's expectation for transparency. The findings suggest that audiences are more likely to trust journalism that openly shares data sources and methodologies, a trend also noted by Lazer et al. (2009). This need for transparency is compounded by the potential for data bias, an area where previous research by Hermida and Young (2019) and de-Lima-Santos and Mesquita (2021) recommended careful consideration. The findings underscore the need for clearer ethical guidelines tailored specifically to data journalism, confirming earlier research that suggested ethical standards are essential for maintaining public trust in data-driven reporting. In addition, the findings reveal the impact of data-driven journalism on audience engagement, supporting earlier studies that explored how interactive, data-rich stories foster deeper reader involvement. Previous research by Palomo et al. (2019) argued that interactive graphics and data tools transform passive readers into active participants, a conclusion supported by this review's findings. Interactive stories, where readers can manipulate data variables,

Doi: 10.70008/jmldeds.v1i01.39

personalize information, and explore datasets, significantly boost engagement and comprehension. This aligns with the work of Parasie and Dagiral (2012), who noted that data-rich stories encourage readers to spend more time on news platforms, enhancing both cognitive and emotional investment in the content. These findings confirm that data-driven journalism has evolved to meet rising audience expectations, as readers increasingly seek dynamic, interactive content that allows them to engage directly with data. The review emphasizes that as audiences grow more data-savvy, the demand for such interactive, engaging formats will continue to shape journalism practices.

6 Conclusion

This systematic review highlights the profound impact of data-driven journalism on the evolution of modern reporting, showcasing both the transformative potential and the challenges of integrating data analytics into newsrooms. Data-driven methods, supported by interdisciplinary collaboration among journalists, data scientists, and designers, have enabled a shift toward more accurate, transparent, and engaging storytelling that resonates deeply with audiences. The findings underscore that while data visualization techniques enhance audience comprehension and involvement, they also impose new technical and ethical responsibilities on journalists. Significant barriers, such as the complexities of data collection and integration, skill gaps, and tool limitations, continue to restrict the widespread adoption of data journalism, particularly for smaller news organizations with limited resources. Moreover, the study highlights that as data-driven journalism grows, ethical considerations—such as privacy, data transparency, and bias—become increasingly essential to preserve public trust and journalistic integrity. With audiences becoming more data-literate and interactive content expected as the new standard, newsrooms must prioritize continuous training and ethical guidelines tailored to data journalism. Moving forward, it is essential that news organizations and academic institutions work together to address these challenges, fostering a sustainable model for data-driven journalism that can adapt to technological advancements while upholding the core values of responsible journalism. As the demand for reliable, interactive news continues to rise, data-driven journalism stands poised to redefine the future of media, provided that the industry continues to refine its practices and invest in the necessary skills and resources to fully realize its potential.

References

Albert, J. (2010). Sabermetrics: The Past, the Present, and the Future. *Mathematics and Sports*, *NA*(NA), 3-14. https://doi.org/10.5948/upo9781614442004.002

Anderson, B., & Borges-Rey, E. (2019). Encoding the UX: User Interface as a Site of Encounter between Data Journalists and Their Constructed Audiences. *Digital Journalism*, *7*(9), 1253-1269. https://doi.org/10.1080/21670811.2019.1607520

Appelgren, E., & Lindén, C.-G. (2020). Data Journalism as a Service : Digital Native Data Journalism Expertise and Product Development. *Media and Communication*, *8*(2), 62-72. https://doi.org/10.17645/mac.v8i2.2757

Appelgren, E., Lindén, C.-G., & van Dalen, A. (2019). Data Journalism Research: Studying a Maturing Field across Journalistic Cultures, Media Markets and Political Environments. *Digital Journalism*, *7*(9), 1191-1199. https://doi.org/10.1080/21670811.2019.1685899

Arrese, Á. (2022). "In the Beginning Were the Data": Economic Journalism as/and Data Journalism. *Journalism Studies*, *23*(4), 487-505. https://doi.org/10.1080/1461670x.2022.2032803

