Global Media Journal-Indian Edition; Volume 7 Issue 1; June 2025. ISSN:2249-5835

Investigating through AI: Transformative Cases that Revolutionised Indian Journalism

Dr Sweta Ghosh

Assistant Professor, Department of Mass Communication and Videography, Rabindra Bharati University, Kolkata
swetamit10@gmail.com

Introduction

Investigative reporting is an intellectual process. For a professional journalist, investigative reporting is the business of gathering and arranging basic concepts and realities, creating displays of facts, evaluating various possibilities available, and finally taking decisions based on logics avoiding emotional attachments. As a logical decision maker, an investigative journalist has to go through several stages. The fundamentally intellectual nature of this discipline demands many layers: at a certain point some reporters have a tip, which later surround them with some circumstantial evidences and leading them to some sensational stories (Williams, 1978, pp. 12–34). According to famous American Investigative journalist Seymour M. Hersh, the investigative journalism is "The hardest thing to do in any reporting." It is expected beyond sitting and around several stories, clipping and indexing them, and cultivating sources by travelling and telephoning.

An investigative journalist has to develop a special style of reasoning, she has to be trained in the subject of Law, and has to act like a researcher. She has to know about the ontological existence of an incidence, and then has to search out her epistemological position in investigation process. If she observes a phenomenon, an effect and wonders about its origin, she must develop a hypothesis and start checking it against the observable facts and figures. Like any communication researcher, an investigative reporter too, works to back up the chains of facts, searches various information that may support or reject her Hypotheses. According to the Environmental Justice Foundation and Tempo Institute (2024), an investigative report can be different from a regular news report and an in-depth report, as a regular news report only explains 5Ws (what, where, when, who, why) and in-depth report further adds an H (How) to the 5Ws, whereas an investigative report explains 5Ws and H along with the expansion to the these i.e. how it originated, how much extended and to what level, and who are involved.

Any investigated incident involves three basic elements. Usually an investigative report begins with a trivial issue, anything that is already known e.g. an unexpected financial report, any complain regarding public affair, or any incident related to sudden wealth increase of a certain government official. The work of investigation is totally an intellectual process, with no magical formula. The methods opted by an investigator forms a pattern, a workable plan, a roadmap along with several steps such as conceptions, feasibility study, go or not-to-go decision, planning and base planning, original research, revaluation, key interviews, final interviews, final go or not-to-go decision making and writing for publication (Williams, 1978, pp. 12–34).

Growth of AI

The process of AI is very complex one. The role of Artificial intelligence is to create intelligent machines that can perform fast and improved works than humans. AI not only about intelligence but it's about resolving difficulties in work. These solutions are generally calculated mathematically and scientifically through devices like computers. Until the 1950s, the term "Artificial Intelligence" was not so popular as it is today. The popularity of the term is obliged to John McCarthy (Father of Artificial Intelligence), who believed machines to can do reasoning and work logically like humans. The famous Dartmouth Summer Research Project on AI unleased the potentiality of the term in 1956.

AI has long history, dated back to 350 BCE, when famous Greece philosopher thinker Aristotle marked about logical and scientific thinking, which has been influencing many scientists and philosophers for years. Since then, the prime concern of researchers and scientists was to investigate the way human beings think. The way they search various avenues to help people for better and faster thinking processes (Harris,2011, pp 5-6). The existence of Scientific Thinking Theory is over 2000 years, which mainly concentrate in understanding and resolving two universal questions i.e. the way human minds work and presence of minds in non-living beings. According to Essential English Dictionary (1990), the term intelligence is defined as the ability to comprehend and acquire things based on rational thinking. Intelligence is the ability to think and understand instead of doing things by instinct or automatically (Collins, 1990). the initial definition decodes intelligence as the qualitative feature owned among humans.

The second definition caters the flexibility in the qualitative nature of the trait. In the broader aspect, it does not provide specific criteria in providing someone or something's ability to think and comprehend (Negnevitsky,2005, pp1-3). One of the early and remarkable achievements of AI is 'ELIZA Speaks'.This famous computer program, was invented in 1966 by a professor of computer science at MIT named, Joseph Weizenbaum. ELIZA was programmed to respond to various basic information, which were entered into the computer. The success of this primary AI program had encouraged many scientists to create more complicated forms of artificial intelligence (Harris,2011, pp 6-8). IBM's Watson has the ability to analyse the unstructured data that is useful for journalism, especially investigative reporting. It can process massive data, generate hypotheses, and can suggest potential news stories. before pursuing a story, a reporter can review all AI-generated leads for data verification. The International Consortium of Investigative Journalists (ICIJ) took 15 months to analyse 2.5 million offshore information totalled more than 260 gigabytes. Watson could automate these offshore data in much lesser time, while extracting relevant articles, mapping relationships, and generating investigative leads quickly (Diakopoulos, 2013).

Research Objectives

1. To investigate specific case studies where AI has played a vital role in investigative journalism
2. To analyse the real-life impact of AI in media both at the Indian and Global scenario.

