

Deepfake crimes in the age of AI: A bibliometric study of emerging risks and research trends

Mishaal Al-Raggad^{a*}

^aJadara University, Irbid, Jordan

CHRONICLE

Received August 10, 2025
Received in revised format
September 12, 2025
Accepted October 28 2025
Available online
October 29 2025

Keywords:

Deepfake crimes
Artificial intelligence
Disinformation
Misinformation
Cybercrime
Deepfake detection
Emerging risks
Bibliometric analysis

ABSTRACT

This study presents a bibliometric analysis of scholarly research on deepfake crimes in the age of AI, examining emerging risks and research trends between 2019 and 2025. Drawing on 349 publications from the Web of Science Core Collection, the study uses VOSviewer to map publication patterns, collaboration networks, and thematic clusters. The results reveal that research on deepfake crimes is very interdisciplinary, as AI-based detection studies are increasingly intersecting with legal, social, and criminal discussions. The cluster analysis highlights that while significant progress has been made in technical detection methods, there are still serious gaps in addressing the societal harms of misinformation and non-consensual content. These results indicate that the field is not only growing rapidly, but is also moving towards a more integrated agenda that combines technological innovation with ethical and regulatory considerations. The study contributes theoretically by framing deep counterfeiting as technological and criminal phenomena, and practically by providing insights to policy makers, researchers and practitioners seeking to mitigate the societal and security risks of artificial media.

© 2026 by the authors; licensee Growing Science, Canada.

1. Introduction

Artificial intelligence (AI) is transforming nearly every area of modern life. AI systems are now powering many of the services and decisions we rely on, from healthcare and finance to classrooms and streaming platforms (Hamdan et al., 2021; Hammouri et al., 2024). However, these developments introduce additional concerns. The most surprising—and disturbing—emergence of "deepfake": Machine-learning technology such as generative adversarial networks (Kalpokas & Kalpokiene, 2022). What began as a technical curiosity with creative potential quickly showed a darker aspect. Deepfakes are now frequently associated with misinformation, online harassment, and other forms of criminal misconduct, and they are causing worry among scholars, regulators, and the public (Shirish & Komal, 2023; Meskys et al., 2020).

Deepfakes are important because they attack the foundation of truth, faith, and safety in the digital era. They can spread fake news, enable fraud and identity theft, and non-consent can facilitate sexual content (Mansoor, 2024). In doing so, they embody the "dual-use dilemma" of artificial intelligence: technologies that could be used for innovation, creativity, and entertainment but are just as easily turned toward harm (Große & Sundberg, 2025). The researchers and IT scholars caution that there may be significant wider political and social repercussions, ranging from encouraging gender-based violence and harassment to compromising democratic processes (Yadlin-Segal & Oppenheim, 2021). By undermining public trust in genuine evidence, the so-called "liar's earnings", which enable violators to reject and decline original recordings as fraudulent, simply increase the danger. Because of this, a multidisciplinary approach is needed to address deepfakes, which are situated at the intersection of computer science, media studies, criminology, and cybercrime (Lin, 2025; Vaccari & Chadwick, 2020).

Even though this topic has received a lot of attention these days, research on deepfake crimes is still dispersed and inconsistent. Several literary compilations on three primary topics. The initial and most important attention is on technological features, such as counteractors and tools. The second is the socio-political ramifications, which include disinformation, public opinion manipulation, and distribution. Third, the new series examines criminal and legal issues, such as non-conscious pornography,

* Corresponding author

E-mail address: M. Al-Raggad (m.alraggad@jadara.edu.jo) YouTube Movie Link: <https://www.youtube.com/watch?v=LtUSiy1LQ>

privacy, and regulatory alternatives. Although each stream offers insightful information, they frequently evolve independently of one another, with only a passing conversation in between.

Another gap is the lack of a comprehensive database overview in this area. Since 2019, the volume of publications has increased, however most of the current reviews are narrative or conceptual, not quantitative. Without a systematic map of the research landscape, it is easy to miss important connections, unexplored topics and opportunities for cross-disciplinary collaboration.

This study seeks to close the gap. As the first large-scale bibliometric investigation of deepfake crimes, we use the Web of Science database as a high-quality data source. We have three things to contribute. In order to provide and draw a comprehensive picture of the field's expansion and global dispersion, we first map publishing trends, citation patterns, and contributions at the global level. Second, we distinguish three primary theme groups: (1) social media disinformation and deepfakes; (2) non-consensual pornography, face manipulation, and deepfake technology; and (3) AI-based security and detection methods. The way these groups are mapped reveals where scientific energy is focused and how research is carried out. Third, we highlight and draw attention to new opportunities and shortcomings, such as the dearth of humanities-based approaches, the small number of partnerships from underdeveloped nations, and the continuous combat between detection systems and creative tools.

When taken as a whole, these findings offer a resource to researchers, scholars, and professionals. For scholars, the study highlights important directions, trends, collaborations, and boundaries, assisting in building future research on strong foundations and exploring underserved regions. Moreover, it illustrates to experts and governments why combating deepfake crimes requires a coordinated strategy that includes technical, legal, social, and sanitation tactics. This text-to-observation in the digital era emphasizes the critical need for an all-encompassing, multidisciplinary research agenda to protect trust, security, and human dignity by drawing attention to the perils of deepfakes in in-depth talks on the risks and dangers of artificial intelligence in deepfake crimes.

2. Methodology

2.1 Data Source and Search Strategy

The study investigates deepfake crimes in the AI era, using data from the Web of Science Core Collection (Clarivate Analytics). The web of science database was chosen because it offers rigorous and important sequencing, wide international coverage, and solid metadata qualities, making it a dependable resource for publishing, transparent and copyable books (Van Eck & Waltman, 2020). Therefore, working with a database that can handle such cross-disciplinary data was critical, as deepfake crime involves areas such as computer science, law, criminology, and media studies.

The research was conducted in September 2025 so that the data set reflects the latest scholarship. We focused on publications between 2019 and 2025. The start date is important: 2019 is when “crimes of deepfake” began to appear constantly as a topic in the network of science, coinciding with the maturation of technology and increasing public and academic concern about its misuse in misinformation, cybercrime and non-consensual content. Expanding the scope of research until 2025 has given us six years of material and allowed us to capture how the field is developing.