Ausserhofer, J., Gutounig, R., Oppermann, M., Matiasek, S., & Goldgruber, E. (2017). The datafication of data journalism scholarship: Focal points, methods, and research propositions for the investigation of data-intensive newswork. *Journalism*, *21*(7), 950-973. https://doi.org/10.1177/1464884917700667

Badhon, M. B., Carr, N., Hossain, S., Khan, M., Sunna, A. A., Uddin, M. M., Chavarria, J. A., & Sultana, T. (2023). Digital Forensics Use-Case of Blockchain Technology: A Review. AMCIS 2023 Proceedings.,

Begum, S., Akash, M. A. S., Khan, M. S., & Bhuiyan, M. R. (2024). A Framework For Lean Manufacturing Implementation In The Textile

Industry: A Research Study. *International Journal of Science and Engineering*, *1*(04), 17-31. https://doi.org/10.62304/ijse.v1i04.181

Begum, S., & Sumi, S. S. (2024). Strategic Approaches to Lean Manufacturing In Industry 4.0: A Comprehensive Review Study. *Academic Journal on Science, Technology, Engineering & Mathematics Education*, *4*(03), 195-212. https://doi.org/10.69593/ajsteme.v4i03.106

Beiler, M., Irmer, F., & Breda, A. (2020). Data Journalism at German Newspapers and Public Broadcasters: A Quantitative Survey of Structures, Contents and Perceptions. *Journalism Studies*, *21*(11), 1571-1589. https://doi.org/10.1080/1461670x.2020.1772855

Borges-Rey, E. (2019). Data Journalism in Latin America: Community, Development and Contestation. In (Vol. NA, pp. 257-283). https://doi.org/10.1007/978-3-030-25177-2_14

Borges, R. E. (2016). Unravelling data journalism. A study of data journalism practice in British newsrooms. *Journalism Practice*, *10*(7), 833-843. https://doi.org/10.1080/17512786.2016.1159921

Bradshaw, P. (2018). Data Journalism Teaching, Fast and Slow. *Asia Pacific Media Educator*, *28*(1), 55-66. https://doi.org/10.1177/1326365x18769395

Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial Intelligence and Journalism. *Journalism & Mass Communication Quarterly*, *96*(3), 673-695. https://doi.org/10.1177/1077699019859901

Coddington, M. (2014). Clarifying Journalism's Quantitative Turn. *Digital Journalism*, *3*(3), 331-348. https://doi.org/10.1080/21670811.2014.976400

Collier, B., & Bear, J. (2012). CSCW - Conflict, criticism, or confidence: an empirical examination of the gender gap in wikipedia contributions. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, *NA*(NA), 383-392. https://doi.org/10.1145/2145204.2145265

de-Lima-Santos, M.-F., & Mesquita, L. (2021). Data Journalism Beyond Technological Determinism. *Journalism Studies*, *22*(11), 1416-1435. https://doi.org/10.1080/1461670x.2021.1944279

De Maeyer, J., Libert, M., Domingo, D., Heinderyckx, F., & Le Cam, F. (2014). Waiting for Data Journalism. *Digital Journalism*, *3*(3), 432-446. https://doi.org/10.1080/21670811.2014.976415

Felle, T. (2015). Digital watchdogs? Data reporting and the news media's traditional 'fourth estate' function. *Journalism*, *17*(1), 85-96. https://doi.org/10.1177/1464884915593246

Fink, K., & Anderson, C. (2014). Data Journalism in the United States. *Journalism Studies*, *16*(4), 467-481. https://doi.org/10.1080/1461670x.2014.939852

Friendly, M. (2008). A Brief History of Data Visualization. In (Vol. NA, pp. 15-56). https://doi.org/10.1007/978-3-540-33037-0_2

Heravi, B. R. (2018). 3WS of Data Journalism Education: What, Where and Who? *Journalism Practice*, *13*(3), 349-366. https://doi.org/10.1080/17512786.2018.1463167

