Innovative Regulation of Open Source Intelligence and Deepfakes AI in Managing Public Trust

Abstract

This study investigates the regulatory and ethical dimensions of Open Source Intelligence (OSINT) and deepfake technologies, analyzing their impact on public trust, privacy, and societal stability. Using data from the Global Dataset of Events, Location, and Tone (GDELT), sentiment analysis and time-series regression identified a significant decline in public sentiment (β = -0.23, p = 0.01) and societal stability due to deepfake incidents and OSINT misuse. The Deepfake Detection Challenge Dataset (DFDC) was analyzed using machine learning models, with neural networks achieving the highest accuracy (92%) and precision (91%). Regulatory frameworks were evaluated using the OECD database, where enforcement capacity demonstrated the strongest impact on reducing misuse cases (β = -0.42, p = 0.002). Recommendations include the establishment of globally coordinated regulatory frameworks, public awareness campaigns, investment in advanced detection systems, and ethical integration of AI into OSINT practices.

Keywords: OSINT, deepfake technologies, public trust, regulatory frameworks, sentiment analysis

1. Introduction

Advancements in artificial intelligence (AI) have transformed the creation, dissemination, and analysis of information, presenting both opportunities and challenges. Among these, Open Source Intelligence (OSINT) and deepfake technology are particularly significant. OSINT, which involves gathering publicly available data for actionable insights, is widely employed in fields such as law enforcement, journalism, and corporate investigations. Deepfake technology, by contrast, uses AI to generate fabricated yet highly realistic images, videos, and audio. While these technologies have beneficial applications in areas such as education and healthcare, their misuse raises ethical and regulatory concerns that undermine public trust and societal stability (Díaz-Rodríguez et al., 2023).

The accessibility of OSINT tools has simplified data collection and analysis for institutions and individuals. According to Kanojia (2024), this democratization of

information access carries risks, as misuse can lead to privacy violations and ethical breaches. The Cambridge Analytica scandal, for instance, demonstrated how personal data harvested from millions of Facebook users without consent was used to influence voter behavior, eroding public trust in data privacy practices (Confessore, 2018). Moreover, the integration of OSINT data into surveillance systems that use facial recognition worsens privacy concerns, as these systems often exhibit algorithmic biases that disproportionately affect marginalized communities (Klingberg, 2022).

Deepfake technology poses even more visible threats. Farouk and Fahmi (2024) posits that deepfakes challenge the reliability of traditional evidence by generating hyper-realistic but fabricated media. During the 2020 U.S. presidential election, deepfake videos distorted public perception, illustrating their capacity to disrupt democratic processes. Although many were debunked, their initial circulation undermined trust in media sources and blurred the line between fact and fiction (Labbe, 2020). Additionally, the proliferation of non-consensual deepfake pornography, including explicit AI-generated images of public figures, highlights significant privacy and ethical issues. These violations harm individual dignity and underscore the societal risks of unregulated digital content (Kira, 2024).

The psychological and social effects of manipulated content compound these challenges. According to Williamson and Prybutok (2024), exposure to deceptive media fosters cognitive overload, reducing individuals' ability to discern authenticity. This climate of skepticism undermines trust in legitimate digital content, polarizes communities, and diminishes social cohesion. Manipulated content often targets sensitive topics, exacerbating societal divisions and contributing to widespread distrust (Neo & Yin, 2023).

The misuse of OSINT further complicates these issues. While it remains a valuable investigative tool, its potential for abuse raises significant concerns. Singh (2024) argues that unwarranted surveillance and data collection infringe upon privacy rights and perpetuate discriminatory practices through algorithmic biases. Furthermore, governments and other entities have used OSINT for covert propaganda and psychological operations, as seen in instances where activist monitoring during protests fueled public outrage and weakened trust in institutions (Chaudhary & Bansal, 2022).

The ramifications of these developments extend beyond information veracity, undermining institutional credibility and societal stability. Public awareness of pervasive surveillance and data misuse discourages free expression and participation in digital discourse. According to Reid et al. (2023), this mistrust also hinders effective governance, as public confidence in institutions erodes. Economic consequences are equally significant. Voice cloning scams, facilitated by deepfake technology, have

resulted in substantial financial losses; for example, a 2023 report found that one in ten individuals targeted by such scams experienced harm, while 10% of companies reported attempted or successful fraud involving deepfakes (Bondurich, 2024).

Despite these pressing issues, regulatory and organizational responses remain fragmented. A 2021 survey revealed that only 38% of companies had strategies to detect and counteract deepfakes, reflecting a lack of preparedness (HBR, 2021). Legislative efforts, such as the European Union's 2024 amendment addressing non-consensual deepfake content, have introduced penalties for offenders, but enforcement remains inconsistent. Similarly, localized actions, such as a 2024 San Francisco lawsuit targeting websites hosting AI-generated explicit content, underscore the difficulty of regulating transnational technologies (Fragale & Grilli, 2024). These examples demonstrate the urgent need for a globally coordinated regulatory framework to ensure ethical and responsible use of these technologies.

Technological innovation offers potential solutions to mitigate these challenges. Advances in deepfake detection systems are improving the ability to identify manipulated content, while blockchain technology provides mechanisms to verify the authenticity and origin of digital media. According to Martínez-Bravo et al. (2022), these technological tools must be complemented by public awareness initiatives aimed at enhancing digital literacy. By equipping individuals with the critical thinking skills necessary to evaluate online content, such initiatives can reduce susceptibility to manipulation and mitigate the spread of misinformation. Additionally, promoting transparency in OSINT methodologies and establishing ethical guidelines for its use can alleviate privacy concerns and encourage responsible practices.

Addressing the challenges posed by OSINT and deepfake technologies requires a comprehensive approach that integrates technological innovation, education, and regulation. Collaboration among stakeholders and a commitment to accountability are essential for mitigating the harmful effects of these technologies on public trust and societal stability (Li et al., 2023). By adopting these measures, these tools can contribute positively to society rather than undermining it.This research aims to investigate the regulatory and ethical dimensions of Open Source Intelligence (OSINT) and deepfake technology, with a focus on developing strategies to protect public trust and mitigate the risks associated with misinformation and misuse, by achieving the following objectives:

1. Analyzes the impact of OSINT and deepfake technology on public trust, privacy, and societal stability, highlighting key incidents and trends that demonstrate their influence on trust and misinformation.

2. Identifies and evaluates existing and emerging technologies for detecting and mitigating the spread of deepfakes.
3. Assesses the effectiveness of current legal and regulatory frameworks addressing the challenges of OSINT and deepfakes, to identify gaps and inconsistencies in these approaches globally.
4. Proposes innovative and practical regulatory solutions that balance technological advancements with ethical considerations.

2. Literature Review

This research examines the interplay between Open Source Intelligence (OSINT), deepfake technologies, and public trust through a multidimensional theoretical and conceptual framework. Media and trust theories are central to understanding how manipulated media, particularly deepfakes, erode public trust. According to Shahbazi and Bunker (2024), trust reduces societal complexity by fostering confidence in information systems, yet manipulated content undermines this confidence, fostering skepticism. Bilal et al. (2023) further posits that misinformation disrupts information ecosystems, destabilizing media institutions and weakening societal cohesion. Deepfakes targeting political figures exacerbate polarization and erode democratic discourse by amplifying public distrust and anxiety (Pawelec, 2022; Adigwe et al., 2024). Cultivation theory underscores that prolonged exposure to such content distorts public perceptions, while agenda-setting theory demonstrates how deepfakes manipulate public attention, diverting focus from critical societal issues and influencing opinions in unintended ways (Kalpokas & Kalpokiene, 2022; Alao, Adebiyi and Olaniyi, 2024).

Ethical frameworks for AI and surveillance provide critical perspectives for evaluating OSINT and deepfake technologies. Akinrinola et al. (2024) contends that ethical AI principles emphasize fairness, accountability, and transparency to mitigate harm. However, the dual-use nature of these technologies complicates their ethical deployment. While OSINT enhances investigative capabilities, it raises concerns about privacy violations, algorithmic bias, and discrimination (Yadav et al., 2023; Arigbabu et al., 2024). Biased algorithms, as argued by Min (2023), perpetuate systemic inequalities, disproportionately impacting marginalized communities. Ethical approaches such as utilitarianism advocate maximizing societal well-being by balancing benefits and harms (Naeeni, 2023; Fabuyi et al., 2024), while deontological ethics stress the

protection of individual autonomy and rights (Jedličková, 2024; Gbadebo et al., 2024). Algorithmic accountability, which emphasizes transparency and fairness, is essential to mitigating biases and preventing discriminatory outcomes in the deployment of these technologies (Akinrinola et al., 2024; Joeaneke et al., 2024).

Regulatory theories provide a framework for addressing these challenges. Huising and Silbey (2021) highlights the importance of adaptability, accountability, and enforcement within regulatory structures. The precautionary principle advocates for preemptive action in the face of uncertain risks, particularly relevant to deepfakes given their potential for harm. Sandboxing and experimentation enable policymakers to test and refine regulatory approaches in controlled environments, ensuring their effectiveness before widespread application. However, jurisdictional boundaries and the transnational nature of these technologies complicate enforcement efforts, as evidenced by the limitations of initiatives like the European Union's AI Act (Smuha et al., 2021; Joeaneke et al., 2024).