To identify the relevant documents, we have built a search query that combines terms related to the technology itself and those that indicate criminal or deceptive uses.

The final query was TS= (deepfake OR “deep fake” OR “synthetic media” OR “AI-generated video” OR “face swap”) AND TS= (crime OR fraud OR cybercrime OR misinformation OR disinformation OR "information disorder" OR "non-consensual pornography" OR sextortion OR "identity theft").

This strategy is designed to capture research that explicitly links deepfake techniques, criminal activities, digital deception and broader forms of information disorder, ensuring objective alignment with the focus of the study.

2.2 Search Criteria

We then applied inclusion criteria to ensure that the dataset is consistent and relevant. Only English-language publications were retained to allow reliable analysis of quotations and keyword interpretation. To focus on peer-reviewed work, we limited the sample to journal articles and conference proceedings indexed in an appropriate network of science categories. Other types of publications were considered only if they clearly fall within the scope of the study.

Book chapters were excluded. Although they can be valuable sources, they often lack standardized bibliometric data such as the number of citations and structured keywords, which are essential for robust mapping. We also verified that each included study explicitly addressed deepfake crimes in the age of artificial intelligence, rather than just mentioning deepfakes in irrelevant contexts such as entertainment or artistic experiments.

2.3 Data Collection and Filtering

Our initial discovery of the web of science core collection produced 433 publications from 2019 to 2025. To ensure that we were honest that we sincerely speak of the crimes of deepfakes in the era of Artificial Intelligence, we applied several rounds of filtering. We first confirmed the deadline 2019-2025. This period not only marks the first appearance of research indexed on deepfakes offenses, but also covers the years when public and regulatory concerns about synthetic media were growing rapidly.

We then limited the sample to English language papers to ensure consistency and avoid translation issues that could distort keyword analysis. Next, we excluded materials outside relevant Web of Science subject categories, along with book chapters, which rarely include the standardized bibliometric information needed for network analysis.

After applying these filters, the dataset was reduced from 433 to 351 publications. This refined collection provides a focused, high-quality basis for our analysis, representing research directly related to crimes that support deepfakes such as fraud, cybercrime, disinformation, identity theft and non-consensual content, while excluding purely peripheral or descriptive references to deepfakes.

2.4 Software Utilized for Data Analysis

The final set of 351 records was exported from the Web of Science database into a structured bibliometric database containing titles, abstracts, authors and researchers' details, institutional affiliations, keywords, and citation data. To map the intellectual structure and track emerging patterns and trends in this field, we used VOSviewer, a software package built for constructing, analyzing, and visualizing bibliometric networks.

Because VOSviewer can handle big information and create comprehensible maps of intricate relationships, it was selected. Using it, we produced co-citation maps that highlighted the scholarly works and journals that shaped the field, co-authorship networks that demonstrated collaboration among scholars, researchers, universities, and academic institutions, and keyword co-occurrence networks that exposed recurrent themes and concepts (Abu Huson et al., 2024; Al-Raggad et al., 2025). This made it possible for us to observe and notice not just who is involved in the deepfake-crime discussion, but also how concepts and approaches are transferring across academic fields.

Using VOSviewer's clustering technique, we categorized and classified linked articles and keywords into thematic clusters, allowing us to discover prevalent and emerging problems such as misinformation, cybercrime, identity theft, and non-consensual explicit content associated with synthetic and fake media. We also applied temporal overlay visualisations to see how research interests have shifted over the 2019–2025 period — for example, from early technical discussions of AI-generated video manipulation to more recent legal, ethical and criminological perspectives.

By combining network and temporal analyses, VOSviewer gave us a rich, visual overview of the research landscape (Van Eck & Waltman, 2010). This approach highlights collaboration patterns, thematic concentrations and gaps in the literature, offering a clearer picture of how scholarly attention to deepfake crimes has evolved and where new risks and debates are emerging.

3. Results

3.1 Publication over Time

The bibliometric analysis shows how quickly academic interest in deepfake crimes has grown. As Figure 1 illustrates, there was virtually no research on this topic before 2019 — the year Web of Science first began indexing papers that explicitly link deepfake technology with criminal misuse. From 2019 to 2021, the number of studies has steadily increased as scientists began to recognize deepfakes as a serious issue, especially about misinformation, fraud and explicit non-consensual content.

The growth became even steeper after 2022, with 2024 marking the highest point in more than 100 publications in one year. This growth reflects public, political and legal debates that have been furious about the regulation of artificial intelligence and digital safety at the same time. Although the count in 2025 decreased slightly, it was significantly higher than the level of early duration, which indicates that research on deepfakes crimes has shifted from a short-term trend to a stable academic field.

The citation strengthens this perception of maturity. Dataset produced 2,861 paragraph quotes (except 2,680 self-consciousness) and a total of 4,457 quotes (except 3,908 except self-condoms). Each publication, average, approximately 12.8 quotes, and the area reached H-Index of 33-evidence of productivity and impact among the leading contributors.

Another striking pattern is the difference between the number of publications and the number of quotes. While the publication number was at its peak in 2024, quotes reflect cumulative effects and will continue to grow, with high impact studies from 2019 to 2022 are still shaping the interaction today with high impact studies. Taken together, these trends show how research on deepfake crimes quickly emerged and then consolidated into a critical interdisciplinary field of study.

