

# FORMS OF EXTREMISM AND VIOLENCE AND THEIR RELATIONSHIP TO HATE SPEECH IN THE VIRTUAL ENVIRONMENTS

<sup>1\*</sup>MOHAMMED EL NADIR ABDALLAH TANI, <sup>2</sup>ZAHIA EL-SSAID YSSAAD, <sup>3</sup>ARIF JUMA BIRO TROUK, <sup>4</sup>AMIRA MOHAMMED AL-KHIDIR, <sup>5</sup>NASSREDINE CHEIKH BOUHENNI, <sup>6</sup>MAHA THAID AL-SHAMMARI

<sup>1234</sup>DEPT OF MEDIA. HA'IL UNIVERSITY, HA'IL, SAUDI ARABIA

<sup>5</sup>DEPT OF ARAB. HA'IL UNIVERSITY, HA'IL, SAUDI ARABIA

<sup>6</sup>DEPT OF SOCIAL SCIENCES. HA'IL UNIVERSITY, HA'IL, SAUDI ARABIA

\*Corresponding Author Email: [m.abdallahtani@uoh.edu.sa](mailto:m.abdallahtani@uoh.edu.sa), ORCID: 0000-0002-4261-4559

## Abstract

The aim of the present review is to explore the forms of extremism and violence in the virtual environment that have developed in recent years and have now grown into an issue of preeminent seriousness for scholars, policymakers, and society at large. The focus of this review article is to synthesise the findings of a number of studies on the existence, prevalence, and use of hate speech as a strategic constitutive element of extremist groups when exploring the best strategies for mitigation. The review emphasises how the digital domain allows for the dissemination and culturing of extremist thoughts by drawing upon real-life instances across different continents under diverse political circumstances. For the review article, a total of 17200 studies were chosen, and after the final screening, 14 studies were chosen as appropriate for the review. These studies were selected from Web of Science and Google Scholar with in the time frame of 2018-2024. On account of the dehumanisation of others, hate speech is said to be central to this process in terms of in-group bonding and calls for violence. The need for broad, multilayered counter-extremism policies covering regulatory interventions, advanced technological interventions, community-based programs, and global cooperation is therefore underlined. Moreover, the review highlights the driving force behind this extremist content and the effect it had on the individuals affected by or exposed to such content. The findings revealed that over time counter violent extremism content strategies have been developed by various countries.

**Keywords:** Virtual Environment, Violence, Extremist Content, Hate Speech, Far-Right Extremism

## INTRODUCTION

The invention of the internet and social media in the digital era revolutionised the way people communicate, share information, and even express one's self to the world (Sadagheyani & Tatari, 2020). While these platforms democratised access to information and united people across borders, they have also been hijacked by some malicious characters at the same time. Anonymity and reach, afforded by virtual environments, have made extremism and violence more elaborately propagated, henceforth posing formidable challenges to the very fibre of societal stability and security (Гоктүг, 2023). With these digital tools, such extremist groups can propagate their ideology recruit, and incite people to commit acts of violence more comprehensively and persuasively than ever before.

Extremism finds expression in right-wing extremism, Islamist extremism, nationalist movements, and other ideologically driven groups (Hutchinson et al., 2023). Australia and Canada are among those countries in which right-wing extremism is mostly active, and these users use Facebook to achieve their target (Hutchinson, 2022). These groups are utilising this capacity of the internet to ensure that their thinking and ideas can create echo chambers, that is, to increase their voices and, at times, browbeat any opposing view. It is in these settings online that hate speech multiplies, further catalysing this process of radicalisation. As some submissions pointed out, hate speech acts as a specific tool in normalising extremist ideas and inciting violent behaviour (Olteanu et al., 2018).

This review article seeks to explore how different forms of extremism and violence relate to one another, as well as the role of hate speech in virtual environments. Drawing on a variety of academic studies, the study aims to find out the dynamics of how extremist ideologies are being shared online and the ways in which hate speech facilitates such

processes. It aims to bring greater clarity around the particular characteristics of different extremist groups, how they operate online, and the related effects on individuals and communities.

The study further discusses the ways through which hate speech travels in an online ecosystem and the contributions from major platforms and technologies that have enabled this problem to spread. The study confers the psychological and social effects of exposure to hate speech and extreme content online, especially among vulnerable groups such as youth. It aims to elaborate on the effective current strategies that are used in detecting, monitoring, preventing, and addressing hate speech and extremism in digital spaces. Some of these strategies include advanced machine learning techniques, moderation policies related to content, and legal frameworks that aim to prevent online radicalisation leading to violence.