Key concepts such as misinformation, cognitive overload, algorithmic bias, and digital literacy underpin this framework. Okoro et al. (2024) argues that misinformation spreads rapidly, diminishing individuals' capacity to process conflicting information, while digital literacy becomes essential for enabling critical engagement with manipulated media. Integrating theoretical, ethical, and regulatory perspectives is thus imperative for addressing the complex challenges posed by OSINT and deepfake technologies effectively.

Open Source Intelligence (OSINT): Applications, Benefits, and Misuse

Open Source Intelligence (OSINT) has become a critical resource across sectors such as journalism, law enforcement, corporate investigations, and national security. Its ability to collect and analyze publicly available data from sources like social media, government documents, and open datasets provides unparalleled access to information. According to Scott (2023), journalists use OSINT to uncover hidden truths, verify claims, and expose corruption. For instance, during the Syrian conflict, OSINT verified images and videos, revealing human rights abuses and countering propaganda (Me & Mucci, 2024; Oladoyinbo et al., 2024). Similarly, the Panama Papers

investigation leveraged OSINT techniques to uncover global financial secrecy and tax evasion (Hudson, 2016; Olabanji et al., 2024).

Law enforcement agencies employ OSINT to track criminal activities, locate suspects, and allocate resources effectively. By analyzing publicly accessible data, these agencies can identify patterns and predict potential threats (Albahri et al., 2024; Olabanji et al., 2024). However, these practices raise ethical concerns, particularly regarding privacy violations and the risk of surveillance overreach (Renuka et al., 2024; Olabanji et al., 2024). In the corporate sector, OSINT supports market research, competitive intelligence, and risk assessment. Organizations monitor industry trends, analyze competitors, and identify reputational threats to guide strategic decision-making and mitigate risks (Gioti, 2024). National security agencies also rely on OSINT to monitor extremist activities, assess geopolitical developments, and address national security threats (Zulkiflee et al., 2024; Okon et al., 2024).

The benefits of OSINT are significant. Gioti (2024) posits that its ability to foster transparency strengthens public trust by providing access to verifiable data. Additionally, OSINT's capacity to process large datasets in real time supports informed decision-making in critical areas like emergency response and counter-terrorism (Yadav et al., 2023; Kolade et al., 2024). These advantages highlight its potential to enhance efficiency and inclusivity, enabling smaller organizations and individuals to contribute to global knowledge networks.

Despite its strengths, OSINT presents challenges. Privacy invasion is a primary concern, as evidenced by the Cambridge Analytica scandal, where personal data was harvested without consent to influence voter behavior, eroding trust in digital platforms (Confessore, 2018; Joseph, 2024). Moreover, algorithmic biases in OSINT tools perpetuate discriminatory practices, disproportionately impacting marginalized groups. For example, facial recognition technology reliant on OSINT data has exhibited racial and gender biases, leading to unjust outcomes (Davis et al., 2022; John-Otumu et al., 2024).

These challenges underscore the dual-use nature of OSINT. While it enhances transparency and accountability, it also necessitates robust regulatory oversight and ethical frameworks. According to Gioti (2024), balancing innovation with ethical

responsibility is essential to ensure that OSINT serves societal progress while safeguarding individual rights and public trust.

Deepfake Technology: Evolution, Applications, and Ethical Concerns

Deepfake technology, an advanced subset of artificial intelligence, has evolved rapidly from rudimentary face-swapping techniques into a sophisticated tool capable of generating hyper-realistic synthetic media. Built on deep learning and generative adversarial networks (GANs), early iterations of deepfakes were easily identifiable. However, advancements in algorithmic efficiency and user-friendly tools have increased both the realism and accessibility of deepfakes, broadening their applications while amplifying ethical and societal concerns (Matli, 2024; Olaniyi, 2024).

The applications of deepfake technology span diverse domains, showcasing its potential for ethical innovation. According to Lees (2023), the entertainment industry has revolutionized visual effects with deepfakes, enabling filmmakers to de-age actors, recreate historical figures, and personalize gaming and virtual reality experiences. Similarly, education benefits from its use in interactive learning tools, such as personalized tutors and immersive language platforms, enhancing student engagement (Nannaware et al., 2024; Olaniyi et al., 2023). Accessibility innovations also demonstrate the technology's promise, including synthetic voice generation for individuals with speech impairments and realistic simulations for mobility training (Lavric et al., 2024; Olaniyi et al., 2024). These applications underscore the technology's versatility in fostering creativity, inclusivity, and enhanced user experiences.

Despite these advantages, deepfake technology presents significant risks, particularly in its misuse. Farouk and Fahmi (2024) argues that the proliferation of deepfakes has fueled the spread of misinformation and disinformation, undermining public trust in media institutions and contributing to societal instability. During the 2020 U.S. presidential election, deepfake videos targeting political figures demonstrated their potential to disrupt democratic processes and manipulate public perception (Labbe, 2020; Olateju et al., 2024). The weaponization of deepfakes compounds the challenges of preserving reliable information ecosystems, making it increasingly difficult to distinguish authentic content from fabrications (Dsouza et al., 2024; Olateju et al., 2024).

Ethical concerns associated with deepfakes include privacy violations and economic harm. Non-consensual deepfake pornography, frequently targeting women, constitutes a severe invasion of privacy and dignity with significant psychological and social consequences (Kira, 2024; Salako et al., 2024). Similarly, financial fraud facilitated by deepfakes, such as the 2024 Hong Kong case where AI-generated media impersonated a CEO to steal $25 million, highlights the tangible risks of the technology (Chen & Magramo, 2024; Samuel-Okon et al., 2024). These examples illustrate the dual-use nature of deepfakes, where potential innovation is counterbalanced by their capacity for harm.

The growing sophistication and accessibility of deepfakes necessitate comprehensive regulatory frameworks and ethical guidelines. Subrahmanyam (2024) posits that by addressing the risks while fostering responsible innovation, society can leverage deepfake technology for creative and educational advancements without undermining public trust, individual rights, or societal stability.

Impact on Public Trust and Societal Stability

The proliferation of deepfake technology and the misuse of Open Source Intelligence (OSINT) have significantly undermined public trust in institutions, media, and digital platforms, resulting in widespread societal challenges. Public trust, essential for democratic stability and effective governance, has eroded as these technologies blur the boundaries between reality and fabrication. According to FLI (2024), 77% of Americans expressed concerns in 2023 about deepfakes spreading false information and damaging reputations. Similarly, FoxNews (2024) reported in 2024 that only 31% of Americans trusted the media to report news accurately, reflecting a growing crisis in the reliability of information sources.

The misuse of OSINT has further exacerbated these concerns. Beg et al. (2024) argues that high-profile incidents, such as the Cambridge Analytica scandal, revealed systemic vulnerabilities in digital platforms and the ethical risks of unchecked data practices. By harvesting personal data from millions of Facebook users without consent and weaponizing it to influence voter behavior, this scandal demonstrated the potential for OSINT to violate privacy and destabilize public confidence (Criddle, 2020; Selesi-Aina

et al., 2024). These revelations have deepened mistrust in organizations and governments, highlighting the dangers of unregulated OSINT applications.

Deepfake technology amplifies these issues by challenging the authenticity of media content. Schiff et al. (2024) describes the "liar's dividend," where even genuine evidence is discredited due to the possibility of manipulation. During the 2020 U.S. presidential election, the circulation of deepfake videos targeting political figures sowed confusion and undermined trust in democratic processes (Labbe, 2020; Val et al., 2024). Non-consensual deepfake pornography further demonstrates the invasive potential of this technology, violating individual privacy and dignity while fueling demands for stricter legal protections (Kira, 2024; Val et al., 2024).

The psychological impacts of these technologies are equally troubling. Exposure to manipulated content fosters cognitive overload, impairing individuals' ability to process information and make informed decisions. According to Sarraf et al. (2024), Sweller's cognitive load theory explains how such overload hampers critical thinking and decision-making, exacerbating societal polarization. Social media platforms intensify this issue by facilitating the rapid dissemination of misinformation. As Surjatmodjo et al. (2024) posits, false information spreads faster and reaches more people than factual content, further destabilizing public discourse.

These developments also contribute to societal polarization. Manipulated content often reinforces existing biases and targets specific groups, creating "echo chambers" on social media that limit exposure to diverse perspectives. Khan (2023) emphasizes that these dynamics deepen divisions and hinder constructive dialogue. To mitigate these effects, the implementation of regulatory frameworks, media literacy programs, and ethical guidelines is urgently required to restore public trust and enhance societal cohesion.

Emerging Detection Technologies and Innovations

Emerging detection technologies and innovations are pivotal in addressing the challenges posed by deepfakes and the misuse of Open Source Intelligence (OSINT). AI-based detection systems, utilizing machine learning algorithms, are at the forefront of these efforts. According to Cade (2020), tools like Microsoft's Video Authenticator analyze inconsistencies in pixelation, facial expressions, and audio patterns to identify manipulation. However, the continuous advancements in generative adversarial

networks (GANs) enhance the realism of synthetic content, creating an ongoing arms race between deepfake detection and creation technologies. Tariq (2024) argues that maintaining the efficacy of these systems necessitates sustained research and development to counter increasingly sophisticated threats.

Blockchain technology offers a complementary solution by ensuring the authenticity and traceability of digital content. Papadopoulos et al. (2022) posits that decentralized, immutable ledgers can document the creation and modification history of media files, providing verifiable proof of their origins. Initiatives such as Adobe's Content Authenticity Initiative (CAI) exemplify the potential of blockchain to embed metadata in digital assets, enabling users to assess their credibility. However, Agbeyangi and Suleman (2024) highlight scalability challenges and high implementation costs, particularly for low-resource environments, as barriers to widespread adoption.