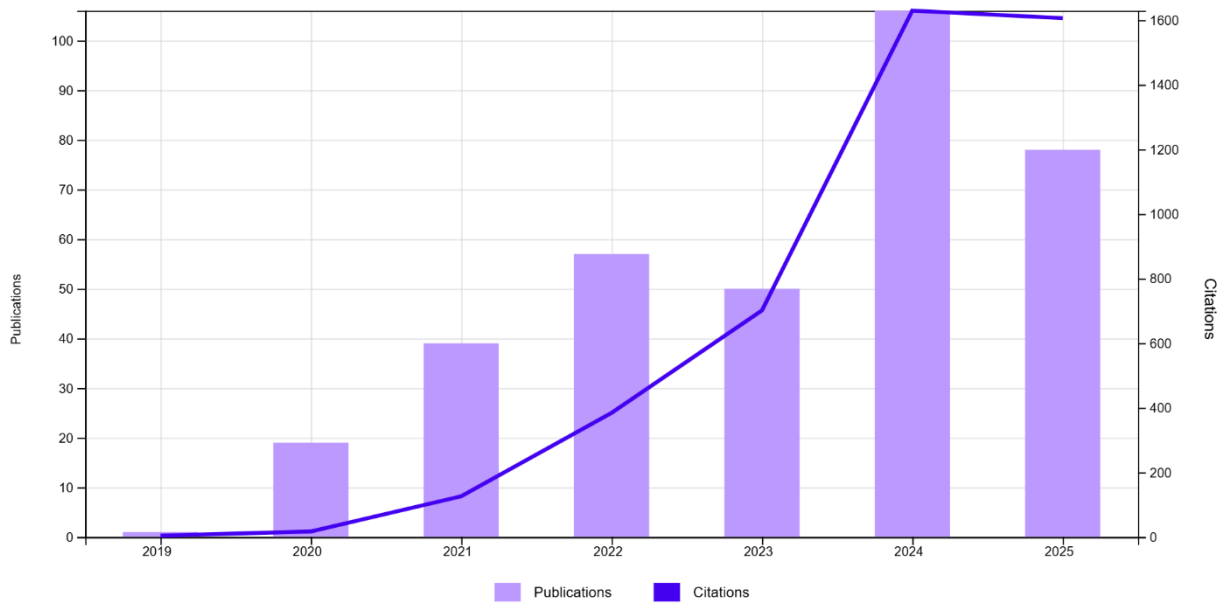


Fig. 1. Annual distribution of publications and citations (2019–2025)

Table 1 shows how the publications in our dataset are distributed across the different Web of Science indexes. Most of the work appears in scientific and technical outlets. The largest share comes from the Conference Proceedings Citation Index – Science (CPCI-S) with 135 records (38.5%), followed by the Science Citation Index Expanded (SCI-EXPANDED) with 119 records (33.9%). Together these two indexes account for more than 70 per cent of the total, reflecting how research on deepfake crimes is still largely rooted in computer science, engineering and related technological fields, and often presented at conferences or in high-impact journals.

The Emerging Sources Citation Index (ESCI) contributes 57 records (16.2%). This is consistent with the field’s youth and rapid expansion, as newer journals begin to cover synthetic media, AI ethics and cybercrime. The Social Sciences Citation Index (SSCI) adds a further 48 records (13.7%), signaling that deepfake crimes are increasingly recognized not just as a technical challenge but also as a social, legal and criminological issue.

By contrast, the Arts & Humanities Citation Index (A&HCI) and the Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH) each contain only four records (1.1%). This very limited presence suggests that cultural, philosophical and humanities-oriented discussions of deepfakes remain relatively underdeveloped compared with the technical and social science perspectives.

Table 1

Distribution of publications across Web of Science Indexes

Web of Science Index	Record Count	% of 351
Conference Proceedings Citation Index – Science (CPCI-S)	135	38.46%
Science Citation Index Expanded (SCI-EXPANDED)	119	33.90%
Emerging Sources Citation Index (ESCI)	57	16.24%
Social Sciences Citation Index (SSCI)	48	13.68%
Arts & Humanities Citation Index (A&HCI)	4	1.14%
Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH)	4	1.14%
Total	351	100%

3.3 Most Contributed Publication Titles and Publishers

Our analysis of publication venues shows that research on deepfake crimes appears across a surprisingly wide range of outlets. As Fig. 2 illustrates, the single largest contributor is IEEE Access, which alone accounts for 18 papers (5.1% of the dataset). This shows that IEEE Access has become an important hub for research at the intersection of artificial intelligence, cybersecurity, and digital ethics in the digital era.

Other prominent outlets are the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops with seven papers (2%), Lecture Notes in Computer Science, Multimedia Tools and Applications, and the Proceedings of SPIE, each publishing five publications (1.4%). A further cluster of outlets — such as ACM Transactions on Multimedia Computing, Communications and Applications, Electronics, Expert Systems with Applications, and New Media & Society — contributed four papers apiece (1.1%), reflecting both the technical and social-science sides of deepfake research.

This spread of venues shows that although high-profile technical outlets still dominate, interdisciplinary journals and conference proceedings are beginning to play a larger role, especially those linking deepfake crimes with wider debates on misinformation, media, and criminology. The presence of journals like PLOS 1 and New Media & Society underlines the growing recognition of deepfake crimes as not just a computing challenge but also a pressing social issue.