The junction of extremism, violence, and hate speech within the digital environment is a complex, evolving challenge that calls for a multi-disciplinary approach. It calls for collaboration between researchers, policymakers, and technology companies to come up with comprehensive solutions that protect freedom of speech while ensuring public safety. The present study brings together current research findings to provide a comprehensive description of the forms of extremism and the strategies that are needed four counter extremism and violence that are prevalent in virtual environments through an extensive review of literature. In the present paper, the researcher analyses counter-online extremism strategies and involvement of hate speech by providing insights and recommendations for future research and policy-making initiatives. In other words, the prospect of digital landscapes taking over with harm or healing to humanity clearly suggests the critical need for a relationship between extremism, violence, and hate speech. Having taken this as a standpoint, this review aims to be able to illuminate such issues in demand of better interventions and the creation of a safer online environment that is inclusive to all.

- The aim of this review is to identify the forms of extremism present in the virtual environment of hail and design strategies to reduce extremism in the virtual environment.

## LITERATURE REVIEW

With the evolution of the digital world, violence and extremism on digital platforms has evolved (Winter et al., 2020). Violence and extremism in virtual environments take place with the use of hate speech in the virtual environment. Hate speech serves as a tool for people to express their radical views in online communities (Fortuna & Nunes, 2018). Among the most important risks that threaten social life and social communication in virtual societies; Violence in all its forms circulated in the virtual space. Social, communicative and psychological studies in this field showed the negative effects that society is witnessing as a result of the increasing phenomenon. These spaces have become a platform for violent and hate speech, instead of raising awareness in combating and preventing it. Manifestations of violence and hate speech have become entrenched in social networks as a result of lack of awareness on the one hand, and on the other hand, due to the user's lack of reservation to use phrases or publish pictures or contents offensive to individuals and institutions through the technique of installing photos and videos ..., as a matter of taking into account the reasons, identities have become Anonymous or borrowed default uses verbal violence to pass its agenda and tear the social and emotional fabric of society. Hate speech that is disseminated through social networks usually leads to intolerance and generates hateful speeches. Before it multiplies, it spreads and amplifies through social media channels, and sometimes these echoes resonate. Contents through TV channels and newspapers alike, and have a special capacity for existing tensions between different groups and ignite violence among their members, as they may raise new concerns about the association of inflammatory discourse through these spaces with acts of violence. It may be a motive for accidents that contribute to psychological, social, economic, emotional, material and physical harm to people. The use of social networks has become aimed at distorting facts, spreading falsehood, and seditious violent discourse, As for the methods of preventing violence, it requires addressing the factors leading to violence and treating them in a scientific way, and preventive efforts must focus on the stages of childhood and adolescence and the role of educational institutions in this, given that the behavior of violence consists mostly of early childhood, according to psychological studies(Abdallah Tani and Alrasheed,2023).A study conducted on the use of extremist language in the virtual environments of the US proposed that there was an increase in uncivil and hostile language in online communities. The study explored the platform Reddit to analyse the use of extremist language. It concluded that the use of such language is due to the political unrest in the country that led people to radicalise their views on online communities (Hiaeshutter-Rice & Hawkins, 2022).

Among the studies conducted on the new media and communication environment in the city of Hail is a study conducted by Abdallah Tani,2023 on The quest for creating a new media and communication environment is at the core of the competitive advantages, mertis, and marketing strengths of the so-called city economy, especially that the effectiveness of new media and institutional communication is reliant on interaction with existing or potential customers or stakeholders as well as maintaining a strong electronic reputation. Interaction can take various forms, such as trade fairs, marketing promotions, customer relation initiatives, and other efforts through which organizations interact with the public using electronic means in order to meet customer needs more effectively. As many large companies have dedicated departments for managing investor relations, a question arises: is the new institutional

media and communication environment in the Hail region appropriate for promoting digital reputation and attaining local and international competitiveness in the region's development? (Abdallah Tani and Ysaad, 2023).

A study conducted on "Detecting Islamic Radicalism in Arabic Tweets Using Natural Language Processing" identified the increase in the challenge to curb online hate speech and extremism through social media. The authors underscore the role of platforms like Twitter in making extremist content available and reaching a wide audience in an instantaneous manner. The paper contributes a manually labelled dataset of 3,000 Arabic tweets to emphasise the role of advanced machine learning techniques in sentiment analysis and classification. The findings of the study have revealed that online hateful content normally acts as clickbait and gets tremendously amplified by automated bots, which fuels real-world violence related to lone-wolf attacks and ethnic cleansing incidents. These findings have put pressure on ensuring that well-implemented measurements for monitoring and controlling hate speech online are in effect since it is doing more harm than ever to society (Mursi et al., 2022).