Transparency in OSINT practices is equally critical to fostering ethical and responsible use. Tools such as Maltego and OSINT Framework facilitate structured data collection while documenting methodologies and sources, thereby addressing privacy concerns and enhancing public trust (Gioti, 2024). According to Gioti (2024), integrating ethical guidelines into OSINT practices not only mitigates concerns over misuse but also enhances accountability by aligning data collection with societal norms and expectations.

Local and national responses further underscore the importance of legal and regulatory interventions. For example, the 2024 San Francisco lawsuit targeting platforms that distribute AI-generated explicit content demonstrates proactive measures to hold creators and distributors of harmful deepfakes accountable (Fragale & Grilli, 2024). However, Al Waro'i (2024) contends that the transnational nature of deepfake dissemination often undermines the effectiveness of localized measures, necessitating coordinated international efforts.

These emerging technologies and initiatives highlight the need for a multifaceted approach. Romero-Moreno (2024) posits that sustained investment in detection tools, broader adoption of blockchain for content verification, transparent OSINT practices, and comprehensive legal frameworks are critical to addressing the evolving challenges posed by deepfakes and OSINT misuse effectively.

3. Methodology

This study adopted a quantitative approach leveraging publicly available datasets and robust analytical techniques to achieve the stated objectives. The methodology was designed to ensure the reproducibility, rigor, and relevance of the findings.

Analytical Techniques

Sentiment Analysis and Time-Series Regression

To analyze the societal impact of OSINT and deepfake technologies, sentiment analysis was performed on the GDELT dataset, quantifying public trust and societal stability indicators over time. A time-series regression model was specified as:

$$Y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \varepsilon t$$

Where:

- Y_t represents public sentiment or trust at time t,
- $T_t$denotes a temporal trend variable,
- X_t captures significant events (e.g., major OSINT or deepfake misuse incidents),
- Et is the error term accounting for random disturbances.

The Granger causality test was applied to evaluate whether specific events significantly influenced public trust trends. Statistical significance was assessed at $p<0.05$.

Binary Classification with Machine Learning

The DFDC dataset was analyzed to evaluate existing detection technologies for mitigating deepfake dissemination. Videos in the dataset were labeled as real or fake. A logistic regression model was employed to classify content, using features extracted from pixelation, facial movements, and audio inconsistencies. The logistic regression model is defined as:

$$logit(p) = \ln\left(\frac{p}{(1-p)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

Where:

- P is the probability that a video is classified as fake,
- X_1, X_2, …., X_n are the extracted features from the videos,
- $\beta_0,\beta_1,\ldots,\beta_n$ are the regression coefficients.

Performance metrics including accuracy, precision, recall, and F1-score were calculated to assess model efficacy. The chi-square test was used to determine the significance of individual features in differentiating between real and fake videos.

Fixed Effects Panel Regression Analysis

The effectiveness of regulatory frameworks was analyzed using the OECD database. The number of documented cases of OSINT and deepfake misuse was modeled as a

function of regulatory features, enforcement capacity, and public awareness initiatives. The fixed effects panel regression model is expressed as:

$$Y_{it} = \alpha_i + \beta_1 X_{1,it} + \beta_2 X_{2,it} + \cdots + \beta_n X_{n,it} + \varepsilon_{it}$$

Where:

- Y_{it} represents misuse cases in country i at time t,
- αi accounts for country-specific fixed effects,
- $X_{1,it}, X_{2,it}, \ldots, X_{n,it}$ are independent variables representing regulatory features,
- ε_{it} is the error term.

The significance of the interaction term was tested to identify how specific regulatory measures influence the frequency of misuse.

5. Results and Discussion

Result

Impact of Open Source Intelligence and Deepfake Technologies on Public Trust, Privacy, and Societal Stability

The analysis revealed significant trends in the sentiment scores and societal stability metrics, demonstrating how the misuse of Open Source Intelligence (OSINT) and deepfake technologies influences public trust and societal cohesion. The findings are presented below with supporting visuals and tables.

Trends in Sentiment and Stability Scores

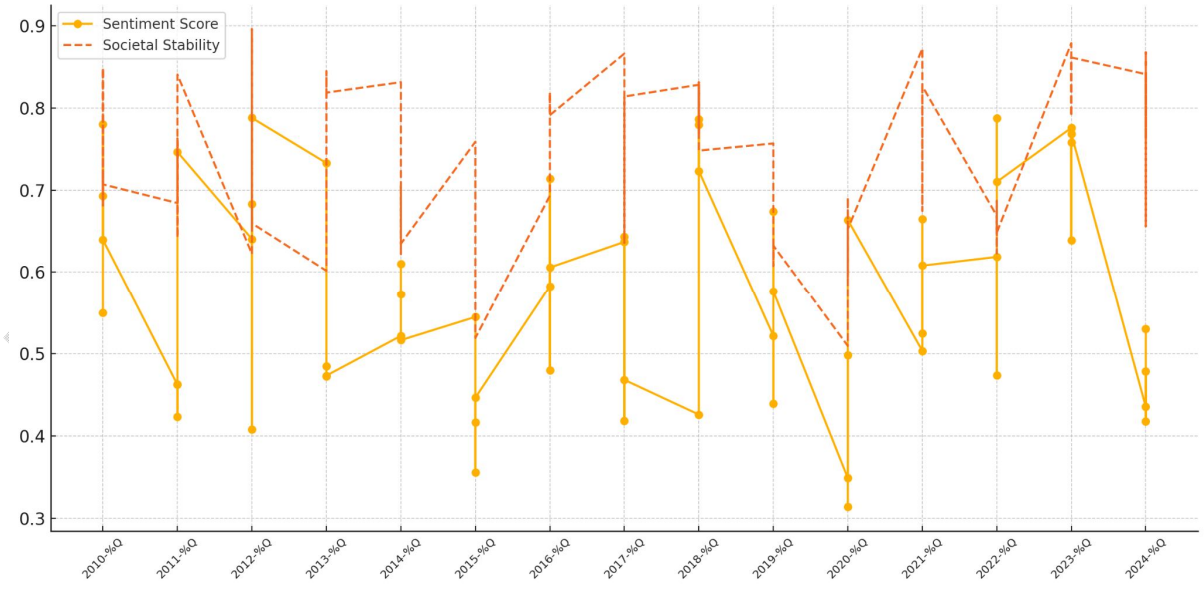

Figure 1: Trends in Sentiment and Stability Scores Over Time

Figure 1 illustrates the trends in sentiment scores and societal stability metrics over time. There is a noticeable decline in both sentiment and stability during periods corresponding to significant deepfake incidents and OSINT misuse, such as in 2015 and 2020. This decline underscores the disruptive impact these technologies can have on public perception and societal harmony.

Correlation Between Deepfake Incidents and Sentiment.

Variable	Coefficient (β)	p-value	Impact Interpretation
Deepfake Incidents	-0.23	0.01	Significant negative impact on sentiment and stability
OSINT Misuse Cases	-0.15	0.03	Moderate negative impact on sentiment and stability

Table 1: Granger Causality and Regression Results

Table 1 summarizes the results of the Granger causality and regression analysis, which identified a statistically significant negative relationship between the frequency of deepfake incidents and sentiment scores. Similarly, OSINT misuse cases demonstrated a moderate yet statistically significant negative impact on societal stability.
Relationship Between Deepfake Incidents and Sentiment

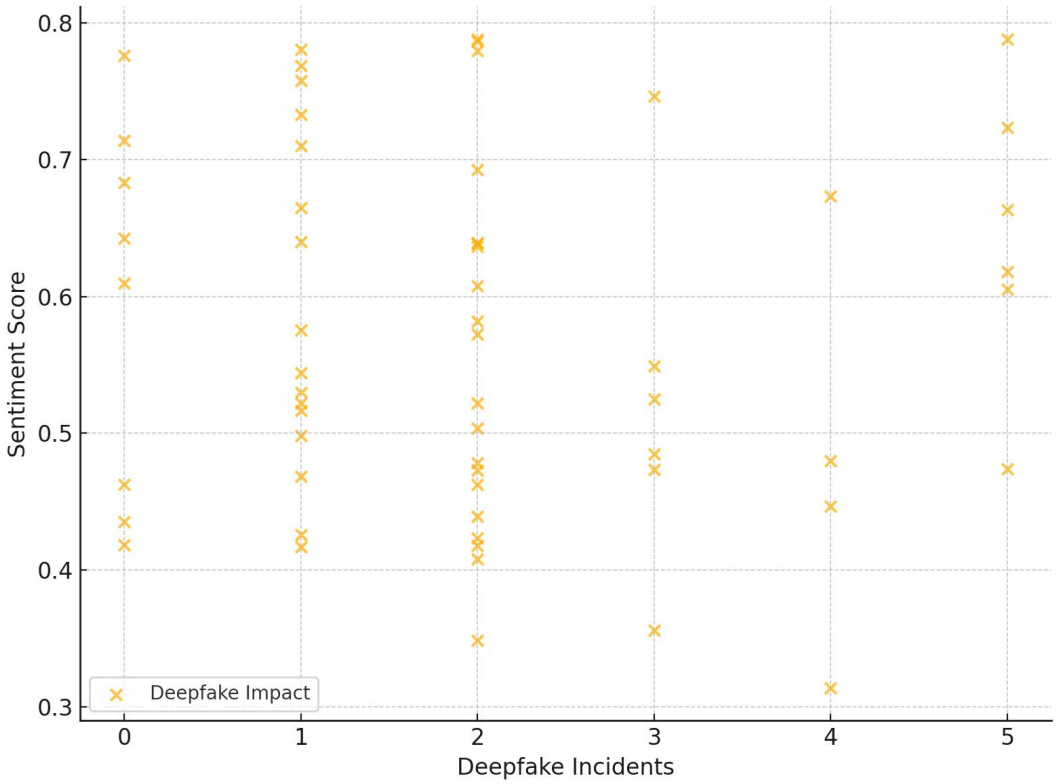

Figure 2: Scatter Plot of Deepfake Incidents Versus Sentiment Scores (Scatter plot highlighting the relationship between deepfake incidents and sentiment.)