Literature Review

Ongsulee (2017, pp. 1-6) has defined AI as the scientific and engineering outcome of of creating intelligent machines, particularly in the intelligence form computer programs. It works in association of computer science and deals with the development of algorithm, software, and peripheral technologies, which is created to stimulate the intelligent behaviour in robots and other electronic devices. Many modern and scientific technologies such as robotics, machine learning, computer vision, Artificial neural networks, and Natural Language Processing (NLP), all comes under the vast umbrella term i.e. Artificial Intelligence. The report of Networking & Information Technology Research and Development Subcommittee & Machine Learning & Artificial Intelligence Subcommittee highlights various opportunities and challenges faced by AI and cybersecurity. The term AI often used synonymously and unified way to another computer system term i.e. Machine Learning (ML). ML is the specialised branch of AI using various algorithm models such statistical machine, empirical and reinforcement learning. Today, Technology is at an inflection point, both AI and ML are advancing faster than society's ability for AI absorption and comprehension. The computer systems employing AI are becoming both critical and ubiquitous. AI technologies become pervasive, both humans and machine are working together efficiently and accurately for every critical works like diagnostic, teaching, or reporting (2020, pp 1-9).
According to Nandini, Indora, and Singh (2024), AI has created a deep penetration in our media system. Media houses are employing AI for critical tasks like automated artificial generation, personalised content recommendation and sentiment analysis. The long tiring routines of both mass media houses and news agencies are well-managed by the AI as it is assisting in news gathering, collecting and distributing process. Usage of Generative AI in Journalism, Deepfakes, Automated Journalism and Robot Reporters, Blockchain and Cloud Journalism, Open Data Journalism, Big Data Journalism, and Aerial Journalism are few terms that emerged due to application of AI in Journalism. Both the Strategic Business model and ethical works too have been changed by AI newsroom (pp.237-242).

Algorithm Journalism

is the term coined by Dörr (2015) where well organised data is transformed into text through Natural Language Generation (NLG). With the growth of Algorithm, critical tasks like editing, aggregating, publishing, and distributing content are being automated. As the media houses especially, the prints are facing challenges such as productivity, profitability and market share, NLG is becoming extremely relevant as an important part of NLP in print production. NLG software can automatically produce human-like language from data. Some tools like Narrative Science or Automated Insights can add multimedia components to the generated texts. This combination of NLG and multimedia enhances content diversity in Journalism (pp 700-722).

Generative AI in Journalism

is a technology driven opportunity that is permitting continuous content development. The generative content in journalism powered by AI and ML have created customization of content possibilities. Presently, AI comes up with a unique and incredible idea to provide a wide range of connected ideas. Small media outlets can offer automated AI customization and dynamic video content. ML is like an intelligent investigator who can monitor the mountain of data and create critical issues that requires necessity to address in Journalism (Naikodi, & Suresh, 2020). Generative AI is a special type of AI algorithms generating new data outputs on the basis of

data-generation training (Routley, 2023). It is an algorithm, which permits the creation of new content in response to prompts. It constructs answers based on queries provided by the users. This algorithm provides platforms to create and generate users' own photos, videos and audios for unique content creation and mimic human natural creativity (**Nandini, Indora, and Singh, 2024, P 240).** For Content Creation in Journalism, Generative AI systems are used as the AI writing tools, image, audio and video generation tool, and as the AI-driven social media bots. Some of the potential harm of this tool are deepfakes, misinformation and disinformation, copyright infringements, intrusion to privacy, impersonation, data protection violations, incitement of hate speech, harassing individuals though illegal and wrongful speech (Schertel Mendes, & Stray, 2024, P 62).

Automated Journalism and Robot Reporters

Robot reporters are the AI-powered automated news reporting system, who can handle repetitive news reporting tasks. A robot reporter can use AI for multi-tasking such as looking to market updates, real estate and property transactions, and sports scores, ensuring real-time fast reporting on various topics like weather, lifestyle, sports and business. AI humanoid robots help in improving news accuracy and consistency. These robots reduce human errors in heavy financial reports permitting human reporters to look into investigative and analytical reporting, leaving the large data on robot investigative reporters (Sjøvaag, 2024, pp 246-247).

Blockchain and Cloud Journalism

Some data of smart contract regulation in the investments are added in the algorithm form to the software. These algorithms reviews in the automatic and permanent contract forms, with all measures in consistency. These rules are stored in decentralised and manageable databases known as blockchain (Gentsch, 2019, P 47). Blockchain developed by Satoshi Nakamoto in 2008, aimed at designing the first reference implementation specification for a distributed database. It mainly consists of records known as blocks, containing a timestamp and linked to the previous block. In this system data is distributed across the entire network, eliminating the central point of failure. This decentralised technology structure makes the blockchain highly resistant to data corruption and cyber-attacks. (Skilton, & Hovsepian, 2018, pp 48-49). The Cloud Computing model provides easy, sharable and demandable access to networks along with parts of computing resources. Cloud as the popular resource based on the large-scale data analytics and a common resource for data science applications (Voulgaris and Bulut, 2018).

Both, Blockchain and Cloud technologies play an important role in journalism. As they both enhance data security, transparency, accessibility and trust. The blockchain security relies on public key cryptography, with a unique address. As there is no centralised copy, it is more credible (Skilton, & Hovsepian, 2018, pp 48-49). Journalists too can rely on these technologies for security in news reporting. They can protect the whistle-blowers and confidential sources. The end-to-end encryption prevents access to sensitive documents by any unauthorized individual or system. Readers can check the authenticity and sources of the investigative report. Along with blockchain technology, cloud computing too provides a flexible and accessible structural arrangement, which is beneficiary to journalists. The cloud-based platforms like Google Drive, Microsoft OneDrive help investigative reporters to store their large volumes of

data of research and analysis securely. Journalists can store multimedia files in the cloud and access them from anywhere.

Open Data Journalism

Data Journalism has reached to a new level with the Precision Journalism. Media news rooms are leveraging open data and computational tools for the enrichment in fact-based investigative storytelling and information transparency. The augmentation of computational system along with human brain has been effective story-generation. Data Journalism is at the heart of Social Science thinking. Precision Journalism methods are assisted by computer while involving in the process of reporting. This method flourished in the 1980s when desktop computers invaded in the newsrooms (Broussard, 2014, pp 9-10).

Big Data Journalism and Aerial Journalism

Big Data is the process enormous information extension such as government records, genetic sequences, internet users trace etc., available in various scientific, analytic, research, public policy making, or business domains (Parasie, 2015, p. 364). The media industry has transformed drastically by the utilization of AI and large-volume of data processing. In case of investigation that require analysis of big data. Big Data analysis techniques help in identification and contextualization of newsworthy events. The implementation of Big Data Journalism, has facilitated media and news organizations to improve investigative reporting, streamline data analysis, and combat misinformation, ensuring a more knowledgeable and mindful society (Silva, 2023, P264).