Abdallah tani,2021 Conducted a study on Information Society, the term spread in the West and USA, primarily it served as a continuation of the theories that spawned the terms of the industrial society and post-industrial, as is evident from the "title" At the basis of the label technological optimism of capacity to resolve many social problems which are intractable by using the information, other than the descriptive, which codifies and inaugurates the objective spread of "computing" and "media" that interspersed with the social fabric as a whole. And often draws attention to focus on the Japanese experience, which provides an example of the information society in Japan, an ambitious program that aims to create and develop the social structure - Computers distinct include production, distribution and transportation services and management, education and livelihoods, and this social structure - Computer permeate networks télé-computing (télé in Greek meaning away and dimension) so that people can through them to solve their own professional problems by direct contact, and there is a national policy in Japan for the creation and establishment of "industry knowledge" as well as projects to reform the country's social development , including the systems of rehabilitation workforce and of re-distribution, so that the community becomes in its Supreme stages "j - Masuda," who is one of the leaders of that program - to a consumer society and prosperous of high class, and then going to solve all personal problems - according to this belief - and its requirement of creativity and self-realization through the global use of planetary information, so they attaches a great hopes on the "globalization" and "internationalization" of the trends of new information - that is for "computing" the society and its global planetary "medias"(Abdallah tani,2021).

Researchers have argued for an all-rounded analysis of extremism on social media in a research conducted by Aldera et al. (2021) on the "Exploratory Data Analysis and Classification of a New Arabic Online Extremism Dataset". They detail how extremist groups use online platforms to promote ideologies, recruit members, and extend their operations beyond physical boundaries. The present study underlines especially those aspects of social media that allow for anonymous contact between extremist individuals and groups to culminate in low-cost recruitment and dissemination of extremist discourse. With the increasing number of subscribers to social media, the potentiality of extremist propaganda influencing public opinion and provoking violence gets fuller. It is, therefore, incumbent upon the authorities to keep a stiff vigil on social media to ensure these areas are not exploited by the extremists (Aldera et al., 2021).

Ring win extremism is a kind of extremist content that is most popular in Western countries. A study conducted on the comparison between Australian and Canadian right-wing extremism groups on Facebook revealed the emerging themes in their content. The researcher compared these two countries notwithstanding their shared cultural and historical identities. The study gathered data from 30 Australian and 29 Canadian right-wing pages on Facebook ranging between 2011 and 2019. The comparative analysis revealed that Canadian right-wing extremists preferred passive engagement in online platforms, while Australian extremists preferred active engagement on online platforms (Hutchinson et al., 2023).

## METHOD

### Search strategy

The literature review provides a comprehensive analysis of the authentic published studies. This process involves a multifaceted approach to finding relevant and authentic sources. For the present study, the latest 7 years' studies were chosen that were published between 2018 and 2024. These studies were accessed using two databases, Google Scholar and Web of Science. The articles were accessed using specific search strings and specific keywords. Table 1 shows the specific search strings and search terms that were used to gather data.

Database	Search strings	Search terms
Google Scholar	"extremism AND violence", "hate speech AND radicalisation", "online extremism AND social media."	Extremism, Violence, Hate speech, Radicalization, Online extremism, Social media and extremism

Web of Science	"counter-extremism strategies", "psychological factors in extremism", "youth AND radicalisation"	Counter-extremism strategies, Psychological factors in extremism, Youth and radicalisation, Gender and extremism
----------------	--	--

Table 1 Search strategy

#### Exclusion & Inclusion criteria

Table 2 shows the inclusion and exclusion criteria chosen for the present study to gather data.

Inclusion Criteria	Exclusion Criteria
Studies published in peer-reviewed journals	Studies not published in peer-reviewed journals.
Studies focused on online extremism, violence, and hate speech	Studies other than extremism, violence, and hate speech
Studies available with full text for the public on Google Scholar, Web of Science, and JSTOR	Studies not available with full text available to the public.
Studies available in English	Studies available in other languages
Studies published between 2018-2024	Studies not published between 2018-2024

Table 2 Exclusion & Inclusion criteria

#### Data Extraction & Analysis

For a comprehensive literature review, a total of 15 studies were included in the study. These studies were chosen based on their relevance with the research topic, title of the research, publishers, and the purpose of the research. These studies were then analysed using thematic analysis by Braun & Clarke. They were extracted from the 15 articles, coded, and then categorised into the relevant themes and sub-themes.