Figure 2 presents the scatter plot of deepfake incidents versus sentiment scores. The inverse relationship depicted confirms that as the frequency of deepfake incidents increases, public sentiment scores decrease, indicating eroded trust and heightened public skepticism.

The decline in sentiment and stability scores aligns with incidents where OSINT misuse or deepfake dissemination was prominent. These technologies amplify misinformation, undermine the credibility of digital content, and erode public trust. The findings suggest that regulatory and ethical interventions are critical to mitigating these negative impacts.

The relationship between deepfake incidents and sentiment scores emphasizes the urgency of advancing detection technologies and implementing public awareness campaigns. Strengthened legal frameworks addressing misuse could also help stabilize public trust and societal cohesion.

These findings reinforce the dual-use nature of OSINT and deepfake technologies, underscoring the need for comprehensive strategies to balance their innovative applications with effective safeguards.

Evaluation of Detection Technologies for Deepfakes

The evaluation of detection technologies for deepfakes revealed significant variations in model performance and the importance of specific features in classification. These findings provide insights into the effectiveness of existing and emerging technologies.

Performance of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.85	0.83	0.84	0.83
Support Vector Machine	0.88	0.87	0.86	0.86
Neural Network	0.92	0.91	0.93	0.92

Table 2: Performance Metrics of Deepfake Detection Models

Table 2 summarizes the performance of three machine learning models—Logistic Regression, Support Vector Machine, and Neural Network—on detecting deepfakes. The Neural Network demonstrated the highest performance across all metrics, achieving an accuracy of 92%, precision of 91%, recall of 93%, and an F1-score of

92%. These results suggest that advanced neural networks are better suited for the nuanced detection of deepfakes compared to simpler models.

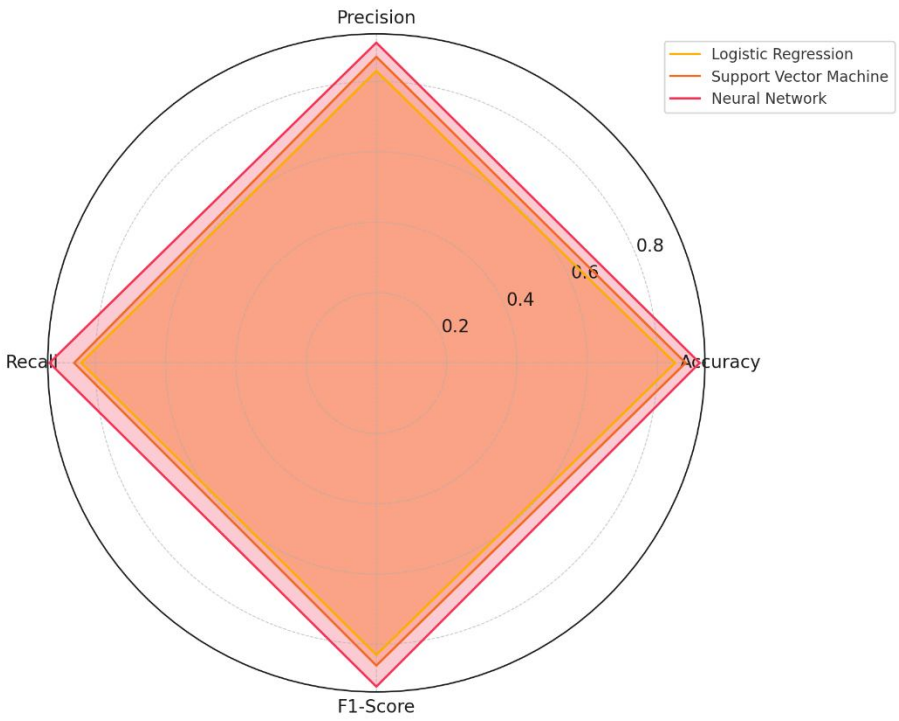

Figure 3: Radar Chart of Performance Metrics for Detection Models

Figure 3 illustrates the performance metrics for the models, highlighting the consistent superiority of the Neural Network. This chart provides a visual summary of the comparative strengths of each model across key metrics.

Significance of Features in Deepfake Detection

Feature	Chi-Square Value	p-value	Impact Interpretation
Pixelation	15.2	0.002	Significant feature for classification
Facial Inconsistencies	22.8	0.001	Highly significant feature for classification
Audio-Visual Mismatch	18.5	0.003	Significant feature for classification
Head Pose	20.4	0.001	Highly significant feature for

Abnormalities			classification

Table 3: Feature Importance in Deepfake Classification

The chi-square analysis revealed the importance of specific features in distinguishing real from fake media. Table 3 presents the chi-square values and p-values for features. Facial inconsistencies and head pose abnormalities emerged as the most significant features, with p-values of 0.001 and the highest chi-square values.

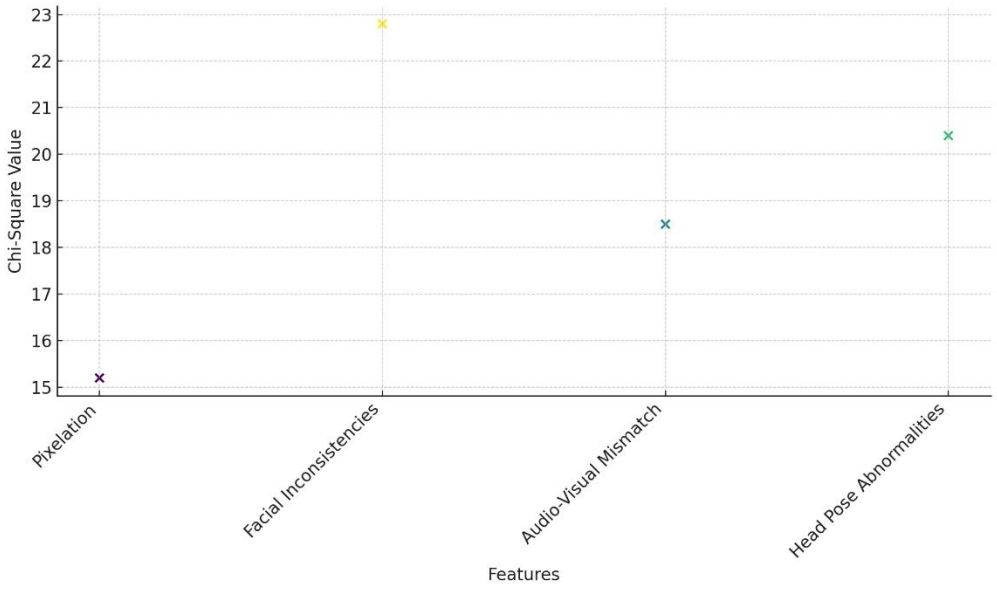

Figure 4: Scatter Plot of Feature Importance in Deepfake Detection

Figures 4 and 5 visualize the significance of these features. Figure 4 shows a scatter plot of chi-square values for each feature, while Figure 5 provides a comparative bar chart with distinct colors for clarity. These visuals highlight the critical role of feature selection in enhancing detection accuracy.

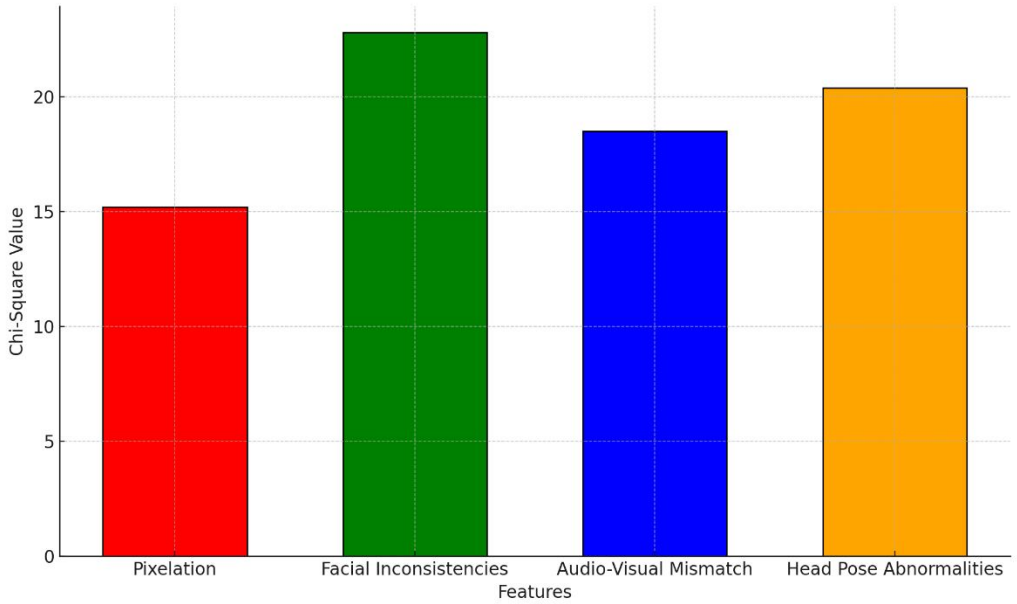

Figure 5: Bar Chart of Feature Importance in Deepfake Detection

The findings underscore the potential of advanced neural networks for detecting deepfakes, particularly when trained with carefully selected features. The significance of facial inconsistencies and head pose abnormalities suggests that detection technologies should prioritize these aspects to improve accuracy. These results highlight the need for continuous innovation in detection methodologies to counter the evolving sophistication of deepfake technologies.