Davidson (2024) states that AI as the super-smart assistant is helping Journalism in every possible way. AI can generate content such as headlines and articles, by analysing patterns in data. It can facilitate a reporter working on an investigative story i.e. trying to uncover financial scandal, by fact-checking and verifying thousands of documents and identifying patterns that are either missed manually. In the age of misinformation, AI algorithms can assist in identifying fake news and verifying facts. It can optimize workflows by aiding in tasks like automation, transcription, proofreading and scheduling. Investigative journalists can investigate companies' financial crimes, analyse financial records and transactions.

Goni and Tabassum (2020) investigates the preparedness and readiness of journalism students in Bangladesh towards the adaptation of AI in the journalistic works. For the theoretical framework the study used Technology Acceptance Model (TAM) (pp. 209-228). Davis (1986, as cited in Chuttur, 2009) has proposed Technology Acceptance Model (TAM) in his PhD thesis, which included Five variables such as Perceived Ease of Use, Perceived Usefulness, Attitude, Intention, and Actual Use, which directly influenced the external stimulus consisting of system's actual attributions and competencies. The extended TAM 2 model of Venkatesh and Davis (2000, pp 186-204) highlighted the rational and cognitive processes such as significance in work, qualitative output, certainty in results, and perceived ease in usage had significantly influenced the acceptance process of the users. The qualitative study conducted by Goni and Tabassum (2020, pp. 209-228) studied 120 representatives from the field Mass Communication and Journalism to identify the way they interact using AI, their voluntarily usage, and about their view regarding Future of Journalism in presence of AI. AI is playing a great role in the detection

of fake news, searching automated content for journalistic production, creating personalised content, and enhancing their improved search capabilities.

Methodology

This study devotes using qualitative case study method to explore the significant role of AI in enriching Investigative Journalism in India. Each case i.e. the usual unit of inquiry has permitted for an extensive and "in-depth" description of the social phenomenon e.g. the way AI tools and techniques have been integrated into journalistic practices. As the Research Questions (RQs) seek to explain current situation (e.g., "how" or "why" AI functions for Investigative Journalism), the case study inquiry becomes more relevant (Yin, 2018).

Research Design

A case research design may be identified with single or multiple, holistic or embedded cases design. The present case study uses multiple cases, along with intrinsic case study approach, as the cases are pre-selected (Stake, 1995, P4). Yin (2018) constituent five elements of a case research design: (a) A case study's questions; (b)Its propositions, if any; (c) Its case(s); (d) The logic linking the data to the propositions; and (e)The criteria for interpreting the findings. The initial task of case study is to clarify precisely the "who," "what," "where," "how," and "why" questions of the study. Based the Research objectives, following Research questions have been created.

RO 1: To investigate specific case studies where AI has played a vital role in investigative journalism.

RQ1: In what ways has AI been integrated into investigative journalism practices in India and abroad?

RQ2: What specific AI tools and techniques have been employed in the detection of various AI-generated fake contents?

RQ3: How has the use of AI affected the investigation outcomes in terms of efficacy, accuracy, and depth?

Case Selection Criteria

Following are few transformative cases that has revolutionized the Indian Investigative Journalism by using AI. These cases have been selected on the basis of their Diversification, significance and AI Integration.

RO 1: To investigate specific case studies where AI has played an essential role in investigative journalism.

The AI-driven systems possess both creativity potential along with the editorial functioning, promoting innovation in Journalism. AI has undertaken a critical role in editorial construction, equability and fact-checking analysis (Leiser, 2022). Computer Vision (CV) is the subset of AI. By using mathematical algorithm, CV allows computer to develop meaningful information through digital images. Image recognition and machine vision are the two strands of CV algorithm. In Ukraine, small media house like Texty.org.ua has used CV models for the detection

of land being transformed into lunar-like structure as a result of illegal amber mining (de-Lima-Santos, & Ceron, 2022, pp 21-23). Texty.org has own several media accreditations for combating disinformation and indulging in innovative investigative journalism practices. It has been awarded with European Press Prize (2024) with the best Innovation Award nominee for its project named "The Stolen Treasures." In 2020, it has gained Sigma Award for the project "Fresh Disinfo from Russia." For its famous "Illegal Amber Mining" project, it has awarded with a Bronze medal in SND Digital Award (2018) (Texty.org.ua, n.d. para. 2). Texty.org has always been recognised for its innovative story telling in investigative journalism. It has used data analysis through technological detailing and computer-based visualization to track the plundered cultural resources and smuggling routes. It also emphasised on cross-border collaborative investigative journalism and developed computer-based interactive tool for tracking the fake information in project Fresh Disinfo from Russia and Fake Cracker. For credibility in investigation, Texty.org has used in-depth data and promoted data-driven journalism. AI played a great role in the identification and tracking of misinformation patterns, especially in the "Fake Cracker" project. In "The Stolen Treasures" investigation was done by using large geospatial datasets for visualising the illegal mining activities.

Various Traditional News Establishments such as The New York Times (NYT) has utilised Applied CV algorithms for the estimating the Three-dimensional poses of sportspersons at live shows. Distinguished news agency Reuters, too has used CV and has tracked down the urban expansion of South China Sea through satellite images. It permits integration of CV for enhancing the search feature and archiving important videos for future references. Xinhua, the Chinese news agency has integrated CV with other AI technologies for the development of Man-Machine collaboration in newsrooms. For the fake image detection, several media houses are using CV for searching out the manipulation in images. Source App developed by Storyful is powered by Google AI, which provides access to an image's public history for the verification of its authenticity and detection of manipulation (de-Lima-Santos, & Ceron, 2022, pp 21-23).