#### Flow Diagram

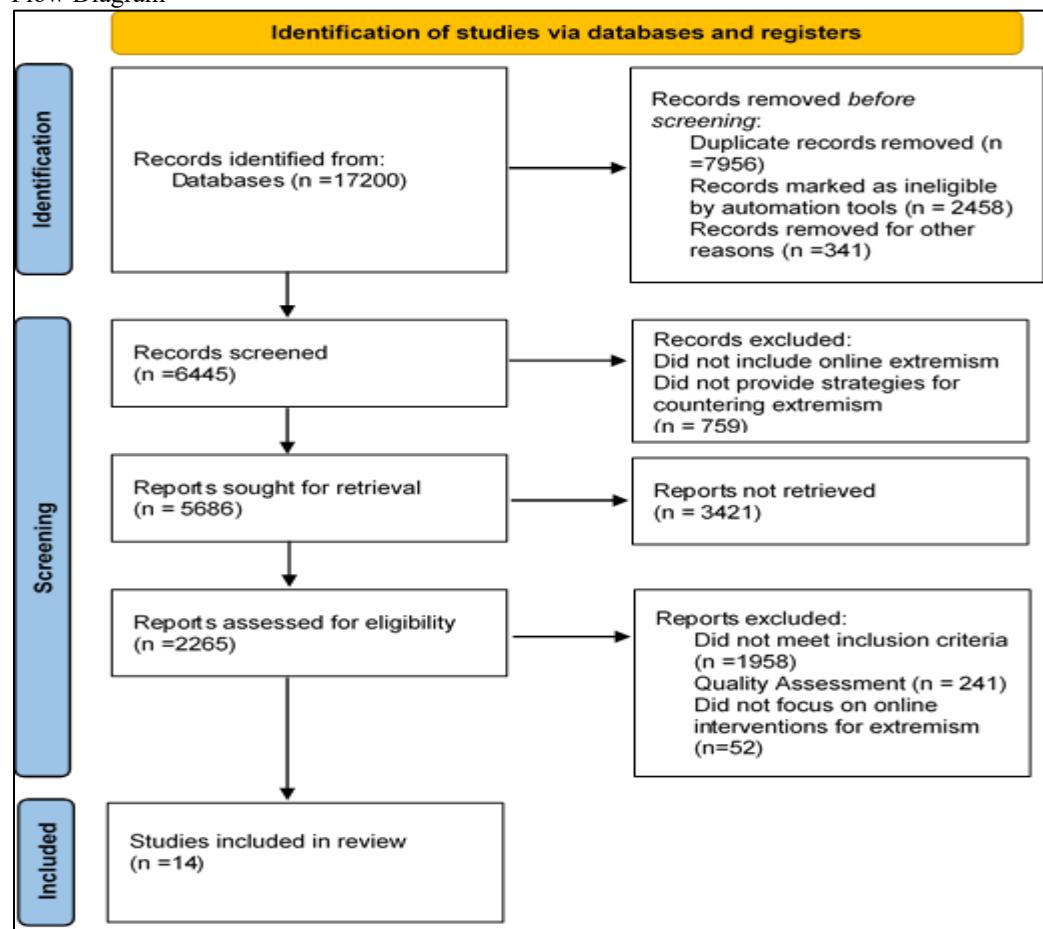


Figure 1 PRISMA Flow chart

## RESULTS

The table 3 shows the key findings of all the studies included in the review. It details the name of the author, publication year, key findings and the forms of extremism identified in the chosen studies.

S.n o	Study	Author & year	Research Focus	Methodolo gy	Key findings	Strengths & Weaknesses	Identified forms of extremism
1	Down the Rabbit Hole: Detecting Online Extremism, Radicalisation, and Politicised Hate Speech	(Govers et al., 2023)	Detection of online extremism and hate speech	Computational analysis/ Literature review and survey of detection methods	Proposes a framework for detecting stages of radicalisation/ Reviews various methods for detecting extremism, radicalisation, and hate speech and evaluates their effectiveness.	Strengths: Broad overview of detection techniques and effectiveness. Limitations: Focus on theoretical approaches rather than practical implementations.	Extremism detection, radicalisation, politicised hate speech.
2	The Role of the Internet in Facilitating Violent Extremism: Insights from Former Right-Wing Extremists	(Gaudette et al., 2022)	Role of the internet in violent extremism	Interviews with former extremists/ Qualitative interviews and case studies	Internet acts as a catalyst for radicalisation/ The internet facilitates violent extremism by providing platforms for radicalisation and community-building among extremists.	Strengths: In-depth insights from former extremists. Limitations: Limited generalizability due to focus on former right-wing extremists.	Right-wing extremism, online radicalisation, violent extremism.
3	Online extremism: research trends in internet activism, radicalisation, and counter-strategies	(Winter et al., 2020)	Role of the internet in violent extremism	Literature review and trend analysis	Provides an overview of research trends in online extremism, including activism, radicalisation, and counter-strategies.	Strengths: Comprehensive overview of research trends. Limitations: Broad focus may not delve deeply into specific issues or solutions.	Research trends, internet activism, radicalisation, counter-strategies.
4	Hate Speech Detection in	(Ahmed et al., 2021)	Mechanisms of hate	A systematic	Highlights the	Strengths: Comprehensive	Hate speech detection