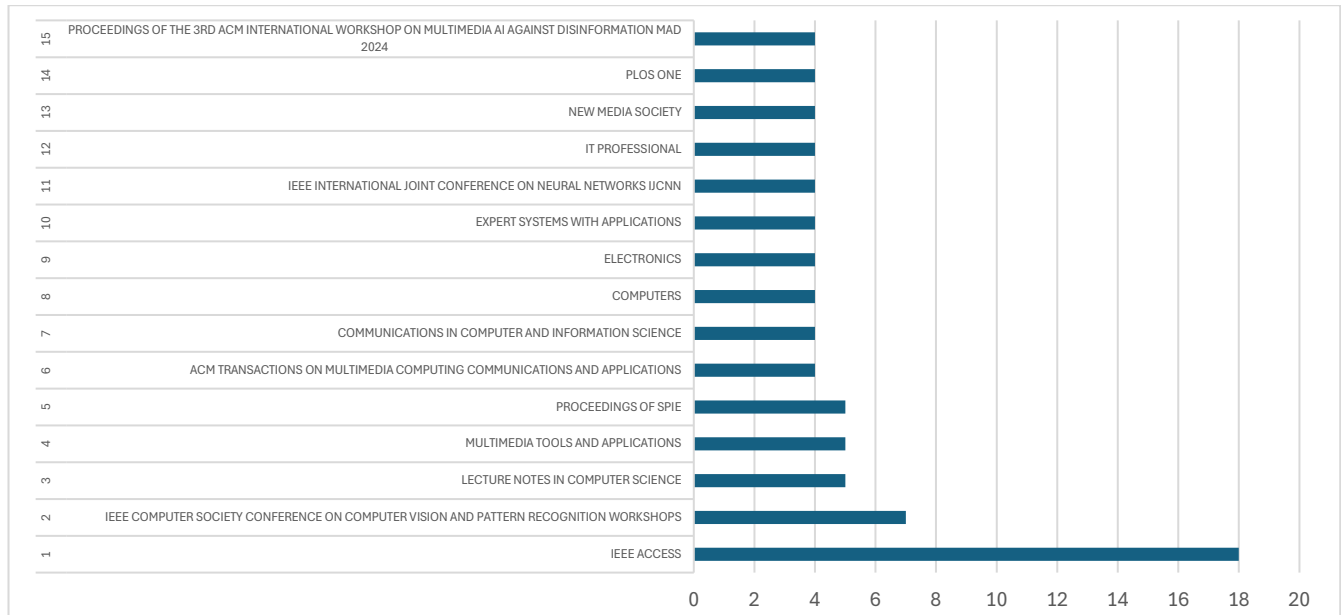


Fig. 2. Leading publication titles

In Fig. 3, the split of publishers demonstrates how substantially this research topic is entrenched in technology-focused publications. IEEE leads with 99 publications (28.2%), highlighting its critical significance in computer science, information security, and engineering research. Springer Nature (43 articles, 12.3%) and the Association for Computing Machinery (ACM) (41 publications, 11.7%) are close behind, serving as key forums for cutting-edge research in multimedia computing, cybersecurity, and AI applications. Elsevier (27 articles, 7.7%) and MDPI (20 publications, 5.7%) also make visible and vital contributions, enhancing and promoting the field's growth through open-access models and interdisciplinary content.

A number of additional publishers make modest but substantial contributions. SAGE (18, 5.1%) and Taylor & Francis (16, 4.6%) offer venues geared toward the social sciences, communication, and criminology. Wiley (8, 2.3%), Oxford University Press (5, 1.4%), and Emerald Group Publishing (4, 1.1%), demonstrate that conventional academic publishers are increasingly concerned and interested in AI ethics and misinformation and disinformation, and cybersecurity. Specialist publications such as SPIE - International Society for Optical Engineering (6, 1.7%) and conference-focused publishers such as Academic Conferences Ltd (4, 1.1%) emphasize the continuous importance of applied and practitioner contributions. Taken together, this distribution demonstrates that research on deepfake crimes in the age of AI remains mostly in technological and applied science disciplines, but is gradually expanding into social science, law, criminal law, and multidisciplinary areas.

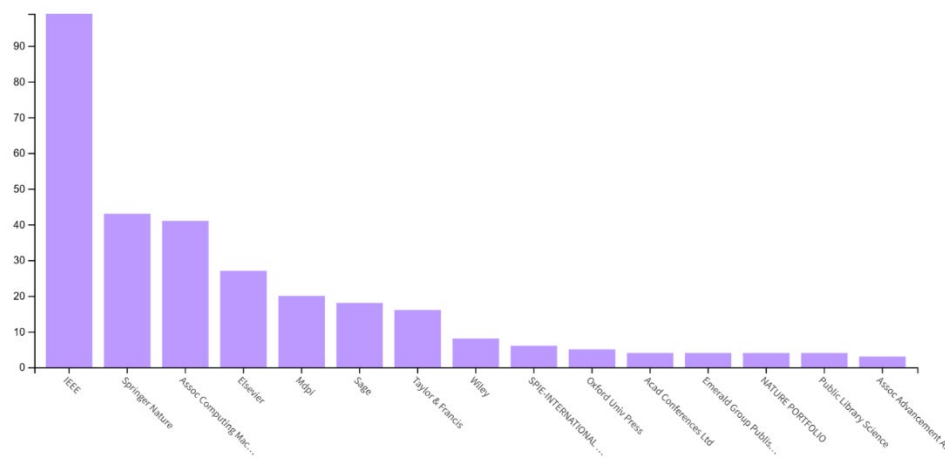


Fig. 3. Leading publishers of research

3.4 Keyword Mapping and Cluster Analysis

A good glimpse into how deepfake criminal research is carried out in the AI era is offered by keyword mapping. Not only can we observe which author and researcher keywords are most frequently employed, but we can also see how they cluster in more and wider general areas. Three primary areas of attention were identified in the dataset we used to create a co-event network using VOSviewer. The network's total of 166 connection strengths and 60 links demonstrated the frequency and strength of the concepts' attachment to the publications, demonstrating the region's multidisciplinary character.

In the center of the map, "deep fake", the most frequently used and most connected keywords sit. Its position as the network of the network shows how it is discussed in very different strands of research, from a computer-science point of view, to work in social science on misinformation, and criminal studies of disadvantages such as non-coordination of clear materials. Three groups or clusters identified by analysis gave a snapshot of how the area has developed:

- Information disorder and digital manipulation. This cluster brings and highlights keywords such as disintegration, misinformation and fake news, reflecting concerns about public trust, democratic discourse and online information manipulation.
- Technical Identification and AI-operated Countermeasures. Here, words such as Artificial Intelligence, Deep Learning, feature extraction, and Deepfake Detection are dominated and highlighted, reflecting the technical "Arms Race" between synthetic media makers and the tools designed for it.
- Disadvantages on the social, moral, and criminal levels. Research violations, identity theft, and research hazards are highlighted by keywords like pornography, technology, and media. These terms are frequently used in discussions about law and policy.

When taken as a whole, these groups imply that research on deepfakes and crime is intersectional. Many studies currently bridge the boundaries between computer science, law, criminal science, criminal law, communication, and social science, such as combining analysis of disinformation operations with AI-based identity approaches. These fields are quickly and increasingly overlapping. This emphasizes how technical, social, and moral difficulties are transformed into the study of ingrained transgressions and the interdisciplinary nature of the interaction region.