Heravi, B. R., & Lorenz, M. (2020). Data Journalism Practices Globally: Skills, Education, Opportunities, and Values. *Journalism and Media*, *1*(1), 26-40. https://doi.org/10.3390/journalmedia1010003

Hermida, A., & Young, M. L. (2016). Finding the Data Unicorn. *Digital Journalism*, *5*(2), 159-176. https://doi.org/10.1080/21670811.2016.1162663

Hermida, A., & Young, M. L. (2019). *Data Journalism and the Regeneration of News* (Vol. NA). https://doi.org/10.4324/9781315163895

Houston, B. (2019). Data Journalism. *The International Encyclopedia of Journalism Studies*, *NA*(NA), 1-9. https://doi.org/10.1002/9781118841570.iejs0119

Knight, M. (2015). Data journalism in the UK: a preliminary analysis of form and content. *Journal of Media Practice*, *16*(1), 55-72. https://doi.org/10.1080/14682753.2015.1015801

Ksiazek, T. B. (2011). A Network Analytic Approach to Understanding Cross-Platform Audience Behavior. *Journal of Media Economics*, *24*(4),

Doi: 10.70008/jmldeds.v1i01.39

237-251. https://doi.org/10.1080/08997764.2011.626985

Lawrence, R. G., Radcliffe, D., & Schmidt, T. R. (2017). Practicing Engagement: Participatory journalism in the Web 2.0 era. *Journalism Practice*, *12*(10), 1220-1240. https://doi.org/10.1080/17512786.2017.1391712

Lazer, D., Pentland, A., Adamic, L. A., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N. A., Contractor, N., Fowler, J. H., Gutmann, M. P., Jebara, T., King, G., Macy, M. W., Roy, D., & Van Alstyne, M. (2009). Computational Social Science. *Science (New York, N.Y.)*, *323*(5915), 721-723. https://doi.org/10.1126/science.1167742

Lesage, F., & Hackett, R. A. (2014). Between Objectivity and Openness—The Mediality of Data for Journalism. *Media and Communication*, *2*(2), 42-54. https://doi.org/10.17645/mac.v2i2.128

Lewis, N. P., & Al Nashmi, E. (2019). Data Journalism in the Arab Region: Role Conflict Exposed. *Digital Journalism*, *7*(9), 1200-1214. https://doi.org/10.1080/21670811.2019.1617041

Lewis, N. P., & Waters, S. (2017). Data Journalism and the Challenge of Shoe-Leather Epistemologies. *Digital Journalism*, *6*(6), 719-736. https://doi.org/10.1080/21670811.2017.1377093

Lewis, S. C., & Usher, N. (2014). Code, Collaboration, And The Future Of Journalism. *Digital Journalism*, *2*(3), 383-393. https://doi.org/10.1080/21670811.2014.895504

Lewis, S. C., & Westlund, O. (2014). Big Data and Journalism: Epistemology, Expertise, Economics, and Ethics. *Digital Journalism*, *3*(3), 447-466. https://doi.org/10.1080/21670811.2014.976418

Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2017). Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016. *Journalism*, *21*(9), 1246-1263. https://doi.org/10.1177/1464884917735691

Lowrey, W., & Hou, J. (2018). All forest, no trees? Data journalism and the construction of abstract categories. *Journalism*, *22*(1), 35-51. https://doi.org/10.1177/1464884918767577

Martin, J. A., Camaj, L., & Lanosga, G. (2024). Audience engagement in data-driven journalism: Patterns in participatory practices across 34 countries. *Journalism*, *25*(7), 1578-1596. https://doi.org/10.1177/14648849241230414

Md Delwar, H., Md Hamidur, R., & Nur Mohammad, A. (2024). Artificial Intelligence and Machine Learning Enhance Robot Decision-Making Adaptability And Learning Capabilities Across Various Domains. *International Journal of Science and Engineering*, *1*(03), 14-27. https://doi.org/10.62304/ijse.v1i3.161