	Social Media Surveillance: A Review of Related Literature		and extremism spread through social media/ Reviews various techniques and approaches for hate speech detection on social media platforms./ Review of methods for hate speech detection.	review of computational hate speech detection methods	strengths and limitations of different detection algorithms, emphasizing the need for context-aware approaches	e review of detection methods. Limitations: Limited discussion on the effectiveness of different methods.	methods and challenges.
5	The Use of Counter Narratives to Combat Violent Extremism Online	(Rees & Montasari, 2023)	Counter-narratives for combating online violent extremism.	Case studies and theoretical analysis	Counter-narratives can be effective in countering online violent extremism when well-targeted and strategic.	Strengths: Offers practical strategies for counter-narratives. Limitations: Effectiveness can vary based on context.	Counter-narratives, violent extremism, online interventions.
6	Ethical and technical challenges of AI in tackling hate speech	(Cortiz & Zubiaga, 2020)	Challenges of AI in detecting hate speech.	Literature review and analysis of existing AI tools	AI tools have limitations in detecting nuanced hate speech and face ethical challenges in ensuring fairness and accuracy.	Strengths: Detailed discussion on AI challenges, broad perspective on ethical issues. Limitations: Limited focus on specific case studies or practical solutions.	Technological extremism, AI ethics in combating online hate.
7	Terrorism, violent extremism, and the internet: Free speech considerations	(Killion & Attorney, 2019)	Free speech considerations in regulating online extremism.	Legal analysis and policy review	Balancing national security interests with protecting free speech is a central challenge in	Strengths: Comprehensive legal analysis of current policy discussion. Limitations: This may not address practical	Terrorism, violent extremism, free speech concerns.

					regulating online extremism.	enforcement challenges in detail.	
8	Hate speech epidemic. The dynamic effects of derogatory language on intergroup relations and political radicalisation	(Bilewicz & Soral, 2020)	Effects of hate speech on intergroup relations and radicalisation.	Experimental studies and literature review	Derogatory language exacerbates intergroup tensions and can lead to political radicalisation.	Strengths: Experimental approach, detailed analysis of language effects. Limitations: Limited scope on long-term effects and broader societal impacts.	Derogatory language, political radicalisation, intergroup conflict.
9	Legal and ethical aspects of violent extremism and terrorism	(Jeknić et al., 2022)	Legal and ethical challenges in countering extremism and terrorism.	Literature review and case study analysis	Identifies legal and ethical challenges in countering extremism and terrorism, focusing on uncertainty and risk.	Strengths: Broad examination of legal and ethical issues, diverse case studies. Limitations: May not offer specific solutions or practical recommendations.	Violent extremism, terrorism, legal and ethical considerations.
10	The same pieces to a different puzzle? Comparing radicalisation towards right-wing and Islamist extremism	(Nouri, 2022)	Comparison of radicalisation processes for different extremisms.	Comparative analysis	Highlights similarities and differences in radicalisation processes between right-wing and Islamist extremism.	Strengths: The comparative approach provides a broad understanding of radicalisation. Limitations: May oversimplify complex radicalisation processes.	Right-wing extremism, Islamist extremism, radicalisation processes.
11	Changing Threat Dimensions: Preventing Extremism in The Digital Age	(Patabendige, 2023)	Adapting prevention strategies for digital age extremism.	Literature review and policy analysis	Discusses new dimensions and strategies for preventing extremism in the digital context.	Strengths: Focuses on current and future challenges in digital extremism prevention. Limitations: May lack specific case studies or empirical evidence.	Digital age challenges, extremism prevention strategies.

12	New media users' awareness of online inflammatory and mobilisation methods for radical and extreme activities	(Elareshi et al., 2021)	Awareness of online radicalisation methods.	Survey and qualitative analysis	New media users are generally aware of inflammatory and mobilisation methods used for radical activities.	Strengths: Insight into user awareness of extremist methods. Limitations: Limited scope on the effectiveness of countermeasures or interventions.	Inflammatory methods, radical activities, online mobilisation.
13	A deep learning approach for automatic hate speech detection in the Saudi Twittersphere	(Alshalan & Al-Khalifa, 2020)	Deep learning for hate speech detection in social media.	Deep learning model development and evaluation	Deep learning models are effective at detecting hate speech, though there are challenges in model accuracy and bias./	Strengths: Advances in hate speech detection technology. Limitations: Models may struggle with nuances of hate speech and cultural context.	Hate speech detection, deep learning models, online hate in the Twittersphere.
14	Predictors of viewing online extremism among America's youth	(Costello et al., 2020)	Predictors of youth engagement with online extremism.	Quantitative survey and statistical analysis	Factors such as political views, social isolation, and exposure to extremist content increase the likelihood of youth engaging in online extremism.	Strengths: Large sample size, diverse youth perspectives. Limitations: Relies on self-reported data, which may introduce biases.	Political extremism, religious extremism, radicalisation of youth.