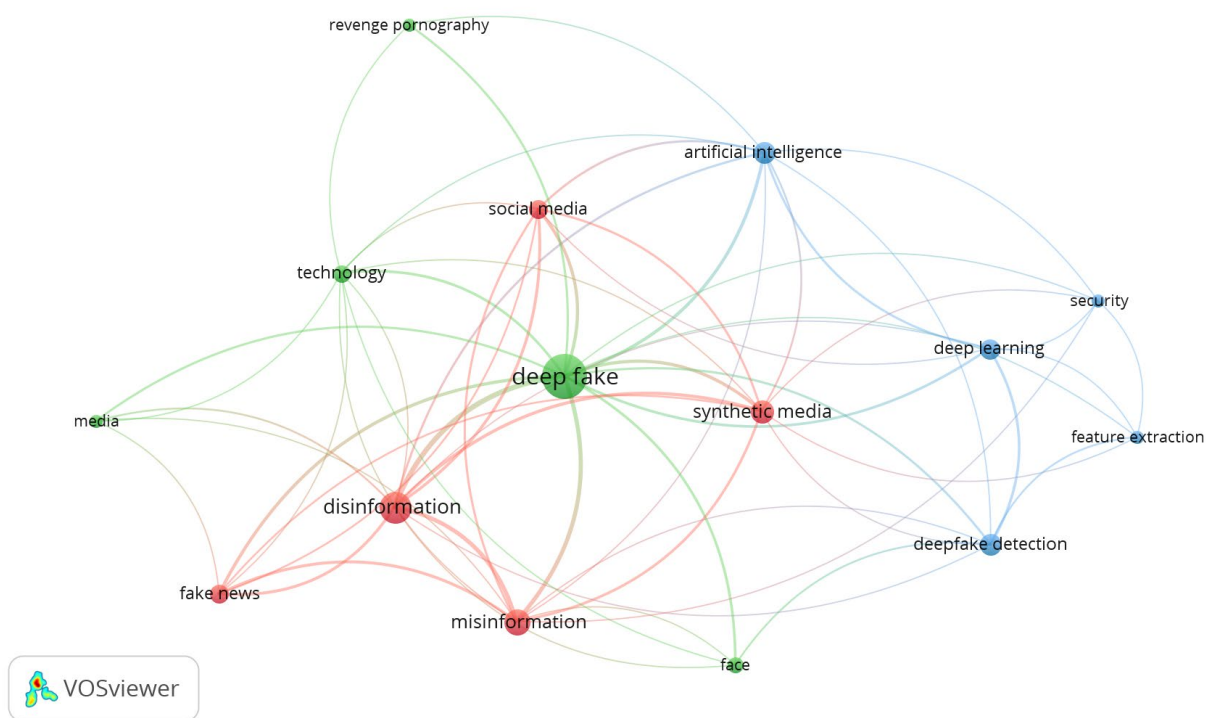


Fig. 4. Visualization of Keyword Relationships

3.5 Interpretation of Clusters Content

Cluster RED: The Role of Social Media in Spreading Fake News and Disinformation

The first group (RED) highlights how deep fakes are combined with a widespread problem of information disorder, especially through fake news, disintegration, and the spread of disinformation via social media and unreliable websites. Social media

platforms are now the main channels of global communication, which play an important and vital role in this ecosystem, serving as catalysts and amplifiers for fake news and liars. Their design makes it easier to make, share, and interact with material, which helps explain why the events related to deepfakes are spreading so quickly.

Researchers' and academics' growing concern over the social and political ramifications of deepfake techniques in the digital age is reflected and highlighted in this cluster. These tools and methods have the potential to erode public trust in information, skew democratic procedures, and exacerbate social division (Alanazi et al., 2025). These impacts have been exacerbated by the digital revolution's quick speed, which has altered how information is created, disseminated, and consumed (Huson et al., 2024). Because deepfakes can improve online echo chambers on social media and capitalize on preexisting biases, they are very powerful (Sophia, 2025).

According to studies, there is a growing trend of the deliberate use of phony images and films for political ends, ranging from leader-created speeches to phony photographs meant to incite unrest or harm a person's reputation (Sylvester, 2020). Interestingly, the threat does not always rely on highly sophisticated and advanced technology; even suggesting that a video "may be deeply fake" can create doubt and erode trust in real content, a phenomenon known as "liar's earnings" (Vaccari & Chadwick, 2020).

This group is also concerned about the practical and psychological repercussions of deepfakes in the era of artificial intelligence in deepfake crimes. Regular exposure to fake and artificial media can lead to audience disappointment, reinforce skepticism about both fake and real news, and foster conspiracy theories (Cover et al., 2022). These problems are exacerbated by social media algorithms that favor sensual or provocative content, which makes it easier for deep fakes to deceive and manipulate in the digital era (Gieseke, 2020). This group generally stresses the importance of providing diverse answers. Developing detection algorithms is only one aspect of combating deepfake-driven misinformation; other strategies include platform-level moderation, technology solutions, legislative frameworks, and public education.

Table 2 shows the top five keywords of this group, along with their frequency, links to other keywords, and overall link strength, showing that the main concern here is the manipulative power of deepfakes in the digital public domain. Social media, in particular, acts as a means of widespread dissemination of disinformation, which makes deep falsification a threat not only to individual victims but also to the integrity of democratic institutions and public confidence in online communication.

Table 2

Core Keywords in Cluster 1 (Red)

Keywords (Cluster 1)	Links	Total Link Strength	Occurrences	Cluster Color
Disinformation	11	50	16	Red
Misinformation	11	38	11	Red
Synthetic Media	10	27	9	Red
Fake News	7	22	6	Red
Social Media	8	24	6	Red

Cluster GREEN: Deepfake Technology and the Manipulation of Faces: Media Challenges and the Rise of Revenge Pornography

The second (green) group shifts the focus from political and societal disinformation to the deep personal damage caused by deepfake technology, especially those involving face manipulation and non-consensual content. Central to this collection are keywords such as technology, media, face, and revenge pornography, highlighting how deepfakes weaponize sexual exploitation, harassment, and identity abuse.

Research shows an alarming increase in non-consensual pornography created using deepfake technology. The faces of the victims are mostly placed on a clear photo or video without consent, which leads to serious violations of privacy and permanent psychological, social, and reputational damage (Okolie, 2023). Contrary to misinformation, which mainly affects public perception and collective belief, Deepfech pornography directly harasses individuals, often with disastrous effects on their safety and dignity (Yadav et al., 2025).

In addition to pornography, this collection also examines the widespread use of facial processing techniques in the media. Scholars have argued that Deepfakes stain the line between reality and imagination, challenging traditional ideas of authenticity in journalism, entertainment, and advertising. While manipulated faces can be used for satire or art, similar equipment can be misused to build identity, fraud, or even for the construction of fake evidence in legal disputes (Farouk & Fahmi, 2024; Nannaware et al., 2025).