RQ1: Integration of AI with Investigation in India and Abroad

Case 1: The Pegasus Project (2021)

Pegasus is a sophisticated spyware (Trojan/Script) was first discovered in 2016, when in January a famous investigative journalist named Carmen Aristegui, received suspicious messages with links after she published an investigation report regarding a property owned by former Mexican President Enrique Pena Nieto (Chawla, 2021). Pegasus developed by the Israel based cyber-based company, can be installed easily in any iOS and Andrid based mobile devices (Hussien, Butt, & Bin Sulaiman, 2023). It was: capable of reading messages, emails, and browser history; enable to listen calls and accessing encrypted audio/text files; able to track GPS location, extract passwords and record audio/video, track GPS location, and extract passwords; monitor multiple applications like Facebook, Skype, WhatsApp, Viber, and BBM (Chawla, 2021). The universal fame of Pegasus came into existence in August 2016 when an Emirati human rights lawyer, Ahmed Mansoor found himself being targeted after receiving a series of suspicious text messages promising "new secrets" to be disclosed about sufferings of United Arab Emirates prisoners. After recognising capability of the cyber threat Mansoor forwarded the message to the Citizen Lab experts working at the University of Toronto. Laboratory investigation revealed it as a highly advanced cyber hacking tool. Citizen Lab tracked more than 1000 matching IP addresses linked to Pegasus, developed by the NSO Group and between 2016-2018, Pegasus was

potentially being used for surveillance operations across 45 countries, many of which were governed by authoritarian regimes (Marzocchi, & Mazzini, 2022).
A collaborative investigation named "Pegasus Project" was conducted by, Amnesty International and 16 international media organizations like *The Guardian (UK), The Washington Post (USA), Le Monde (France), Süddeutsche Zeitung (Germany), Die Zeit (Germany), The Wire (India), Haaretz (Israel), Radio France (France), Daraj (Lebanon), Proyectos de Periodismo (Mexico), Knack (Belgium), Le Soir (Belgium), Forbidden Stories (International coordination), Frontline (Hungary), The Organized Crime and Corruption Reporting Project (OCCRP), El País (Spain)*. This extraordinary media consortium investigated the use of Pegasus spyware. This investigative project mainly relied on the traditional investigative methods, technical analysis and human intelligence rather than on AI. The project used techniques like formation of a small investigative team of dedicated reporters, using caller detection apps like TrueCaller and CallApp, checking all profiles on WhatsApp and other social media platforms, searching all legal and government documents manually (Varadarajan, 2021). Amnesty International's Security Lab used both technical and artificial intelligence for analysing large datasets, identifying patterns, and detecting anomalies that facilitated the exposure of surveillance activities. AI enhanced Amnesty International's Security Lab's forensic analysis capabilities as it fastened the detection and monitoring process. In June 2023, the renewed version of Pegasus spyware threats was first detected by the Lab. AI aided forensic analysis also successfully recovered evidence of a zero-click exploit sent via iMessage to investigative journalist and Regional Editor of *South Asia*, Anand Mangnale's mobile (Amnesty International, 2023). The Wire used AI-driven forensic tools like Amnesty International's Mobile Verification Toolkit (MVT), which was developed in July 2021 for the identification of infected devices for the confirmation of surveillance claims (Kaldani, & Prokopets, 2022, P 9).
Pegasus Project gave birth to Collaborative Journalism driven by AI technologies. The collaborative journalism is the explosion of high-profile data journalism bringing together a wide range of news publishers across the territorial boundaries (Clements, 2018, P 78). With the development of Journalism, media research, and technology, each and every journalist in the collaborative team started understanding the role of single individual in the journalistic team operating transnationally and promoting cross-border journalism. The cross-border journalism includes four features: (1) Journalists belonging to different nations, who (2) unite to work on a theme or story (3) They compile, validate and merge all mutually gathered outcomes as (4) an individually fact-check data and ultimately publish these outcomes to their local, regional, national, local or any group of specialization(Alfter, 2018, pp 41-43). Rise of AI and Algorithm Journalism have facilitated many transnational collaborations of Indian and Foreign journalists and organizations to opt for AI-assisted investigative journalism.

Case 2: Electoral Bonds Investigation (2018)

The Founding Editor of 'The Reporters' Collective Nitin Sethi, who is famous for investigative journalism focused on political accountability, has used AI-driven data analysis technique to investigate documents related to electoral bonds. The case regarding Electoral bonds started in 2019, when Commodore Lokesh Batra, a retired navy officer, delivered a collection of documents to Sethi, which he acquired through multiple requests based on Right to Information (RTI) Act, 2005. This cache of documents exposed noteworthy indiscretions in a newly introduced scheme facilitated anonymous political donations (Dore, 2024).

In 2017, the concept "electoral bonds" was introduced in the Lok Sabha by the then Finance Minister Arun Jaitley as the Financial Bill 2017. On 29th of January, 2018, NDA government led by P.M Narendra Modi notified about the Electoral Bond Scheme, 2018, as a scheme to create an alternative to the cash donations, ensuring the transparency in political funding (Vachhani & Rana, n.d.), by reducing black money circulation during the election process while preserving the voters' anonymity (Mittal, & Agarwal, 2021, pp 21-22).

Nitin Sethi's investigation produced six-part investigative series, exposing illegal bond sales before elections, and uncovering various government misrepresentations regarding the bond sales. It also revealed the way the instructions of the Reserve Bank of India had been disregarded in endorsing this controversial scheme. Initially Huffington Post India published this report. Later numerous investigative reports were published through The Reporters' Collective. Sethi worked overnight with his team analysing the new data and produced over dozen investigative stories within 48 hours, identifying connections between top fund donors and major conglomerates. Stories revealed many renowned companies donating more than their declared profits. The prime investigative approaches were: reporters conducting thorough examination of the public documents and; using Right to Information (RTI) requests as a key information gathering tool. The main areas of investigations were mismanagement in Political funding, Government accountability, Coal auction irregularities, Environmental violations and COVID malpractices. This investigative project too led to collaborative journalism as many media outlets such as The Quint, Scroll, Newslaundry and The News Minute collaborated with the project with other regional media houses for wider reach and impact. The Major part of the investigations revealed that Companies under investigation by law enforcement donated large amount to the ruling party, false and deceptive pharmaceutical companies bought electoral bonds and Pro-BJP social media web-pages spent significant money on misleading advertisements. Collaborative digital platforms with multi-room facilities were created for collaborative investigation (Dore, 2024). Along with traditional investigative journalism approach, the teams used digital algorithm and computer-driven tools for data analysis and verification of electoral bonds information, and cross-referencing of public records along with organizational data.