Table 3 Results

## DISCUSSION

### Extremism and Violence Prevalence

Bringing together the different sources and case studies shows that there is a high and increasing prevalence of extremism and violence in the virtual realm. According to the articles, over the years, sites online have proved to be the breeding grounds for critical radicals and espousing extremist ideologies. Such realms are anonymous and have high reach; therefore, there is sharing of cause violent content in laying violent rhetoric. The rise and spread of such practices are not regionally centred or particular to a determined group or area, which is visible from the very diverse geographical and sociopolitical contexts represented in the literature (Killion & Attorney, 2019).

Exposure to hate speech is linked with the age factor. Children between the ages of 15 to 24 are most likely to be exposed to violence and extremism in the virtual environment of America. The exposure to extremism increases with

the increase in the amount of time children spend online and the political views they hold (Costello et al., 2020). White people are said to be most exposed to extremist content online. Furthermore, Asian countries are no more prone to violent extremism. Bangladesh is one of those countries that is widely affected by violent extremism in the virtual environment, especially on Facebook. These extremists lure people towards their content, recruit them for their violent activity and become a symbol of terror (Fernandez et al., 2018).

Radicalisation is often used as an interchangeable term for extremist violence. A study conducted in the Netherlands showed that among the most affected individuals are the youth. They are the most vulnerable group when exposed to extremist ideologies. The research conducted comparing radicalisation among the far right extremists and Islamic radicalisation focused on the reason behind these radicalised groups. Their findings revealed that there exists a high similarity between the far-right youth and Islamist extremists. It also revealed the role perceptions of injustice play in grievances in both groups (Nouri, 2022).

In the prevailing digital age, where the use of the internet has been essential for everyone. The digital age is characterised by the age of information. The use of extremism in virtual environments poses a threat to the national security of the country. A study conducted in Sri Lanka addresses the possible threats of extremist content that prevail in the social environment of the country. The study discusses the case of the Easter Sunday attacks in 2019 that were carried out by the ISIS group. This incident resulted in a form of present online extremism in the country (Patabendige, 2023).

### **Use of Hate Speech as a Tool**

Hate speech is recognised as one important tool used by extremists in driving violence and escalating ideological divides. The use of virtual environments makes hate speech pervasive; through dehumanising a particular group of people, it is meant to make the perpetrator groups more united and call them into action to do hostile acts. The linkage between speech and violence is now well established, and to that extent, it becomes important to develop mechanisms that monitor and contain such speech online.

In Arab countries, Twitter is the most used platform for social interactions. Based on the nature and policies of the Twitter account, it was exploited for disseminative and hateful content. Abusive language in social media corresponds to a complex phenomenon with an extended, quite diffuse spectrum of overlapping forms or targets. For instance, it can be called hate speech, offensive language, or cyberbullying; several works have been conducted to detect and locate such types of languages in the literature (Alshalan & Al-Khalifa, 2020).

Hate speech serves as a tool for extremist content on online platforms. It involves pointing towards minority groups and derogating people from different sectors of society. (Bilewicz & Soral, 2020) Identify hate speech as a harmful language; it promotes prejudice and deteriorates well-being.

### **Driving Force towards Violent Extremism**

There is disagreement over what motivates people to engage in Violent extremist activities. Scholars concur that there are several contributing elements rather than just one. Macro-, meso-, and micro-level drivers are the three categories into which the Australian Government's Department of Foreign Affairs and Trade has separated violent extremism factors. The macro-level factors encompass a country's overall socio-economic and political patterns and can also be acknowledged as driving forces. Identity groupings, that is, specific groups brought together by common beliefs based on religion or ethnicity, have the power to pull or enable the meso-level processes. Ultimately, the pull factors or individual factors resulting from social isolation, interpersonal interactions, individual vulnerabilities, or openness to radical narratives make up the micro-level drivers (Gordon & True, 2019).

### **Counter-Extremism Strategies**

Effective counter-extremism strategies must address both the digital and offline dimensions of the issue. A review of most of these articles brings out a multidimensional approach to designing policy measures, technological solutions, and community engagements. Well-noted strategies include:

**Regulatory Frameworks:** Enforcement and implementation of laws that criminalise hate speech and control activity on extremist grounds, but importantly, ones that do not infringe upon the rights of others. Countering extremist content online is divided into three categories that involve positive, negative, and monitoring. This involves blocking, censoring, removing and identifying and analysing extremist content (Amit et al., 2021).

**Technological Interventions:** Innovate cutting-edge algorithms and AI tools to trace and bring down extremist content on the internet at the earliest opportunity. Collaboration among tech companies with governments to moderate content in a more effective manner. The counter-extremist project is to identify extremist content online and deal with the ideology of such extremism(Costello et al., 2020; Project, 2024). convolutional neural network (CNN) and recurrent neural network (RNN) were employed to detect hate speech in Arabic tweets. The study also evaluated the recent language representation model bidirectional encoder representations from transformers (BERT) in hate speech detection (Alshalan & Al-Khalifa, 2020). The use of AI in detecting hate speech results in identifying those challenges and countering extremist content online (Cortiz & Zubiaga, 2020).