One of the main concerns here is the moral and legal challenge of addressing these crimes. Sexual abuse associated with deepfakes is present in a legal gray field, as many courts lack clear laws against non-consensual synthetic pornography (McGLYNN & Toparlak, 2025). Researchers emphasized the need for strong legal protection and affected approaches, including digital removal mechanisms, legal treatment, and public awareness campaigns (Tuliakov, 2023). In short, this collection makes it clear that the crime of deep falsezation is not only an abstract threat to the integrity of information, but

they are a direct attack on personal privacy and autonomy. To address them, the misuse of AI-borne facial manipulation requires policies, technologies, and advocacy measures to protect vulnerable people.

Table 3 shows the five most important keywords in Group 2, which reflect the focus of this group. Discussion here emphasizes the dual nature of deep fakes: they represent innovative development in media and artificial intelligence-based facial synthesis, but they can also be made into weapons to cause serious personal losses. Thus, to understand the crimes of deepfakes, not only technical knowledge is required, but also a moral, legal and social outlook with immediate attention to the safety and political intervention of the victims.

Table 3

Core Keywords in Cluster 2 (Green)

Keywords (Cluster 2)	Links	Total Link Strength	Occurrences	Cluster Color
Deep Fake	14	72	33	Green
Technology	10	13	5	Green
Face	5	9	4	Green
Media	5	8	3	Green
Revenge Pornography	3	5	3	Green

Cluster BLUE: Harnessing Deep Learning for Deepfake Detection and Security Enhancement

The third group (blue) focuses on technical and engineering research that aims to develop a defense against deepfakes crimes. The major words in this group- such as artificial intelligence, deep education, convenience, safety, and deep fake detections, focused their focus before focusing on technical strategies before focusing on technical strategies.

Research here often applies machine learning and computer vision techniques to detect fake videos and subtle artifacts in photos. Analysis of facial asymmetry in methods, discrepancies, or eyelids in eye activities, and mismatch in hearing and visual synchronization (Farid, 2022). More advanced methods use conventional neural networks and generative adversarial network models based on the detection of techniques used to generate deep falsehood. Since generative models are constantly improving, discovery should continuously adapt their methods to maintain a running arms race-makers with realistic outputs constantly with realistic outputs (Peck et al., 2023).

Another attention in this group is the integration of the detection devices in the comprehensive safety structure. Scholars have discovered the integration of these systems in social media platforms, forensic software, and legal evidence verification processes. This approach frames deep fake not only as a media challenge, but also as a concern for cybersecurity and national security, in view of the possibility of misuse of publicity, terrorism, and cyber warfare (Digmelashvili, 2023).

Beyond the detection, the group highlights the preventive strategies, including the original media, the blockchain-based certification systems, and the source of the material (Al-Raggad et al., 2024). The purpose of these methods is not only to identify fake materials, but also to restore confidence in the original media by providing verification of original verification.

In general, this collection displays the dual role of Artificial Intelligence in the era of Deep Fake: it is possible to create extremely realistic artificial media from Artificial Intelligence, but also provides the most promising tools to protect against them. Research suggests that the fight against deepfakes crimes requires continuous innovation, interdisciplinary cooperation and continuous investment in research and development. Table 4 shows the five most important keywords in this group.

Table 4

Core Keywords in Cluster 3 (Blue)

Keywords (Cluster 3)	Links	Total Link Strength	Occurrences	Cluster Color
Artificial Intelligence	10	22	8	Blue
Deep Learning	8	15	8	Blue
Deepfake Detection	7	15	7	Blue
Feature Extraction	5	6	3	Blue
Security	6	6	3	Blue

Together, these three clusters provide a comprehensive view of deepfake crimes in the AI era. They reveal how the issue is approached from multiple angles: the spread of disinformation on social media (red), the direct personal harms caused by facial manipulation and non-consensual content (green), and the development of technical detection and security solutions (blue).

3.6 The Most Contributed Countries

The bibliometric analysis of contributing countries sheds light on both the global distribution of research on deepfake crimes and the patterns of international collaboration. A co-authorship map generated with VOSviewer reveals a network of 208 links, with a total link strength of 4,830, showing that the study of deepfake crimes has quickly become a worldwide research focus, supported by strong collaborative ties across continents.

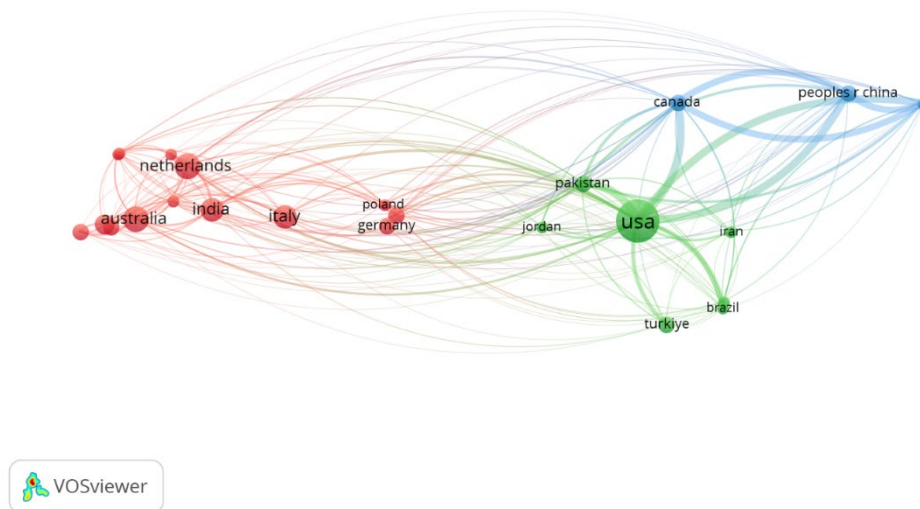


Fig. 5. Country collaboration network

Fig. 5 shows three distinct clusters of international cooperation. The Red Group, including countries such as the Netherlands, Australia, India, Italy and Germany, highlights strong contributions in Europe and Asia-Pacific with significant cross-country partnerships. A diversified network linking North America with research groups in the Middle East and South America is reflected and highlighted in the green group, which is led by the United States and has linkages to Pakistan, Jordan, Iran, Brazil, and Turkey. China and Canada are members of the Blue Group, which highlights the growing significance of Chinese research in deepfakes crimes and its close collaboration with other technologically advanced nations.