Milan, S., & Treré, E. (2020). The Rise of the Data Poor: The COVID-19 Pandemic Seen From the Margins. *Social Media + Society*, *6*(3), 2056305120948233-2056305120948233. https://doi.org/10.1177/2056305120948233

Mosleuzzaman, M., Shamsuzzaman, H. M., & Hussain, M. D. (2024). Engineering Challenges and Solutions In Smart Grid Integration With Electric Vehicles. *Academic Journal on Science, Technology, Engineering & Mathematics Education*, *4*(03), 139-150. https://doi.org/10.69593/ajsteme.v4i03.102

Mutsvairo, B., Bebawi, S., & Borges-Rey, E. (2019). *Data Journalism in the Global South* (Vol. NA). https://doi.org/10.1007/978-3-030-25177-2

Nandi, A., Emon, M. M. H., Azad, M. A., Shamsuzzaman, H. M., & Md Mahfuzur Rahman, E. (2024). Developing An Extruder Machine Operating System Through PLC Programming with HMI Design to Enhance Machine Output and Overall Equipment Effectiveness (OEE). *International Journal of Science and Engineering*, *1*(03), 1-13. https://doi.org/10.62304/ijse.v1i3.157

Ojo, A., & Heravi, B. R. (2017). Patterns in Award Winning Data Storytelling. *Digital Journalism*, *6*(6), 693-718. https://doi.org/10.1080/21670811.2017.1403291

Palmer, M. B. (2003). Parisian newsrooms in the late nineteenth century: how to enter from the agency back office, or inventing news journalism in France. *Journalism Studies*, *4*(4), 479-487. https://doi.org/10.1080/1461670032000136578

Palomo, B., Teruel, L., & Blanco-Castilla, E. (2019). Data Journalism Projects Based on User-Generated Content. How La Nacion Data

Transforms Active Audience into Staff. *Digital Journalism*, *7*(9), 1270-1288. https://doi.org/10.1080/21670811.2019.1626257

Parasie, S. (2019). Data journalism and the promise of transparency. In (Vol. NA, pp. 263-272). https://doi.org/10.4324/9781351173001-27

Parasie, S., & Dagiral, É. (2012). Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago. *New Media & Society*, *15*(6), 853-871. https://doi.org/10.1177/1461444812463345

Peng, T.-Q., Liu, M., Wu, Y., & Liu, S. (2016). Follower-Followee Network, Communication Networks, and Vote Agreement of the U.S. Members of Congress. *Communication Research*, *43*(7), 996-1024. https://doi.org/10.1177/0093650214559601

Pereira, F. H., & Mastrella, B. (2022). The Practice of Data-Driven Journalism in Brazil: Between Disruption and the Reinstatement of Professional Values. *Digital Journalism*, *10*(4), 599-625. https://doi.org/10.1080/21670811.2022.2061550

Poell, T., & Rajagopalan, S. (2015). Connecting Activists and Journalists: Twitter communication in the aftermath of the 2012 Delhi rape. *Journalism Studies*, *16*(5), 719-733. https://doi.org/10.1080/1461670x.2015.1054182

Porlezza, C., & Splendore, S. (2019). From Open Journalism to Closed Data: Data Journalism in Italy. *Digital Journalism*, *7*(9), 1230-1252. https://doi.org/10.1080/21670811.2019.1657778

Radchenko, I., & Sakoyan, A. (2014). *AIST - The View on Open Data and Data Journalism: Cases, Educational Resources and Current Trends* (Vol. NA). https://doi.org/10.1007/978-3-319-12580-0_4

Rahman, M. M. (2024). Systematic Review of Business Intelligence and Analytics Capabilities in Healthcare Using PRISMA. *International Journal of Health and Medical,*, *1*(4), 34-48. https://doi.org/10.62304/ijhm.v1i04.207

Ruellan, D. (2017). A Professional. Or How to Recognize One. *Brazilian journalism research*, *13*(1), 6-19. https://doi.org/10.25200/bjr.v13n1.2017.978