**Community Programs:** Involvement of local communities in the efforts toward counter-radicalisation through education and outreach campaigns; creation of a safe environment through the empowerment of community leaders and influencers to promote tales of tolerance and inclusiveness. A review of several studies revealed programs against

VE, Exit White Power; Think Again, Turn Away, Muflehun; Don't be a Puppet, and Open Letter to Our Sons and Daughters in Syria and Iraq are among those programs that are implemented to counter violent extremism in western countries (Amit et al., 2021). Since the use of hate speech is directed towards extremism, it is therefore recommended to create an environment where minority groups are treated equally.

**International Collaboration:** Deepening country-to-country relationships to harness intelligence-sharing, best practices, and resources in this fight against cyber-radicalisation. This is especially important because the internet is a global technology, and extremist groups are transnational operators.

#### Effects of Extremism Content

The exposure to extremist content leads to several effects on the individuals who are exposed to such content. A study on the effects of hate speech argues that exposure to such content makes the individual desensitisation, which means the more a person is exposed, he becomes prone to such content and doesn't react going further. Such people start believing that violence is a norm and it doesn't harm anyone. As a result, these individuals are more likely to be involved in derogatory acts (Bilewicz & Soral, 2020).

## CONCLUSION

The review of extremism and violence in virtual environments has a trend of an increase in the prevalence and impact of the phenomena. Online platforms represent important fields for both the spreading of extremist ideologies and the organisation of violent activities. Anonymity and the broad reach, facilitated by the power vested in the internet, have quickly spread hate speech, which has become a cardinal tool to incite violence and attract members, deepen ideological gaps, and rally like-minded actors. To counter the extremist content various strategies to counter extremism must be implemented. Key aspects include the implementation of regulatory frameworks that are stringent in the criminalisation of and curbing of hate speech and extremism propaganda, the development of cutting-edge technological interventions for content detection and removal, and community-based programs for tolerance and inclusivity. International cooperation is also paramount to dealing effectively with the transnational aspect of online extremism.

The integration of all such insights from diverse disciplines becomes imperative in developing resilient societies that can counter modern risks related to online extremism. Of course, without a multi-disciplinary approach to understanding the nature and consequences of these dynamics, it is intriguing how sources of risk can be approached and threats tackled with effectiveness.

#### Limitations and Future recommendation

The study is limited to 14 studies that may not comprehend the presence of extremism content in the virtual environment. The study focuses on a limited country and, therefore, does not account for other countries. Future studies should conduct research using more studies. A comprehensive literature review or PRISMA shall be used for further research.

#### Credit authorship contribution statement

Mohammed EL Nadir Abdallah Tani, zahia El-saaid Yssaad: Conceptualization, Methodology, Software, Validation, Formal analysis, Resources.

Arif Juma Biro Trouk, Amira Mohammed Al-Khidir: Supervision, writing – original draft

Nassredine Cheikh Bouhenni , Maha Thaid Al-Shammari : Writing – review & editing, Visualisation.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgement

This research has been funded by the Deputy for Research & Innovation, Ministry of Education, through the initiative of institutional funding at the University of Hail – Saudi Arabia, through project number RG-23034.

#### Data availability

No data was used for the research described in the article.

## REFERENCES

1. ABDALLAH TANI , M, E. N., ALRASHEED,T.K.A.2023. Violence in digital communication social media through the of installing technique, Measurement: Sensors,Volume 25,100646, ISSN 2665-9174, <https://doi.org/10.1016/j.measen.2022.100646>.
2. ABDALLAH TANI , M, E. N. 2021. Information society and knowledge gap in the third world–A conceptual approach–. Online Journal of Communication and Media Technologies, 5(September 2015-Special Issue), 53-63. <https://doi.org/10.30935/ojcmt/5690>.

3. ABDALLAH TANI , M, E. N., Yssaad, Z.ES. & Mesawa, M.T.A.2023. Communication Environment and Development of the Hail Region: Using IoT Technique. *Wireless Pers Commun* . <https://doi.org/10.1007/s11277-023-10728-2>

4. Ahmed, U., Bibi, R., Ullah, O., & Bano, R. (2021). Hate Speech Detection in Social Media Surveillance: A Review of Related Literature. *Innovative Computing Review*, 1(1), 01–11-01–11.

5. Aldera, S., Emam, A., Al-Qurishi, M., Alrubaian, M., & Alothaim, A. (2021). Exploratory data analysis and classification of a new arabic online extremism dataset. *IEEE Access*, 9, 161613-161626.