One of the main hubs for these deepfake crime and technology contributors is the United States. With 14 publications and 766 quotes, it shows both activity and importance in shaping the region. Its 2,400 overall connection strength and 23 linkages show a great deal of global cooperation as well as the addition of several regional research communities in these academic areas. The United States' dual contributions—shaping the conversation on politics, media manipulation, privacy, criminal law, and digital ethics—as well as creating technological solutions like enhanced detection and AI-powered operations—have contributed to the country's crucial position.

China is a major player, especially in the Blue Group, and has strong relations to Canada and other countries that make large investments in technology and deepfake research. The country has made large expenditures in artificial intelligence and digital forensics, which has increased its productivity. Meanwhile, European countries such as Germany, Italy, and the Netherlands contribute significantly to the Red Group. Unlike North America and China, which tend to focus more on technical aspects of deepfakes, they often emphasize the ethical, legal, and societal aspects of these fakes. Emerging contributions from outside of established centers, like Pakistan, Jordan, and Brazil, are featured in the Green Group, underscoring the global expansion of study interest in information technology and deepfake crimes. Their participation indicates that the threats posed by synthetic media are acknowledged in a variety of political, cultural, and legal contexts, highlighting the significance of tackling deepfake crimes on a worldwide scale.

4. Discussion

The results of this bibliometric study offer a comprehensive evaluation of the emerging research on deepfakes in the era of artificial intelligence. The study charts the thematic groups of intellectual structure, cooperation patterns, and literature in the field to show how scientific discourses on deepfakes crimes have changed since 2019. The three main categories are social media manipulation, AI-operated sector and security-deepfakes, and face manipulation and non-consensual pornography. When taken as a whole, they highlight the diversity of deepfakes and the pressing need for interdisciplinary collaboration.

4.1 Theoretical Implications

This study adds to the theoretical understanding of deepfakes as a technical reality and a criminal threat in many societies. Computer science and criminal science may not be the only fields contributing to the rise in deepfake crimes, according to our bibliometric analysis. Furthermore, as a reflection of their multidisciplinary complexity, these instances are situated at the nexus of criminal law, cybersecurity, media studies, and artificial intelligence research. The Red Group, which deals with misinformation, disinformation, and theoretical considerations of false news, highlights the part deepfakes play in information disorder. Moreover, it draws attention to the necessity of reconsidering the fundamentals of digital propagation, media trusts, and political communication in the age of deceit fueled by artificial intelligence.

In addition to matching deepfake offenses with victims, gender-based violence, and digital damage principles, the Green Group focuses on face manipulation and pornography. In this case, the technology is used for personal gain, reinforcing the theoretical ideas of disadvantage, consent, and digital amplification of power. In addition to promoting theoretical discussion in computer science and cybersecurity on adversarial learning, system flexibility, and dual-use AI morality, Blue Group illustrates the race of technical weapons between generic AI and detection systems. This speaks to a fundamental paradox of deepfake crimes: learning about them requires the same artificial intelligence technology that makes deception possible. Therefore, when combined, these clusters broaden the theoretical definitions of loss, trust, and criminality in the digital era. They strengthen and enhance the approach that deepfakes form the convergence of crime technology and criminality and require outlines that integrate technical sophistication with human and social weaknesses.

4.2 Practical Implications

The practical implications of this study are equally important. By identifying the current trends and future research directions and main publication centers, contribution countries, and thematic groups, the study offers a roadmap for policymakers, technologists, and regulatory bodies.

First, the strong presence of the United States, China, and European countries highlights that the leadership in deepfake research focuses on technically advanced areas. This global policy has implications for coordination: developing countries, although they have less literature, face similar risks from deepfakes, but often have deficient technical and educational infrastructure to react effectively. Therefore, it is necessary to enhance and strengthen international cooperation to overcome these crimes.

Secondly, the dominance of research focused on detection confirms immediate demand for technical countermeasures. Social media platforms, programmers, law enforcement agencies, and cybersecurity companies can use these findings to design real-time monitoring systems, forensic tools, and certification protocols to reduce and mitigate risks of deep fakes.

Third, the identity of groups associated with revenge porn and non-conscious materials highlights the need for afflicted-centered interventions. Along with technical solutions, legal reforms, public awareness campaigns, and digital rights protection are important and necessary. These results confirm that the fight against deepfake crimes requires extensive strategies: technical measures alone are inadequate without parallel social, legal, and moral guarantees.

Practically, in this study list, it shows that in the era of artificial intelligence, action is required from many stakeholders, including research institutes, technology developers, regulators, and civil society, to deal with deepfake crimes.

5. Conclusion

This study provides the first comprehensive literature list of deepfakes crimes in the AI era, based on 349 articles indexed in the Web of Science Core Collection from 2019 to 2025. By mapping keyword groupings, posting patterns, and country contributions, the hazards linked with the study and research trends have been identified.

The results suggest that deepfakes is a versatile challenge in crimes, including technical, social, legal, and moral dimensions. Three major thematic currents of analysis emerge: (i) deepfakes and dissolution in the digital public sector; (ii) deepfakes technique, facial manipulation, and non-absorption pornography; and (iii) AI-Powered Identification and Safety Counters. Each stream collectively highlights different risks, depicting the complex differences between technology, crime, and society.

5.1 Study Limitations

Despite these contributions, several limitations should be acknowledged. First, the dataset was restricted to the Web of Science Core Collection and English-language publications. While this ensures consistency and quality, it excludes relevant research from other databases and non-English sources, which may bias the results towards western and English-speaking contexts. Secondly, bibliometric indicators measure productivity, citations and cooperation, but they cannot fully capture the qualitative depth of scientific discussions. Finally, the analysis reflects the state of research until September 2025, and given the rapid pace of AI innovation, new developments may quickly reshape the field.

5.2 Future Research Directions

Future research should address these limitations by integrating multiple databases—such as Scopus, IEEE explore, and Google Scholar—to expand coverage, and by combining quantitative bibliometric analyses with qualitative content studies. The inclusion of non-English publications would provide richer insights into how to study deepfake crimes in non-Western contexts, especially in areas most vulnerable to digital exploitation and information manipulation.