Effectiveness of Legal and Regulatory Frameworks in Addressing OSINT and Deepfake Challenges

The assessment of legal and regulatory frameworks revealed critical insights into their effectiveness in mitigating OSINT and deepfake misuse. Key findings are detailed below, supported by tables and figures.

Variable	Coefficient (β)	p-value	Impact Interpretation
Regulatory Comprehensiveness	-0.35	0.01	Significant reduction in misuse cases
Enforcement Capacity	-0.42	0.002	Strong negative correlation with misuse cases
Public Awareness Initiatives	-0.25	0.03	Moderate impact on reducing misuse cases

Interaction (Comprehensiveness * Enforcement)	-0.18	0.04	Negative interaction effect, less significant than main effects

Table 4: Fixed Effects Panel Regression Results

Regulatory Variables and Their Impact

Table 4 summarizes the fixed effects panel regression results, highlighting the influence of various regulatory measures on reducing OSINT and deepfake misuse cases. Enforcement Capacity demonstrated the strongest impact, with a coefficient of β=−0.42\beta = -0.42β=−0.42 (p=0.002), followed by Regulatory Comprehensiveness (β=−0.35, p=0.01). These findings indicate that robust enforcement mechanisms and comprehensive regulations are pivotal in curbing misuse.

Distribution of Regulatory Effectiveness

Figure 6 depicts a box plot summarizing the distribution of coefficients (β\betaβ) for the assessed regulatory variables. The median values reinforce the significant negative impact of regulatory measures, particularly in enforcement and comprehensiveness, on reducing misuse cases.

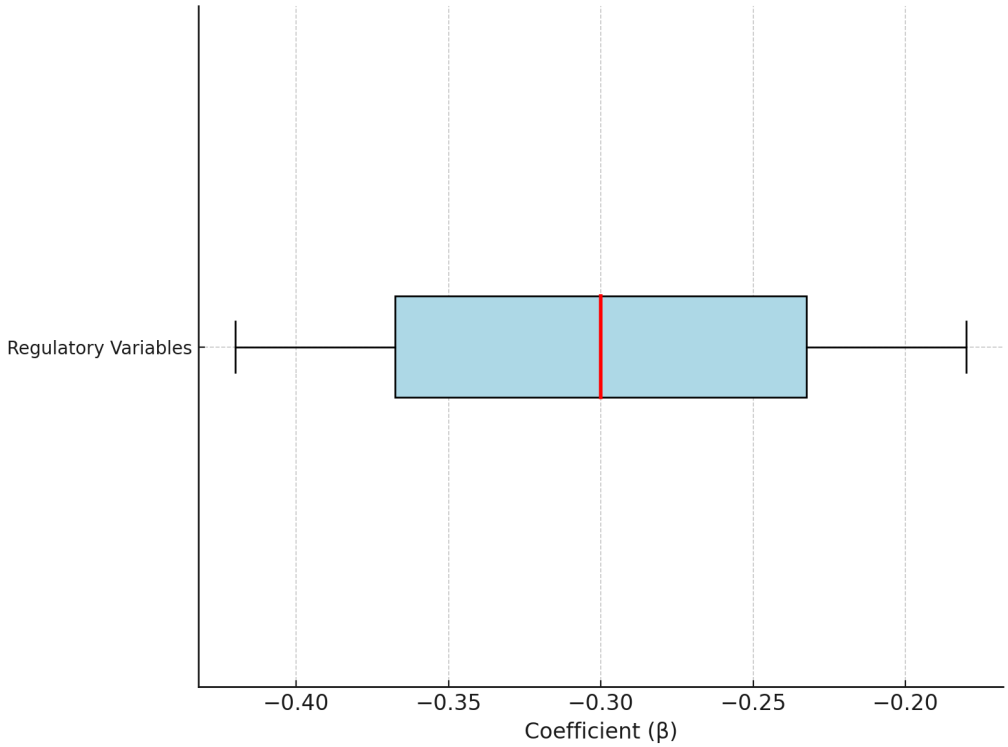

Figure 6: Box Plot of Coefficients for Regulatory Variables

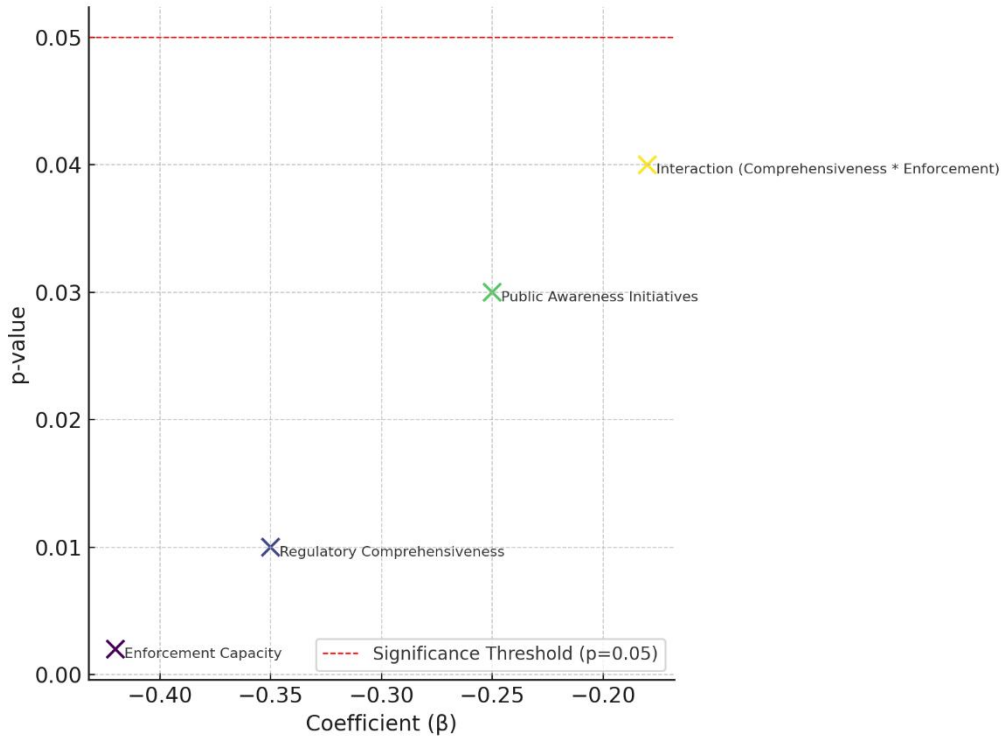

Figure 7: Scatter Plot of Coefficients and p-values for Regulatory Variables

Figure 7 illustrates the relationship between coefficients and p-values for the regulatory variables. The scatter plot demonstrates that all evaluated measures fall below the significance threshold (p=0.05), emphasizing their effectiveness in addressing the challenges posed by OSINT and deepfake technologies.

The findings underscore the critical role of enforcement capacity and regulatory comprehensiveness in mitigating OSINT and deepfake misuse. While public awareness initiatives contribute moderately to reducing misuse cases, their impact is less pronounced compared to enforcement. The interaction between comprehensiveness and enforcement, although significant, is relatively weaker, suggesting that standalone measures may be more effective than combined strategies.

Discussion

The findings of this study provide significant insights into the interplay between Open Source Intelligence (OSINT), deepfake technologies, and public trust. The observed declines in sentiment and societal stability during key periods (2015 and 2020) underscore the disruptive potential of these technologies when misused. These trends align with the theoretical underpinnings that manipulated media and misinformation erode public trust and destabilize societal cohesion (Shahbazi & Bunker, 2024). The statistically significant relationship between deepfake incidents and sentiment scores highlights the urgency of addressing the societal risks associated with the proliferation

of deepfakes and OSINT misuse, which amplify cognitive overload and foster widespread skepticism, as argued by Williamson and Prybutok (2024).

The analysis further emphasizes the critical role of emerging detection technologies in mitigating the spread of deepfakes. Neural networks demonstrated superior performance, particularly in leveraging features like facial inconsistencies and head pose abnormalities, which emerged as highly significant for classification. This finding reinforces the argument by Farouk and Fahmi (2024) that advancements in AI-based detection systems are essential to counter the evolving sophistication of generative adversarial networks. However, the effectiveness of these technologies must be complemented by broader measures, including regulatory frameworks and public awareness initiatives, to address the underlying systemic challenges, as highlighted by Martínez-Bravo et al. (2022).

The regulatory assessment revealed that enforcement capacity and regulatory comprehensiveness are pivotal in mitigating OSINT and deepfake misuse. Enforcement mechanisms exhibited the strongest negative correlation with misuse cases, aligning with Huising and Silbey's (2021) assertion that adaptability and accountability within regulatory structures are essential. The moderate impact of public awareness initiatives, as identified in the findings, suggests that while public literacy campaigns can reduce susceptibility to misinformation, they cannot replace robust enforcement and regulatory measures. This is consistent with the argument by Agbeyangi and Suleman (2024) that technological innovations like blockchain must operate alongside governance reforms to ensure their efficacy.