6. Alshalan, R., & Al-Khalifa, H. (2020). A deep learning approach for automatic hate speech detection in the saudi twittersphere. *Applied Sciences*, 10(23), 8614.

7. Amit, S., Barua, L., & Kafy, A.-A. (2021). Countering violent extremism using social media and preventing implementable strategies for Bangladesh. *Heliyon*, 7(5).

8. Bilewicz, M., & Soral, W. (2020). Hate speech epidemic. The dynamic effects of derogatory language on intergroup relations and political radicalisation. *Political Psychology*, 41, 3-33.

9. Cortiz, D., & Zubiaga, A. (2020). Ethical and technical challenges of AI in tackling hate speech. *The International Review of Information Ethics*, 29.

10. Costello, M., Barrett-Fox, R., Bernatzky, C., Hawdon, J., & Mendes, K. (2020). Predictors of viewing online extremism among America's youth. *Youth & Society*, 52(5), 710-727.

11. Elareshi, M., Habes, M., & Ziani, A.-K. (2021). New media users' awareness of online inflammatory and mobilisation methods for radical and extreme activities. *Ilkogretim Online*, 20(5), 5567-5576.

12. Fernandez, M., Asif, M., & Alani, H. (2018). Understanding the roots of radicalisation on twitter. *Proceedings of the 10th ACM conference on web science*,

13. Fortuna, P., & Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys (CSUR)*, 51(4), 1-30.

14. Gaudette, T., Scrivens, R., & Venkatesh, V. (2022). The role of the internet in facilitating violent extremism: insights from former right-wing extremists. *Terrorism and Political Violence*, 34(7), 1339-1356.

15. Gordon, E., & True, J. (2019). Gender stereotyped or gender responsive? Hidden threats and missed opportunities to prevent and counter violent extremism in Indonesia and Bangladesh. *The RUSI Journal*, 164(4), 74-91.

16. Govers, J., Feldman, P., Dant, A., & Patros, P. (2023). Down the rabbit hole: Detecting online extremism, radicalisation, and politicised hate speech. *ACM Computing Surveys*, 55(14s), 1-35.

17. Hiaeshutter-Rice, D., & Hawkins, I. (2022). The language of extremism on social media: An examination of posts, comments, and themes on Reddit. *Frontiers in Political Science*, 4, 805008.

18. Hutchinson, J. (2022). Australian and Canadian far-right extremism: a cross-national comparative analysis of social media mobilisation on Facebook [Macquarie University].

19. Hutchinson, J., Amarasingam, A., Scrivens, R., & Ballsun-Stanton, B. (2023). Mobilising extremism online: comparing Australian and Canadian right-wing extremist groups on Facebook. *Behavioral Sciences of Terrorism and Political Aggression*, 15(2), 215-245.

20. Jeknić, G. P., Tomović-Šundić, S., & Jeknić, I. (2022). **LEGAL AND ETHICAL ASPECTS OF VIOLENT EXTREMISM AND TERRORISM**. How to deal with uncertainties in increasingly complex environment?(The new cartography of risk and crises), 25.

21. Killion, V. L., & Attorney, L. (2019). Terrorism, violent extremism, and the internet: Free speech considerations. *Congressional Research Service*, 6.

22. Mursi, K. T., Alahmadi, M. D., Alsubaei, F. S., & Alghamdi, A. S. (2022). Detecting Islamic radicalism Arabic tweets using natural language processing. *IEEE Access*, 10, 72526-72534.

23. Nouri, M. (2022). The same pieces to a different puzzle? Comparing radicalisation towards right-wing and Islamist extremism.

24. Olteanu, A., Castillo, C., Boy, J., & Varshney, K. (2018). The effect of extremist violence on hateful speech online. *Proceedings of the international AAAI conference on web and social media*,

25. Patabendige, C. (2023). **Changing Threat Dimensions; Preventing Extremism in The Digital Age**.

26. Project, C. E. (2024). <https://www.counterextremism.com/>

27. Rees, J., & Montasari, R. (2023). The Use of Counter Narratives to Combat Violent Extremism Online. In *Digital Transformation in Policing: The Promise, Perils and Solutions* (pp. 15-31). Springer.

28. Sadagheyani, H. E., & Tatari, F. (2020). Investigating the role of social media on mental health. *Mental health and social inclusion*, 25(1), 41-51.

29. Winter, C., Neumann, P., Meleagrou-Hitchens, A., Ranstorp, M., Vidino, L., & Fürst, J. (2020). Online extremism: research trends in internet activism, radicalisation, and counter-strategies. *International Journal of Conflict and Violence (IJCV)*, 14, 1-20.

30. Гоктүг, С. (2023). **ONLINE VISIBILITY AND RADICALISATION: CONCEPTS, CONTEXTUALISATION, AND CASES**. Адам әлемі, 95(1), 132-141.