Application of AI is not new in investigative journalism. Previously large datasets like Panama Papers with 13.5 million documents or the ICIJ's data set with 2.5 million documents involving massive offshore holdings and huge accounts (over 100,000 bodies) were analysed for the identification of the relationships between various organizations and renowned individuals. AI has enhanced journalistic efforts by handling vast amounts of data, permitting reporters to focus on deeper investigative work. It not only helps in countering any misinformation by guiding reporters to reliable sources, but also helps in detecting patterns and anomalies. Modern AI techniques and its usage in journalism especially investigative journalism have preoccupied various models of Ethical AI. Old concepts like 'accountability,' 'algorithmic transparency,' 'fairness in ML' and 'explainability' are invading in the modern journalistic arenas especially in investigative genres, as they require thorough, detailed and credible analysis (Leiser, 2022).

Case 3: Deepfake Detection During Elections (2024)

The 2024, Indian General Election has witnessed a massive upsurge in the generation of AI-based deepfake videos of various political figures including PM Narendra Modi and other opposition leaders like Rahul Gandhi, and Mamata Banerjee. Approximately $16 billion in total were spent in the election by various political parties. AI was leveraged to produce highly convincing but fake videos, such as one showing Prime Minister Modi pirouetting on a

Bollywood dance number (Anadi, 2024) or resurrection of a deceased leader Muthuvel Karunanidhi, who passed away in 2018, endorsing for the leadership of his son, Muthuvel Karunanidhi (MK) Stalin (Kohlman, n.d.). Coined as India's first AI election several manipulative videos appeared during 2024 election.

a) **Fake Videos of Modi's Aides:** Two manipulative clipping of PM Modi leading police investigation for the arrest of Congress party workers.

b) **AI-Generated Fake Voices:** Several deepfake voices of top leaders appeared as a part of conspiracy to create tension prior to elections.

c) **Bollywood Celebrity Deepfakes:** Fake videos of many Bollywood celebrities such as Anil Kapoor, Aamir Khan and Ranveer Singh were seen criticizing the Prime Minister.

d) **Amit Shah Deepfake:** A doctored clip falsely showed Amit Shah, the Central Home Minister claiming sojourn on certain minority rights by the BJP government. Such videos are highly sensitive for a multicultural democratic country like India.

e) **Yogi Adityanath Deepfake:** An AI-operated video falsely showed Uttar Pradesh Chief Minister Yogi Adityanath criticizing Modi for not supporting the families of those who died in a 2019 militant attack (The Hindu, 2024).

f) **Resurrection of Dead Politicians through Deepfakes:** Apart from M. Karunanidhi, another deceased political person J. Jayalalithaa (died in 2016) was resurrected through AI. The Communist Party of India (Marxist) in West Bengal, too had released an AI-generated video of former Chief Minister Buddhadeb Bhattacharjee, urging to citizens to "save the country and the state" (Sakunia, 2024).

Deepfakes are the consequences of the AI and advancement of computer science. Deepfake videos are the new, innovative and highly persuasive method of disinformation dissemination. These videos and footages are reformed synthetically in the portrayed face or physical features. It is modified artificially on the subject to depict someone or something else. For any democratic nation like USA, India or any extensive international community, AI generated contents and Deepfakes are at very vulnerable period of time (Helmus, 2022). Kavanagh & Rich (2018), identify four trends that eventually decreasing the importance of truth in American society: increase in disagreements in facts and data evaluation; an obscuring line between opinion and fact; a bulk increase in opinion and consequential influence of such bulk on facts and; a declination in trusting sources, which were considered trustworthy previously. With the changes in technology and growth of highly susceptible audiences, synthetic videos are increasingly lifelike, and increasingly became threat to the both domestic and international disinformation. Apart from deepfake videos, generative texts and voice cloning are growing concern for political scientists and research analyst.

Waldemarsson (2020, P 5) states that the Foreign election interference is not new. Since the days of Gutenberg, disinformation has been used to manipulate public opinion. But the issue regarding Election and fake information upsurge especially after Russia's intervention in the Presidential Election of United States in 2016. Since then, many global incidents regarding public opinion and disinformation emerged. In 2020, reports regarding Chinese disinformation campaign against a China-critical candidate in the Taiwanese presidential election appeared. In

2019, there was a suspicion Russian attempts that came into sight to influence UK election and in the same year several proofs regarding anti-democracy campaigns by several fake Chinese accounts came into notice in Hong Kong. Waldemarsson (2020) also warned by the year 2022, preponderance number of people will consider false information as authentic one. The false information, defined as Disinformation, intentionally spread to deceive public and it is becoming a new normal in the society especially in the political arena.

Deepfake videos can manipulate elections. For instance, any manipulated video can potentially influence the election's outcome, if suddenly on the election day eve, a scandalous video involving any candidate emerges on the online platform (Helmus, 2022). During the election of 2024, India, with over 968 million voters has witnessed biggest massive AI-based election contents. Various political parties reported spent about $50 million on different AI-generated election content, defining the Election 2024 as the AI-Election (Sakunia, 2024). Election campaigners used AI-generated various Hindi and Punjabi dance songs in Prime Minister's voice, blending politainment content. This blending of political and entertainment content, entangled the political celebrities, issues, and practices with an entertaining culture. This hybrid content engages entertainment with political reality to enhance the images of politicians along with promotion of certain issue through media access (Riegert and Collins, 2015). Various Memes and political satire based social media communities also leveraged generative AI tools to produce content related to Indian politics. Though PM Modi marked his deep concern about the misuse of AI in election, still he endorsed the trend by retweeting a meme, where he himself was dancing to a popular Bengali track 'Paglu Dance', for expressing his hilarity at the creativeness. The election 2024 witnessed a massive surge in Politainment or more minutely one can say "Electertainment (Election+Entertainment) utilising AI technologies. Still there is a concern for AI-based misinformed political content in the country.