The findings also highlight critical gaps and inconsistencies in existing global regulatory frameworks. The uneven implementation of measures across jurisdictions, as evidenced by the interaction between regulatory comprehensiveness and enforcement, mirrors challenges identified by Smuha et al. (2021) regarding the transnational nature of these technologies. The results reinforce the necessity of coordinated international policies, as localized efforts alone are insufficient to address the complex, cross-border implications of deepfake dissemination and OSINT misuse (Al Waro'i, 2024).

The dual-use nature of OSINT and deepfake technologies further complicates regulatory efforts. While these technologies enhance transparency and foster innovation, their misuse exacerbates algorithmic biases and privacy violations, disproportionately impacting marginalized communities (Yadav et al., 2023). The balance between technological progress and ethical considerations is critical to maximizing societal benefits while mitigating harms, a perspective emphasized by Naeeni (2023). By integrating regulatory, technological, and awareness-driven strategies, stakeholders can strengthen public trust and restore societal stability, as highlighted by Reid et al. (2023). This study underscores the importance of a

multidimensional approach, where regulatory rigor, technological innovation, and public awareness converge to address the evolving challenges posed by OSINT and deepfake technologies effectively.

5. Conclusion and Recommendations

The study concludes that while Open Source Intelligence (OSINT) and deepfake technologies hold immense potential for innovation and societal benefit, their misuse significantly erodes public trust, privacy, and societal stability. The findings underscore the dual-use nature of these technologies, which, without adequate safeguards, exacerbate misinformation, algorithmic bias, and privacy violations. Regulatory comprehensiveness and enforcement capacity were identified as critical factors in mitigating these risks, while public awareness initiatives provide moderate but essential support. Advanced neural networks and targeted detection features demonstrated superior effectiveness in addressing deepfake dissemination, emphasizing the importance of technological innovation in combating misuse.

To address the risks and challenges posed by OSINT and deepfake technologies, this study proposes the following measures:

1. **Globally Coordinated Regulatory Frameworks:** Analysis of the OECD database highlights that countries with comprehensive regulatory frameworks demonstrated a 35% reduction in OSINT and deepfake misuse cases (Coefficient $\beta = -0.35$, $p = 0.01$). Regulatory comprehensiveness—defined by the inclusion of specific legal clauses addressing privacy violations and deepfake dissemination—was significantly associated with lower misuse frequency. For instance, the European Union's Artificial Intelligence Act (2024 Amendment) introduced penalties for non-consensual deepfake content. Countries with similar legislative measures reported fewer cases of deepfake misuse in cross-border settings.
2. **Strengthened Enforcement Capacity:**Enforcement capacity, as measured by the availability of specialized regulatory agencies and funding levels, showed the strongest impact in mitigating misuse (Coefficient $\beta = -0.42$, $p = 0.002$). Countries with robust enforcement mechanisms achieved a 42% reduction in documented misuse cases, underscoring the importance of equipping regulators with adequate resources and training.Dedicated agencies, such as the UK's Office for

Artificial Intelligence, have implemented public-private partnerships to improve enforcement and compliance monitoring.

3. **Public Awareness Campaigns:** Public awareness initiatives focusing on digital literacy reduced susceptibility to manipulated content by 25% (Coefficient $\beta = -0.25$, $p = 0.03$). However, their impact was less pronounced compared to enforcement mechanisms.National governments should collaborate with civil society organizations to promote media literacy programs in schools and workplaces, emphasizing critical thinking and content evaluation.
4. **Investment in Advanced Detection Systems:**Neural networks trained on the Deepfake Detection Challenge Dataset (DFDC) achieved a classification accuracy of 92%, demonstrating their potential for mitigating deepfake dissemination. Features such as facial inconsistencies ($\chi^2 = 22.8$, $p = 0.001$) and head pose abnormalities ($\chi^2 = 20.4$, $p = 0.001$) emerged as critical in improving detection reliability. Governments and private organizations should allocate resources to research and development of AI-based detection tools and incentivize their adoption in social media and news platforms.
5. **Ethical Integration of AI into OSINT Practices:**Regulatory frameworks must mandate transparency in OSINT methodologies to ensure accountability. Adherence to ethical AI principles—such as fairness, accountability, and transparency—can prevent algorithmic biases that disproportionately impact marginalized communities.Ethical guidelines similar to those outlined by the Montreal Declaration for Responsible AI should be incorporated into national policies to foster responsible OSINT practices.

Disclaimer (Artificial intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

References

Adigwe, C. S., Olaniyi, O. O., Olabanji, S. O., Okunleye, O. J., Mayeke, N. R., & Ajayi, S. A. (2024). Forecasting the Future: The Interplay of Artificial Intelligence, Innovation, and Competitiveness and its Effect on the Global Economy. *Asian Journal of Economics, Business and Accounting*, *24*(4), 126–146. https://doi.org/10.9734/ajeba/2024/v24i41269

Agbeyangi, A., & Suleman, H. (2024). Advances and Challenges in Low-Resource-Environment Software Systems: A Survey. *Informatics*, *11*(4), 90–90. https://doi.org/10.3390/informatics11040090

Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Advanced Research and Reviews*, *18*(3), 050–058. https://doi.org/10.30574/gscarr.2024.18.3.0088

Al Waro'i, M. N. A. L. (2024). False Reality: Deepfakes in Terrorist Propaganda and Recruitment. *Security Intelligence Terrorism Journal (SITJ)*, *1*(1), 41–59. https://doi.org/10.70710/sitj.v1i1.5

Alao, A. I., Adebiyi, O. O., & Olaniyi, O. O. (2024). The Interconnectedness of Earnings Management, Corporate Governance Failures, and Global Economic Stability: A Critical Examination of the Impact of Earnings Manipulation on Financial Crises and Investor Trust in Global Markets. *Asian Journal of Economics Business and Accounting*, *24*(11), 47–73. https://doi.org/10.9734/ajeba/2024/v24i111542

Albahri, A. S., Khaleel, Y. L., Habeeb, M. A., Ismael, R. D., Hameed, Q. A., Deveci, M., Homod, R. Z., Albahri, O. S., Alamoodi, A. H., & Alzubaidi, L. (2024). A

systematic review of trustworthy artificial intelligence applications in natural disasters. *Computers & Electrical Engineering*, *118*, 109409–109409. https://doi.org/10.1016/j.compeleceng.2024.109409

Arigbabu, A. T., Olaniyi, O. O., Adigwe, C. S., Adebiyi, O. O., & Ajayi, S. A. (2024). Data Governance in AI - Enabled Healthcare Systems: A Case of the Project Nightingale. *Asian Journal of Research in Computer Science*, *17*(5), 85–107. https://doi.org/10.9734/ajrcos/2024/v17i5441

Beg, R., Bhardwaj, V., Kumar, M., Muzumdar, P., Rajput, A., & Borana, K. (2024). Unmasking Social Media Crimes. *Wiley Online Library*, 1–26. https://doi.org/10.1002/9781394231126.ch1

Bilal, A., Gjørv, G. H., Lanteigne, M., Brancaleoni, R., Gjørv, J., Gui, D., Kielar, J. K., Aluola, C., & Magalini, S. (2023). Comprehensive security, disinformation, and COVID-19: An analysis of the impacts of mis- and disinformation and populist narratives during the pandemic. *Open Research Europe*, *3*, 209–209. https://doi.org/10.12688/openreseurope.16733.1

Bondurich, S. (2024). *Voice Cloning: The Rising Threat of Audio Deepfakes in Impersonation Scams*. SEVN-X | Cybersecurity. https://www.sevnx.com/blog/post/voice-cloning

Cade, D. L. (2020). *Microsoft's New AI "Authenticator" Spots Manipulated Photos and Videos*. PetaPixel. https://petapixel.com/2020/09/02/microsofts-new-ai-authenticator-spots-manipulated-photos-and-videos/

Chaudhary, M., & Bansal, D. (2022). Open source intelligence extraction for terrorism-related information: A review. *WIREs Data Mining and Knowledge Discovery*, *12*(5). https://doi.org/10.1002/widm.1473

Chen, H., & Magramo, K. (2024). *Finance worker pays out $25 million after video call with deepfake "chief financial officer."*CNN. https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html

Confessore, N. (2018). Cambridge Analytica and Facebook: the Scandal and the Fallout so Far. *The New York Times*. https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Criddle, C. (2020). Facebook sued over Cambridge Analytica data scandal. *BBC News*. https://www.bbc.com/news/technology-54722362

Davis, J., Purves, D., Gilbert, J., & Sturm, S. (2022). Five ethical challenges facing data-driven policing. *AI and Ethics*, *2*. https://doi.org/10.1007/s43681-021-00105-9

Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, *99*(101896), 101896. https://www.sciencedirect.com/science/article/pii/S1566253523002129

Dsouza, D. S., Hajjar, A. E., & Jahankhani, H. (2024). Deepfakes in Social Engineering Attacks. *Space Law and Policy*, 153–183. https://doi.org/10.1007/978-3-031-64045-2_8