Sah , B. P., Shirin, B., Minhazur Rahman, B., & Shahjalal, M. (2024). The Role of AI In Promoting Sustainability Within the Manufacturing Supply Chain Achieving Lean And Green Objectives. *Academic Journal on Business Administration, Innovation & Sustainability*, *4*(3), 79-93. https://doi.org/10.69593/ajbais.v4i3.97

Schwabish, J. A. (2014). An Economist's Guide to Visualizing Data. *Journal of Economic Perspectives*, *28*(1), 209-234. https://doi.org/10.1257/jep.28.1.209

Shamim, M. (2022). The Digital Leadership on Project Management in the Emerging Digital Era. *Global Mainstream Journal of Business, Economics, Development & Project Management*, *1*(1), 1-14.

Sikder, M. A., Begum, S., Bhuiyan, M. R., Princewill, F. A., & Li, Y. (2024). Effect of Variable Cordless Stick Vacuum Weights on Discomfort in Different Body Parts During Floor Vacuuming Task. *Physical Ergonomics and Human Factors*, *44*.

Stalph, F. (2017). Classifying Data Journalism. *Journalism Practice*, *12*(10), 1332-1350. https://doi.org/10.1080/17512786.2017.1386583

Stalph, F., & Borges-Rey, E. (2018). Data Journalism Sustainability: An outlook on the future of data-driven reporting. *Digital Journalism*, *6*(8), 1078-1089. https://doi.org/10.1080/21670811.2018.1503060

Tandoc, E. C., & Oh, S.-K. (2015). Small Departures, Big Continuities? *Journalism Studies*, *18*(8), 997-1015. https://doi.org/10.1080/1461670x.2015.1104260

Uddin, M. M., Auyon, M. O. S., Al Adnan, A., & Akter, F. (2024). Strategies for Information Systems Development: Analyzing Requirements Determination and Project Selection. *International Journal for Multidisciplinary Research*, *6*(2). www.ijfmr.com

Uddin, M. M., Ullah, R., & Moniruzzaman, M. (2024). Data Visualization in Annual Reports–Impacting Investment Decisions. *International*

Doi: 10.70008/jmldeds.v1i01.39

Journal for Multidisciplinary Research, *6*(5). https://doi.org/10.36948/ijfmr

Usher, N. (2016). *Interactive Journalism* (Vol. NA). https://doi.org/10.5406/illinois/9780252040511.001.0001

Weber, W., Engebretsen, M., & Kennedy, H. (2018). Data stories. Rethinking journalistic storytelling in the context of data journalism. *Studies in Communication Sciences*, *18*(1), NA-NA. https://doi.org/10.24434/j.scoms.2018.01.013

Webster, J. G., & Ksiazek, T. B. (2012). The Dynamics of Audience Fragmentation: Public Attention in an Age of Digital Media. *Journal of Communication*, *62*(1), 39-56. https://doi.org/10.1111/j.1460-2466.2011.01616.x

Weeks, B. E., & Holbert, R. L. (2013). Predicting Dissemination of News Content in Social Media: A Focus on Reception, Friending, and Partisanship. *Journalism & Mass Communication Quarterly*, *90*(2), 212-232. https://doi.org/10.1177/1077699013482906

Westlund, O., & Hermida, A. (2021). Data Journalism and Misinformation. In (Vol. NA, pp. 142-150). https://doi.org/10.4324/9781003004431-16

Wong, F. M. F., Tan, C. W., Sen, S., & Chiang, M. (2016). Quantifying Political Leaning from Tweets, Retweets, and Retweeters. *IEEE Transactions on Knowledge and Data Engineering*, *28*(8), 2158-2172. https://doi.org/10.1109/tkde.2016.2553667

Zhang, X., & Ho, J. (2020). Exploring the Fragmentation of the Representation of Data-Driven Journalism in the Twittersphere: A Network Analytics Approach. *Social Science Computer Review*, *40*(1), 089443932090552-089443932090560. https://doi.org/10.1177/0894439320905522