RQ2: Usage of specific AI tools and techniques in detection of various AI-generated fake content

There are multiple investigative tools and techniques available for the investigative journalist to dig out the fake AI generated content during the election. As on 2024, more than 150 countries were conducting elections and since all these nations are ruled by different civic freedoms and regulations, no single technique could aid them all. Journalists all over the global for their investigation formed small groups to work on the basis of different fact digging tools and methods. Following are the examples of different fact-digging techniques for investigations that do not require any advanced computer knowledge:

1. Boolean Google search helps investigator reporters to concentrate on Google's data-finding power, by in combining the reputable search operators, or by using advanced Google "dorks."
2. The "excluding trick"; One can detect disinformation by using "-site:" technique such as: "-site:youtube.com" or "-site:Instagram.com" to track the origin of political content across platforms.
3. Simple Copy-Paste Technique: This technique is used to search text, which can reveal hidden influence campaigns by finding similar content across websites. It is the easiest way to detect any election-associated site with an arbitrary or flammable content, with no information regarding the owner. The investigator can just copy any chunk of text from the homepage to Google search bar to detect any similarity of the

content with other websites. This search also suggests hidden content, logos and layouts related to the arbitrary information.

4. Using the "Geocode" Trick to Track Campaign Incidents: an investigative reporter can track any suspicious political event by using X Pro. Previously, one could gather automatically social media posts from specific location by using TweetDeck. But currently, ad tweeter changed to X, one requires a paid subscription to access TweetDeck (now called X Pro). X Pro is a powerful tool for monitoring real time political raids and rallies. An investigative reporter can easily use it to identify the culprits and can debunk false allegation blaming the victims.
5. Hootsuite: Works similar to X tools for real time monitoring of political incidents. It can be used for location based surveillance and aid in debunking false claims about political violence.
6. Junkipedia: With the declining trust of the reporters about social media like Facebook, Meta's most public insight powerful tool for social content exploration named CrowdTangle was shut down on 14th August 2024 and a new versatile information digging tool of Algorithmic Transparency Institute

named Junkipedia emerged to capture the social media content. Junkipedia permits users to form social media accounts from 12 different applications, along with the mainstream platforms like Facebook, Telegram, TikTok, and alternative platforms like Gab and GETTR, for tracking election related disinformation.

7. Psiphon: It is an innovative tool for reaching out the oppressed and distressed citizens. This sophisticated tool of Canada enables voters' accessibility to free information about their rights and responsibilities along with all data about the governance without any risk or menace.
8. **Bellingcat Auto Archiver: It is a useful tool for preserving video evidences. It permits investigators simply store and copy-pasting videos related to election-related violence. The system of Bellingcat automatically detect the best method for downloading and archiving each video clip for total security** (Philp, 2024).
9. Aleph database: It includes a "cross-reference" feature to search hundreds of databases for investigating financial irregularities and searching out the hidden connections. Various financial investigators have successfully searched into different election-related financial scams and global indiscretions with Aleph document leaks database. The Organized Crime and Corruption Reporting Project's (OCCRP), data platform named Aleph, aids in investigative journalism permitting journalists to conduct search, organize, and analyse information efficiently. It simplifies complicated investigations process especially in "monetary scams and malpractices" in large corporate sectors by searching out the connections in corruption (Strozyk, 2024).

RQ3: The way usage of AI has affected various investigation outcomes in terms of efficacy, accuracy, and depth

Various media houses have witnessed massive transformation in their journalistic practices due to proliferation of AI, and ML. These modern technologies have permitted the arrival of new investigative journalism era. Journalists are now capable of creating data driven contents by using a smart set of data gathering and analysis tool. The most versatile and investigative journalism tool named SociaLens is gaining popularity. This AI- powered Investigative Journalism tool aids in identification and extraction of specific data from online resources, for responding to various inquisitive information in order to draw conclusions entailed by large data

using ML analytics autonomously. SociaLens is an intelligent chatbot, which is empowered with ML that enables autonomous gathering of data, identification of features, generation of model, prediction of information and visualization of data. It incorporates situational awareness and cognitive-based forecasting to provide automated analysis from multi-modal data sources. This tool basically uses a user-accessible graphical interface for the transformation of unstructured test into structure data, promoting easy fact extraction. With the combination of stored-data with inferences, SociaLens supports contextual conversational interrogation, relying on real-time prediction of events. Jamil & Rubaiat (2024) have conducted study on the way SociaLens has been applied on a social science research on child rape incidents in a developing country like Bangladesh. Various incidents related to Child Abuse cases in Bangladesh are facing severe challenges. The poor democratic structure and political instability are constantly hindering the justice. There is a limited availability of reliable data regarding child abuse cases. The press freedom is nominal. The mainstream media often face threats and intimidations. Though digital media are only sources of reporting incidents, still digital journalists often pressurised and facing repression through media restriction laws (pp 1-15).