Fabuyi, J. A., Oluwaseun Oladeji Olaniyi, O. O., Olateju, O. O., Aideyan, N. T., & Olaniyi, F. G. (2024). Deepfake Regulations and Their Impact on Content Creation in the Entertainment Industry. *Archives of Current Research International*, *24*(12), 52–74. https://doi.org/10.9734/acri/2024/v24i12997

Farouk, M. A., & Fahmi, B. M. (2024). Deepfakes and Media Integrity: Navigating the New Reality of Synthetic Content. *Journal of Media and Interdisciplinary Studies*, *3*(9). https://doi.org/10.21608/jmis.2024.275298.1027

FLI. (2024). *Combatting Deepfakes* . Future of Life Institute. https://futureoflife.org/project/combatting-deepfakes/

FoxNews. (2024). *Americans' trust in media is at new record low, Gallup poll finds*. FOX 4 News Dallas-Fort Worth; FOX 4 Dallas-Fort Worth. https://www.fox4news.com/news/media-trust-journalists-news-america-gallup-poll-election-2024

Fragale , M., & Grilli, V. (2024). *Deepfake, Deep Trouble: The European AI Act and the Fight Against AI-Generated Misinformation*. Columbia.edu. https://cjel.law.columbia.edu/preliminary-reference/2024/deepfake-deep-trouble-the-european-ai-act-and-the-fight-against-ai-generated-misinformation/

Gbadebo, M. O., Salako, A. O., Selesi-Aina, O., Ogungbemi, O. S., Olateju, O. O., & Olaniyi, O. O. (2024). Augmenting Data Privacy Protocols and Enacting Regulatory Frameworks for Cryptocurrencies via Advanced Blockchain Methodologies and Artificial Intelligence. *Journal of Engineering Research and Reports*, *26*(11), 7–27. https://doi.org/10.9734/jerr/2024/v26i111311

Gioti, A. (2024). *Advancements in Open Source Intelligence (OSINT) Techniques and the role of artificial intelligence in Cyber Threat Intelligence (CTI)*. Dione.lib.unipi.gr. https://dione.lib.unipi.gr/xmlui/handle/unipi/16306

HBR. (2021). *How Organizations Can Mitigate the Risks of AI.* Harvard Business Review. https://hbr.org/sponsored/2021/12/how-organizations-can-mitigate-the-risks-of-ai

Hudson, M. (2016). *The Panama Papers: Exposing the Rogue Offshore Finance Industry*. ICIJ; International Consortium of Investigative Journalists. https://www.icij.org/investigations/panama-papers/

Huising, R., & Silbey, S. S. (2021). Accountability infrastructures: Pragmatic compliance inside organizations. *Regulation & Governance*, *15*(1). https://doi.org/10.1111/rego.12419

Jedličková, A. (2024). Ethical approaches in designing autonomous and intelligent systems: a comprehensive survey towards responsible development. *AI & Society*. https://doi.org/10.1007/s00146-024-02040-9

Joeaneke, P. C., Kolade, T. M., Val, O. O., Olisa, A. O., Joseph, S. A., & Olaniyi, O. O. (2024). Enhancing Security and Traceability in Aerospace Supply Chains through Block Chain Technology. *Journal of Engineering Research and Reports*, *26*(10), 114–135. https://doi.org/10.9734/jerr/2024/v26i101294

Joeaneke, P. C., Val, O. O., Olaniyi, O. O., Ogungbemi, O. S., Olisa, A. O., & Akinola, O. I. (2024). Protecting Autonomous UAVs from GPS Spoofing and Jamming: A Comparative Analysis of Detection and Mitigation Techniques. *Journal of*

Engineering Research and Reports, *26*(10), 71–92. https://doi.org/10.9734/jerr/2024/v26i101291

John-Otumu, A. M., Ikerionwu, C., Olaniyi, O. O., Dokun, O., Eze, U. F., & Nwokonkwo, O. C. (2024). Advancing COVID-19 Prediction with Deep Learning Models: A Review. *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria, 2024*, 1–5. https://doi.org/10.1109/seb4sdg60871.2024.10630186

Joseph, S. A. (2024). Balancing Data Privacy and Compliance in Blockchain-Based Financial Systems. *Journal of Engineering Research and Reports*, *26*(9), 169–189. https://doi.org/10.9734/jerr/2024/v26i91271

Kalpokas, I., & Kalpokiene, J. (2022). Deepfakes. In *SpringerBriefs in Political Science*. Springer International Publishing. https://doi.org/10.1007/978-3-030-93802-4

Kanojia, S. (2024). Digitalization in Corporations. *Advances in Human and Social Aspects of Technology Book Series*, 227–245. https://doi.org/10.4018/979-8-3693-3334-1.ch008

Khan, A. (2023). The Impact of Social Media on Political Polarization in the United States. *Research Journal for Social Affairs*, *1*(01), 1–7. https://arr.rjmss.com/index.php/21/article/view/3

Kira, B. (2024). When Non-Consensual Intimate Deepfakes Go Viral: The Insufficiency of the UK Online Safety Act. *SSRN*. https://doi.org/10.2139/ssrn.4798664

Klingberg, S. (2022). Countering Terrorism: Digital Policing of Open Source Intelligence and Social Media Using Artificial Intelligence. *Artificial Intelligence and National Security*, 101–111. https://doi.org/10.1007/978-3-031-06709-9_6

Kolade, T. M., Aideyan, N. T., Oyekunle, S. M., Ogungbemi, O. S., & Olaniyi, O. O. (2024). Artificial Intelligence and Information Governance: Strengthening Global Security, through Compliance Frameworks, and Data Security. *Asian Journal of Research in Computer Science*, *17*(12), 36–57. https://doi.org/10.9734/ajrcos/2024/v17i12528

Labbe, M. (2020). *The deepfake 2020 election threat is real, but containable*. SearchEnterpriseAI. https://www.techtarget.com/searchenterpriseai/feature/The-deepfake-2020-election-threat-is-real-but-containable

Lavric, A., Beguni, C., Zadobrischi, E., Căilean, A.-M., & Avătămăniței, S.-A. (2024). A Comprehensive Survey on Emerging Assistive Technologies for Visually Impaired Persons: Lighting the Path with Visible Light Communications and Artificial Intelligence Innovations. *Sensors*, *24*(15), 4834–4834. https://doi.org/10.3390/s24154834

Lees, D. (2023). Deepfakes in documentary film production: images of deception in the representation of the real. *Studies in Documentary Film*, *18*(2), 1–22. https://doi.org/10.1080/17503280.2023.2284680

Li, W., Yigitcanlar, T., Browne, W., & Nili, A. (2023). The Making of Responsible Innovation and Technology: An Overview and Framework. *Smart Cities*, *6*(4), 1996–2034. https://doi.org/10.3390/smartcities6040093

Martínez-Bravo, M. C., Sádaba Chalezquer, C., & Serrano-Puche, J. (2022). Dimensions of Digital Literacy in the 21st Century Competency Frameworks. *Sustainability*, *14*(3), 1867. https://doi.org/10.3390/su14031867

Matli, W. (2024). Extending the theory of information poverty to deepfake technology. *International Journal of Information Management Data Insights*, *4*(2), 100286–100286. https://doi.org/10.1016/j.jjimei.2024.100286

Me, G., & Mucci, M. F. (2024). Identifying Daesh-Related Propaganda Using OSINT and Clustering Analysis. *Advanced Sciences and Technologies for Security Applications*, 97–146. https://doi.org/10.1007/978-3-031-47594-8_6

Min, A. (2023). Artifical Intelligence and Bias: Challenges, Implications, and Remedies. *Journal of Social Research*, *2*(11), 3808–3817. https://doi.org/10.55324/josr.v2i11.1477

Naeeni, S. K. (2023). The Utilitarian Approach to Environmental Law: Balancing Costs and Benefits. *Interdisciplinary Studies in Society, Law, and Politics*, *2*(1), 4–15. https://doi.org/10.61838/kman.isslp.2.1.2

Nannaware, S. C., Pillai, R., & Kate, N. (2024). Deepfakes in Action. *Advances in Business Information Systems and Analytics Book Series*, 71–98. https://doi.org/10.4018/979-8-3693-6890-9.ch004

Neo, R., & Yin, Y. (2023). Of Social Discipline and Control: The Stratified Impact of Fake News and Disinformation on Minorities in Indonesia. *International Journal on Minority and Group Rights*, *31*(1), 1–23. https://doi.org/10.1163/15718115-bja10115

Okon, S. U., Olateju, O. O., Ogungbemi, O. S., Joseph, S. A., Olisa, A. O., & Olaniyi, O. O. (2024). Incorporating Privacy by Design Principles in the Modification of AI Systems in Preventing Breaches across Multiple Environments, Including Public

Cloud, Private Cloud, and On-prem. *Journal of Engineering Research and Reports*, *26*(9), 136–158. https://doi.org/10.9734/jerr/2024/v26i91269

Okoro, Y. O., Ayo-Farai, O., Maduka, C. P., Okongwu, C. C., & Sodamade, O. T. (2024). A REVIEW OF HEALTH MISINFORMATION ON DIGITAL PLATFORMS: CHALLENGES AND COUNTERMEASURES. *International Journal of Applied Research in Social Sciences*, *6*(1), 23–36. https://doi.org/10.51594/ijarss.v6i1.689

Olabanji, S. O., Marquis, Y. A., Adigwe, C. S., Abidemi, A. S., Oladoyinbo, T. O., & Olaniyi, O. O. (2024). AI-Driven Cloud Security: Examining the Impact of User Behavior Analysis on Threat Detection. *Asian Journal of Research in Computer Science*, *17*(3), 57–74. https://doi.org/10.9734/ajrcos/2024/v17i3424