Additional research on policy and management responses is also needed, including comparative studies of legal frameworks across jurisdictions. Investigating the psychological and social effects on victims of deepfake crimes can bridge the gap between technical detection research and human-centered consequences. Longitudinal studies are especially valuable for

tracking the development of detection technologies and misinformation strategies, highlighting the ongoing arms race between the creators of deepfakes and defenders of digital authenticity.

Acknowledgement

The authors are grateful to the Deanship of Scientific Research at Jadara University for providing financial support for this publication.

References

- Abu Huson, Y., Sierra-García, L., & Garcia-Benau, M. A. (2024). A bibliometric review of information technology, artificial intelligence, and blockchain on auditing. *Total Quality Management & Business Excellence*, 35(1-2), 91-113.
- Alanazi, S., Asif, S., Caird-daley, A., & Moulitsas, I. (2025). Unmasking deepfakes: a multidisciplinary examination of social impacts and regulatory responses. *Human-Intelligent Systems Integration*, 1-23.
- Al-Raggad, M., Albalawee, N., Al-Mahasneh, A., Abu Huson, Y., & Albajaly, A. (2025). Unveiling financial crimes: advancing forensic accounting practices and ethical integrity through bibliometric insights. *Safer Communities*, 24(3), 244-264.
- Al-Raggad, M., Al-Raggad, A., Al-Raggad, M., Alraggad, A., & Huson, Y. A. (2024). Examining Bribery as a Financial Crime in the Private Sector within the Framework of Jordanian Penal Legislation. *Pakistan Journal of Criminology*, 16(1).
- Cover, R., Haw, A., & Thompson, J. D. (2022). Fake News and Conspiracy Theories. In *Fake News in Digital Cultures: Technology, Populism and Digital Misinformation* (pp. 77-91). Emerald Publishing Limited.
- Digmelashvili, T. (2023). The impact of cyberwarfare on the national security. *Future Human Image*, (19), 12-19.
- Farid, H. (2022). Creating, using, misusing, and detecting deep fakes. *Journal of Online Trust and Safety*, 1(4).
- 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).
- Gieseke, A. P. (2020). "The new weapon of choice": Law's current inability to properly address deepfake pornography. *Vand. L. Rev.*, 73, 1479.
- Große, C., & Sundberg, L. (2025). Generative AI and digital resilience: a research agenda. *Journal of Risk Research*, 1-26.
- Hamdan, A., Hassanien, A. E., Khamis, R., Alareeni, B., Razzaque, A., & Awwad, B. (2021). *Applications of Artificial Intelligence in Business, Education and Healthcare*. Springer. <https://doi.org/10.1007/978-3-030-72080-3>
- Hammouri, J. A., Almahasneh, A. A. A., Khwaileh, K. M., & Al-Raggad, M. M. (2024). The Criminal Liability of Artificial Intelligence Entities. *Pakistan Journal of Life and Social Sciences*, 22(2), 8785-8790.
- Huson, Y. A., Aljawarneh, N., Albajaly, A., Alkrarha, A., Alqmool, T., & Alqudah, M. (2024, April). Auditor and audit report: does business intelligence matter?. In *International Conference on Business and Technology* (pp. 1-15). Cham: Springer Nature Switzerland.
- Kalpokas, I., & Kalpokiene, J. (2022). *Deepfakes: a realistic assessment of potentials, risks, and policy regulation*. Springer Nature.
- Lin, L. S. (2025). Organisational Challenges in US Law Enforcement's Response to AI-Driven Cybercrime and Deepfake Fraud. *Laws*, 14(4), 46.
- Mansoor, S. I. U. (2024). Legal implications of deepfake technology: In the context of manipulation, privacy, and identity theft. *Central University of Kashmir Law Review*, 4, 65-92.
- McGLYNN, C. L. A. R. E., & Toparlak, R. T. (2025). The 'new voyeurism': criminalizing the creation of 'deepfake porn'. *Journal of Law and Society*, 52(2), 204-228.
- Meskys, E., Kalpokiene, J., Jurcys, P., & Liaudanskas, A. (2020). Regulating deep fakes: legal and ethical considerations. *Journal of Intellectual Property Law & Practice*, 15(1), 24-31.
- Nannaware, S. C., Pillai, R., & Kate, N. (2025). Deepfakes in action: Exploring use cases across industries. In *Deepfakes and Their Impact on Business* (pp. 71-98). IGI Global Scientific Publishing.
- Okolie, C. (2023). Artificial intelligence-altered videos (deepfakes), image-based sexual abuse, and data privacy concerns. *Journal of International Women's Studies*, 25(2), 11.
- Peck, J., Goossens, B., & Saeys, Y. (2023). An introduction to adversarially robust deep learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(4), 2071-2090.
- Shirish, A., & Komal, S. (2023). A socio-legal inquiry on deepfakes. *Cal. W. Int'l LJ*, 54, 517.
- Sophia, L. I. (2025). The Social Harms of AI-Generated Fake News: Addressing Deepfake and AI Political Manipulation. *Digital Society & Virtual Governance*, 1(1), 72-88.
- Sylvester, S. (2020). Don't Let Them Fake You Out: How Artificially Mastered Videos Are Becoming the Newest Threat in the Disinformation War and What Social Media Platforms Should Do About It. *Fed. Comm. LJ*, 73, 369.
- Tuliakov, V. (2023). Criminal Law and its Victim-Oriented Development: an Academic Inquiry. *Copernicus Political and Legal Studies*, 2(3), 70-74.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social media+ society*, 6(1), 2056305120903408.
- Van Eck, N. J., & Waltman, L. (2010). *Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538. <https://doi.org/10.1007/s11192-009-0146-3>

- Van Eck, N. J., & Waltman, L. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, 1(1), 363–376. https://doi.org/10.1162/qss_a_00018
- Yadav, G., Sadique, M. Z., Kumar, S., Sharma, R., Sharma, M., Sharma, R., & Rattan, T. (2025). Psychological Trauma and Legal Challenges of Deep fake Technology. *Sciences of Conservation and Archaeology*, 37(1), 143-150.
- Yadlin-Segal, A., & Oppenheim, Y. (2021). Whose dystopia is it anyway? Deepfakes and social media regulation. *Convergence*, 27(1), 36-51.



© 2026 by the authors; licensee Growing Science, Canada. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).