SociaLens is designed to combat journalistic pressure in data gathering process, analysis and reporting sensational issues like child abduction and abuses, rape, and human rights violations. Investigative journalists working in repressive and authoritarian regimes can use this tool for the aggregation and structuring of fragmented data collected from different and diverse sources such as news reports from local and small media, information from NGOs and online archives. It provides anonymised and encrypted investigative research tools for the reporters working under surveillance. It identifies emerging trends and pattern of abuses, to picturise and create a systematic storytelling process. The study had used data sources collected from Bangladesh's leading newspaper Prothom Alo, data set from NGORep and various archival materials available online. SociaLens combined 2,811 rape incidents recorded from 213 to 2023, various NGOs data available from 2001 to 2021, and all real-time trending news regarding child abuse and rape available in archival portals like NewsAPI & NewsData.io. Jamil & Rubaiat (2024) collected all data based on rape incidents and categorised the news based on 3 classifications of incidents such as gang rape, domestic abuse and harassment at the workplace.

Conclusion

According to famous Pulitzer Prize Awardee, James R. Polk, a reporter is responsible for the truth of every word in that investigative article, that allegation, any second-hand bit of knowledge is not a fact. Polk, who has received the prize for his National Reporting of the Watergate scandal in *The Washington Star*, also suggested that in investigative journalism there is no excuse to apologize for the mistakes of the sources. With the growth of AI (Artificial Intelligence) investigative journalism has changed its working process. AI has penetrated into the media system and several new terms such as Generative AI, Deepfakes, Automated Journalism, Robot Journalism, Blockchain and Cloud Journalism, Open Data Journalism, and Big Data Journalism, have emerged. Investigators can use SociaLens for creating geographical tagging to search out the patterns of criminal offences. It can leverage ChatGPT-4o for enquiring intent recognition, generating retrieval-augmented, and responses. This AI-powered tool permits investigative journalists to enquiry about the system in plain language, and translating the query into SQL (Structured Query Language), MQL (MongoDB Query Language), or Python depending on the essentiality of investigation. SociaLens permits clustering of news information

based on the categories. On the negative side issues like deepfakes, disinformation and fabricated news are the consequences of application AI in journalism in a wrongful way.

References

Anadi. (2024, September 11). *Deep fakes, deeper impacts: AI's role in the 2024 Indian General Election and beyond.* GNET. Retrieved from https://www.gnet.org/deep-fakes-ai-indian-election-2024

Alfter, B. (2018). New method, new skill, new position? Editorial coordinators in cross-border collaborative teams. In R. Sambrook (Ed.), Global teamwork: The rise of collaboration in investigative journalism. Reuters Institute for the Study of Journalism

Amnesty International. (2023). India: Damning new forensic investigation reveals repeated use of Pegasus spyware to target high-profile journalists. Amnesty International. Retrieved from https://www.amnesty.org

Broussard, M. (2014). Artificial intelligence for investigative reporting: Using an expert system to enhance journalists' ability to discover original public affairs stories. Digital Journalism, 2(3). https://doi.org/10.1080/21670811.2014.985497

Chawla, A. (2021). Pegasus Spyware – 'A Privacy Killer'. Delhi High Court. Retrieved from http://dx.doi.org/10.2139/ssrn.3890657

Chuttur, M.Y. (2009). Overview of the Technology Acceptance Model: Origins, Developments and Future Directions. Sprouts: Working Papers on Information Systems, Vol.9 Issue 37, ISSN: 1535-6078. P 9-37. Retrieved from http://sprouts.aisnet.org/9-37

Clements, J. (2018). Collaborative journalism and the law in the UK. In R. Sambrook (Ed.), Global teamwork: The rise of collaboration in investigative journalism. Reuters Institute for the Study of Journalism

Collins. (1990). Essential English dictionary. London: Collins.

Davidson, D. (2024). AI in the newsroom: Revolutionizing journalism in the digital age. Pure Water Books

Davis, F.D. (1986). A Technology Acceptance Model for Empirically Testing New End- User Information Systems: Theory and Results. Doctoral dissertation, MIT (Massachusetts Institutes of Technology) Sloan School of Management, Cambridge, MA.

de-Lima-Santos, M.-F., & Ceron, W. (2022). Artificial intelligence in news media: Current perceptions and future outlook. Journal. Media, 3(1). https://doi.org/10.3390/journalmedia3010002

Diakopoulos, N. (2013). *Can artificial intelligence like IBM's Watson do investigative journalism?* Fast Company. Retrieved from https://www.fastcompany.com/3021545/can-artificial-intelligence-like-ibms-watson-do-investigative-journalism

Dore, B. (2024). Investigating India: How smaller, independent news outlets are digging into politics in a key election year. Global Investigative Journalism Network. Retrieved from https://gijn.org/2024/04/03/investigating-india-independent-news-outlets.

Dörr, K. N. (2015). Mapping the field of Algorithmic Journalism. Digital Journalism, 4(6), Retrieved from https://doi.org/10.1080/21670811.2015.1096748

Environmental Justice Foundation & Tempo Institute. (2024). Investigative journalism training manual: Environmental issues in Indonesia. Environmental Justice Foundation. Bureau of International Narcotics and Law Enforcement Affairs of the U.S. Embassy, Jakarta.

Gentsch, P. (2019). AI in marketing, sales and service: How marketers without a data science degree can use AI, big data and bots. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-89957-2

Goni, M. A., & Tabassum, M. (2020). Artificial Intelligence (AI) in journalism: Is Bangladesh ready for it? A study on journalism students in Bangladesh. Athens Journal of Mass Media and Communications, 6(4). Retrieved from https://doi.org/10.30958/ajmmc.6-4-1

Harris, M. C. (2011). Artificial intelligence. New York: Marshall Cavendish Benchmark.

Helmus, T. C. (2022). *Artificial intelligence, deepfakes, and disinformation: A primer.* RAND Corporation. Retrieved from https://www.rand.org/pubs/perspectives/PEA1043-1.html

Hussien, O., Butt, U., & Bin Sulaiman, R. (2023). Critical analysis and countermeasures tactics, techniques and procedures (TTPs) that target civilians: A case study on Pegasus. arXiv. Retrieved from https://doi.org/10.48550/arXiv.2310.00769

Jamil, H. M., & Rubaiat, S. Y. (2024). Online digital investigative journalism using SociaLens. arXiv preprint arXiv:2410.11890. Retrieved from https://arxiv.org/abs/2410.11890

Kaldani, T., & Prokopets, Z. (2022). *Pegasus spyware and its impacts on human rights*. Council of Europe, Information Society Department.