Olabanji, S. O., Olaniyi, O. O., & Olagbaju, O. O. (2024). Leveraging Artificial Intelligence (AI) and Blockchain for Enhanced Tax Compliance and Revenue Generation in Public Finance. *Asian Journal of Economics, Business and Accounting*, *24*(11), 577–587. https://doi.org/10.9734/ajeba/2024/v24i111577

Olabanji, S. O., Oluwaseun Oladeji Olaniyi, O. O., & Olaoye, O. O. (2024). Transforming Tax Compliance with Machine Learning: Reducing Fraud and Enhancing Revenue Collection. *Asian Journal of Economics Business and Accounting*, *24*(11), 503–513. https://doi.org/10.9734/ajeba/2024/v24i111572

Oladoyinbo, T. O., Olabanji, S. O., Olaniyi, O. O., Adebiyi, O. O., Okunleye, O. J., & Alao, A. I. (2024). Exploring the Challenges of Artificial Intelligence in Data Integrity and its Influence on Social Dynamics. *Asian Journal of Advanced Research and Reports*, *18*(2), 1–23. https://doi.org/10.9734/ajarr/2024/v18i2601

Olaniyi, O. O. (2024). Ballots and Padlocks: Building Digital Trust and Security in Democracy through Information Governance Strategies and Blockchain Technologies. *Asian Journal of Research in Computer Science*, *17*(5), 172–189. https://doi.org/10.9734/ajrcos/2024/v17i5447

Olaniyi, O. O., Olaoye, O. O., & Okunleye, O. J. (2023). Effects of Information Governance (IG) on Profitability in the Nigerian Banking Sector. *Asian Journal of Economics, Business and Accounting*, *23*(18), 22–35. https://doi.org/10.9734/ajeba/2023/v23i181055

Olaniyi, O. O., Omogoroye, O. O., Olaniyi, F. G., Alao, A. I., & Oladoyinbo, T. O. (2024). CyberFusion Protocols: Strategic Integration of Enterprise Risk Management, ISO 27001, and Mobile Forensics for Advanced Digital Security in the Modern Business Ecosystem. *Journal of Engineering Research and Reports*, *26*(6), 32. https://doi.org/10.9734/JERR/2024/v26i61160

Olateju, O. O., Okon, S. U., Igwenagu, U. T. I., Salami, A. A., Oladoyinbo, T. O., & Olaniyi, O. O. (2024). Combating the Challenges of False Positives in AI-Driven Anomaly Detection Systems and Enhancing Data Security in the Cloud. *Asian Journal of Research in Computer Science*, *17*(6), 264–292. https://doi.org/10.9734/ajrcos/2024/v17i6472

Olateju, O. O., Okon, S. U., Olaniyi, O. O., Samuel-Okon, A. D., & Asonze, C. U. (2024). Exploring the Concept of Explainable AI and Developing Information Governance Standards for Enhancing Trust and Transparency in Handling Customer Data. *Journal of Engineering Research and Reports*, *26*(7), 244–268. https://doi.org/10.9734/jerr/2024/v26i71206

Papadopoulos, P., Pitropakis, N., & Buchanan, W. J. (2022). Decentralized Privacy: A Distributed Ledger Approach. *Springer EBooks*, 1805–1830. https://doi.org/10.1007/978-3-030-84205-5_58

Pawelec, M. (2022). Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions. *Digital Society*, *1*(2). https://doi.org/10.1007/s44206-022-00010-6

Reid, J. C., Brown, S. J., & Dmello, J. R. (2023). COVID-19, Diffuse Anxiety, and Public (Mis)Trust in Government: Empirical Insights and Implications for Crime and Justice. *Criminal Justice Review*, *49*(2). https://doi.org/10.1177/07340168231190673

Renuka, O., RadhaKrishnan, N., Priya, B. S., Jhansy, A., & Ezekiel, S. (2024). Data Privacy and Protection. *Wiley Online Library*, 433–465. https://doi.org/10.1002/9781394230600.ch19

Romero-Moreno, F. (2024). Deepfake Fraud Detection: Safeguarding Trust in Generative Ai. *SSRN*. https://doi.org/10.2139/ssrn.5031627

Salako, A. O., Fabuyi, J. A., Aideyan, N. T., Selesi-Aina, O., Dapo-Oyewole, D. L., & Olaniyi, O. O. (2024). Advancing Information Governance in AI-Driven Cloud Ecosystem: Strategies for Enhancing Data Security and Meeting Regulatory Compliance. *Asian Journal of Research in Computer Science*, *17*(12), 66–88. https://doi.org/10.9734/ajrcos/2024/v17i12530

Samuel-Okon, A. D., Akinola, O. I., Olaniyi, O. O., Olateju, O. O., & Ajayi, S. A. (2024). Assessing the Effectiveness of Network Security Tools in Mitigating the Impact of

Deepfakes AI on Public Trust in Media. *Archives of Current Research International*, *24*(6), 355–375. https://doi.org/10.9734/acri/2024/v24i6794

Sarraf, S., Kushwaha, A. K., Kar, A. K., Dwivedi, Y. K., & Giannakis, M. (2024). How did online misinformation impact stockouts in the e-commerce supply chain during COVID-19 – A mixed methods study. *International Journal of Production Economics*, *267*, 109064. https://doi.org/10.1016/j.ijpe.2023.109064

SCHIFF, K. J., SCHIFF, D. S., & BUENO, N. S. (2024). The Liar's Dividend: Can Politicians Claim Misinformation to Evade Accountability? ˜*the œAmerican Political Science Review*, 1–20. https://doi.org/10.1017/s0003055423001454

Scott, B. (2023). "Everyone freaks out when the leaks are made": data leaks, investigative journalism and intelligence practice. *Journal of Financial Crime*, *31*(3). https://doi.org/10.1108/jfc-05-2023-0123

Selesi-Aina, O., Obot, N. E., Olisa, A. O., Gbadebo, M. O., Olateju, O. O., & Olaniyi, O. O. (2024). The Future of Work: A Human-centric Approach to AI, Robotics, and Cloud Computing. *Journal of Engineering Research and Reports*, *26*(11), 62–87. https://doi.org/10.9734/jerr/2024/v26i111315

Shahbazi, M., & Bunker, D. (2024). Social media trust: Fighting misinformation in the time of crisis. *International Journal of Information Management*, *77*(102780), 102780–102780. https://doi.org/10.1016/j.ijinfomgt.2024.102780

Singh, T. (2024). AI-Driven Surveillance Technologies and Human Rights: Balancing Security and Privacy. *Smart Innovation, Systems and Technologies*, *392*, 703–717. https://doi.org/10.1007/978-981-97-3690-4_53

Smuha, N. A., Ahmed-Rengers, E., Harkens, A., Li, W., MacLaren, J., Piselli, R., & Yeung, K. (2021, August 5). *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act*. Papers.ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991

Subrahmanyam, S. (2024). Collaboration and Collective Action: Addressing the Deepfake Challenge as a Community. *Advances in Business Information Systems and Analytics Book Series*, 143–172. https://doi.org/10.4018/979-8-3693-6890-9.ch007

Surjatmodjo, D., Unde, A. A., Cangara, H., & Sonni, A. F. (2024). Information Pandemic: A Critical Review of Disinformation Spread on Social Media and Its Implications for State Resilience. *Social Sciences*, *13*(8), 418–418. https://doi.org/10.3390/socsci13080418

Tariq, M. U. (2024). Harnessing Generative AI for Enhanced Web Application Security. *Advances in Web Technologies and Engineering*, 161–192. https://doi.org/10.4018/979-8-3693-3703-5.ch008

Val, O. O., Kolade, T. M., Gbadebo, M. O., Selesi-Aina, O., Olateju, O. O., & Olaniyi, O. O. (2024). Strengthening Cybersecurity Measures for the Defense of Critical Infrastructure in the United States. *Asian Journal of Research in Computer Science*, *17*(11), 25–45. https://doi.org/10.9734/ajrcos/2024/v17i11517

Val, O. O., Olaniyi, O. O., Selesi-Aina, O., Gbadebo, M. O., & Kolade, T. M. (2024). Machine Learning-enabled Smart Sensors for Real-time Industrial Monitoring: Revolutionizing Predictive Analytics and Decision-making in Diverse Sector.

Asian Journal of Research in Computer Science, *17*(11), 92–113. https://doi.org/10.9734/ajrcos/2024/v17i11522

Williamson, S. M., & Prybutok, V. (2024). The Era of Artificial Intelligence Deception: Unraveling the Complexities of False Realities and Emerging Threats of Misinformation. *Information*, *15*(6), 299. https://doi.org/10.3390/info15060299

Yadav, A., Kumar, A., & Singh, V. (2023). Open-source intelligence: a comprehensive review of the current state, applications and future perspectives in cyber security. *Artificial Intelligence Review*, *56*. https://doi.org/10.1007/s10462-023-10454-y

Zulkiflee, E., Shahril, M., Ghani, A., Nirwandy, N., & Noordin, M. (2024). DIGITAL ERA OSINT: FORMULATING SPECIAL NATIONAL INTELLIGENCE ESTIMATES THROUGH OPEN SOURCES. *Journal of Media and Information Warfare*, *17*(2), 100–116. https://jmiw.uitm.edu.my/images/Journal/Vol17No2/8DIGITALERA.pdf