Kavanagh, J., & Rich, M. D. (2018). *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life.* RAND Corporation. https://www.rand.org/pubs/research_reports/RR2314.html

Kohlman, E. (n.d.). *Indian voters inundated with deepfakes during the largest democratic exercise in the world.* Blackbird.AI. Retrieved from https://blackbird.ai/blog/india-election-deepfakes/

Leiser, M. R. (2022). *Bias, journalistic endeavours, and the risks of artificial intelligence.* In T. Pihlajarinne & A. Alén-Savikko (Eds.), *Artificial intelligence and the media: Reconsidering rights and responsibilities*. Edward Elgar Publishing.

Marzocchi, O., & Mazzini, M. (2022). Pegasus and surveillance spyware: In-depth analysis for the PEGASUS committee. Policy Department for Citizens' Rights and Constitutional Affairs, Directorate-General for Internal Policies, European Parliament.

Mittal, P., & Agarwal, V. (2021). Electoral bonds: Efficacious or a camouflaged road to tyranny. Indian Politics & Law Review Journal (IPLRJ), Vol. 6. Retrieved from https://lawbrigade.com/

Naikodi, C., & Suresh, L. (2020). The today and future of WSN, AI, and IoT: A compass and torchbearer for the technocrats. BPB Publications.

Nandini, Indora, P., & Singh, R. K. (2024). Artificial intelligence in journalism: An overview of its applications and uses. Journal of Communication and Management, 3(3), ISSN: 2583-617X (Online). Retrieved from https://doi.org/10.58966/JCM2024337

NDTV. (2024). Nearly half of the fake news stories in India are political: Study. Retrieved from https://www.ndtv.com/india-news/nearly-half-of-the-fake-news-stories-in-india-are-political-study-7291481

Negnevitsky, M. (2005). Artificial intelligence: A guide to intelligent systems (2nd ed.). Pearson Education Limited.

Networking & Information Technology Research and Development Subcommittee, & Machine Learning & Artificial Intelligence Subcommittee. (2020, March). Artificial intelligence and cybersecurity: Opportunities and challenges: Technical workshop summary report. National Science & Technology Council. Retrieved from https://www.nitrd.gov/pubs/AI-Cybersecurity-Workshop-2020.pdf

Ongsulee, P. (2017). Artificial Intelligence, Machine Learning and Deep Learning. 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), Bangkok, Thailand.

Parasie, S. (2015). Data-driven revelation? Epistemological tensions in investigative journalism in the age of "big data." Digital Journalism, 3(3). Retrieved from https://doi.org/10.1080/21670811.2014.976408

Philp, R. (2024). Revised elections reporting guide for investigative reporters: New digging tools. Global Investigative Journalism Network. Retrieved on December, 22, 2024, from Revised Elections Reporting Guide for Investigative Reporters: New Digging Tools – Global Investigative Journalism Network

Riegert, K. and Collins, S. (2016). Politainment. In Mazzoleni, G (ed.), The International Encyclopaedia of Political Communication. John Wiley & Sons, Inc.

Routley, N. (2023). What is generative AI? An AI explains. World Economic Forum. Retrieved from https://www.weforum.org/agenda/2023/02/generative-ai-explain-algorithms-work/

Sakunia, S. (2024). *AI and deepfakes played a big role in India's elections: From resurrecting dead politicians to creating satirical content and translating speeches, campaigns experimented in multiple ways.* New Lines Magazine. Retrieved from https://newlinesmag.com/spotlight/ai-and-deepfakes-played-a-big-role-in-indias-elections

Schertel Mendes, L., & Stray, J. (2024). AI as a public good: Ensuring democratic control of AI in the information space. Forum on Information & Democracy.

Silva, J. R. (2023). Connect-the-dots: Artificial Intelligence and Automation in Investigative Journalism (Doctoral dissertation). Universidade do Porto.

Sjøvaag, H. (2024). The business of news in the AI economy. AI Magazine, 45(2), 246–255. https://doi.org/10.1002/aaai.12172

Skilton, M., & Hovsepian, F. (2018). The 4th industrial revolution: Responding to the impact of artificial intelligence on business. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-62479-2

Stake, R. E. (1995). The art of case study research. SAGE Publications, Inc.

Strozyk, J. (2024, October 9). Tipsheet for reporters: How to get the best from OCCRP's Aleph. Global Investigative Journalism Network. Retrieved from https://gijn.org/resource/using-aleph/

Texty.org.ua. (n.d.). About us. Retrieved February 7, 2025, from https://texty.org.ua/p/about-en/

The Hindu. (2024). *Fake videos of Modi aides trigger political showdown in India election. The Hindu.* Retrieved from https://www.thehindu.com/sci-tech/technology/fake-videos-of-modi-aides-trigger-political-showdown-in-india-election/article68144464.ece

Vachhani, H., & Rana, P. (n.d.). Analysing the Electoral Bond Scheme, 2018. *Manupatra*. Retrieved on December, 23, 2024, from https://www.manupatra.com

Varadarajan, S. (2021). Revealed: How The Wire and its partners cracked the Pegasus project and what it means for India. The Wire. Retrieved from https://thewire.in

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management Science, 46(2). Retrieved from https://doi.org/10.1287/mnsc.46.2.186.11926

Voulgaris, Z., & Bulut, Y. E. (2018). AI for data science: Artificial intelligence frameworks and functionality for deep learning, optimization, and beyond. Technics Publications.

Waldemarsson, C. (2020). Disinformation, Deepfakes & Democracy: The European response to election interference in the digital age. The Alliance of Democracies Foundation.

Williams, P. N. (1978). Investigative Reporting and Editing. Prentice-Hall.

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE Publications, Inc.