False Information as a Threat to Modern Society: A Systematic Review of False Information, Its Impact on Society, and Current Remedies

M Ahmed[1], SD Bachmann[2], C Martin[1], T Walker[1], J Rooyen[1], A Barkat[3]

[1]*School of Science, Edith Cowan University*
Perth, Australia

E-mail: mohiuddin.ahmed@ecu.edu.au, cwmarti0@our.ecu.edu.au, tristraw@our.ecu.edu.au, jvanroo0@our.ecu.edu.au

[2] *Professor in Law and Co-Convenor National Security Hub*
University of Canberra
Canberra, Australia

Visiting Research Fellow, The Security Institute for Governance and Leadership in Africa
Stellenbosch University
Stellenbosch, South Africa

E-mail: Sascha.Bachmann@canberra.edu.au

[3]*Faculty of Science and Technology, University of Canberra*
Canberra, Australia

E-mail: Abu.Barkat@canberra.edu.au

Abstract: *False information and by extension misinformation, disinformation and fake news are an ever-growing concern to modern democratic societies, which value the freedom of information alongside the right of the individual to express his or her opinions freely. This paper focuses on misinformation, with the aim to provide a collation of current research on the topic and a discussion of future research directions. It argues that a major current issue is the lack of reliable verified datasets and, by extension, the algorithmic learning models based on them. Through the understanding of different variations of false information, it has been possible to see the impact this absence is currently having on society. At the core is a focus on misinformation. A deeper investigation shows that artificial intelligence-based techniques are widely used to counter the effects and confusion caused by misinformation. Potential consequences of misinformation are being highlighted to allow insight into potential issues that may arise from this form of false information to highlight the severity of the issue and to demonstrate the growing need for reliable detection and prevention systems.*

Keywords: *Artificial Intelligence, Deepfakes, False Information, Infodemic, Disinformation, Fake News, Infosphere, Misinformation, Cyber Domain, Covid-19 Disinformation, Artificial Intelligence, Bots*

ISSN 1445-3312 Print/ISSN 1445-3347 Online

Introduction

Misinformation is generally defined in the *Oxford Dictionary* as the act of giving wrong information regarding something, or wrong information itself. The scope of misinformation and the impact it has on modern society is significant. While the spreading of rumours and false information has been around since the beginning of humankind (Burkhardt 2017), modern technology has amplified the impact of false information and has turned it into what we now refer to as misinformation and disinformation. When one considers that once information was 'etched into stone' only by the most learned, passed down from one generation to another mostly through recitation (Burkhardt 2017), modern information and by extension its dissemination have also evolved, becoming more efficient as we have become more dependent on the cyber information domain. There can be no doubt that the innovation of the physical printing press is becoming more obsolete in a digitally driven world (Burkhardt 2017). News consumption through broadcast television and online resources is rising rapidly (Truong 2020), creating a whole new infosphere: one where false information can be distributed more easily and can reach a variety of people faster. As society takes evolutionary steps involving information, new doors have been opened for misinformation that need to be examined and controlled carefully to allow negative consequences to be minimized.

The Internet and social media provide the platform for rapid information dissemination, creating a new breeding ground for misinformation and disinformation that can be easily shared between very different types and groups of people. Since the 2016 U.S. election, public interest has grown around the dissemination of misinformation and disinformation in cyberspace facilitating increased research on the topic (Allen *et al.* 2020). This has been accredited to the new understanding that misinformation and fake news increased political polarization, undermined democracy, and reduced public trust in general (Allen *et al.* 2020).

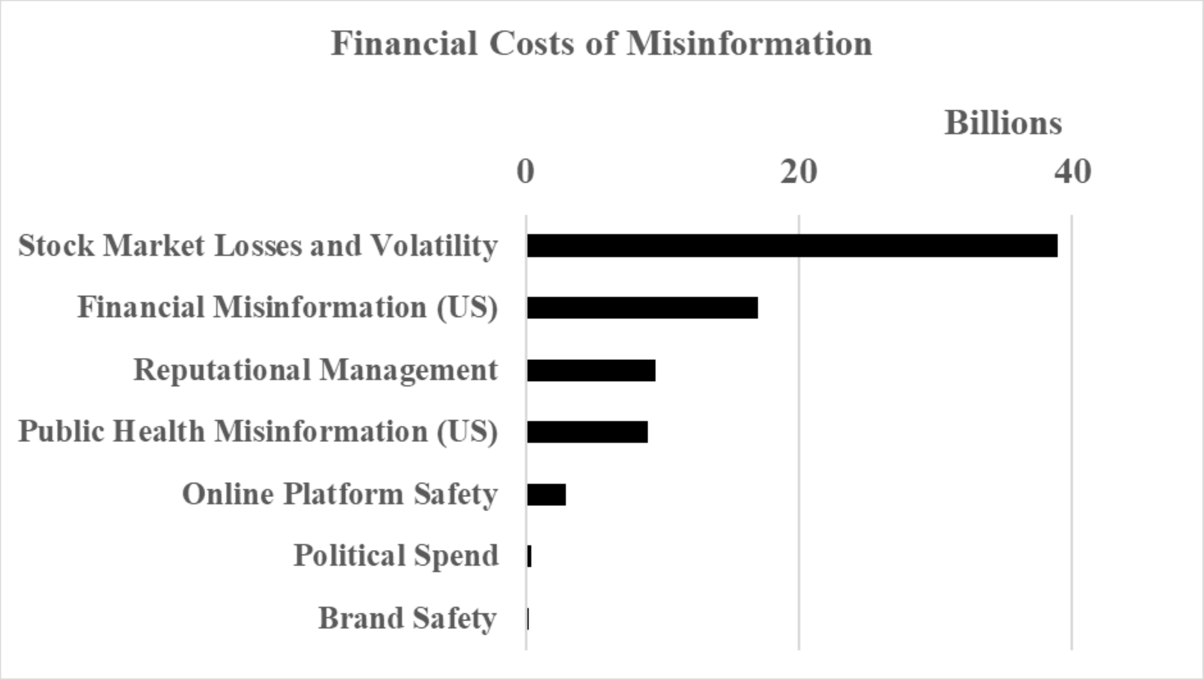

Figure 1: Statistics showing the financial cost of misinformation on various sectors. Note "(US)" signifies within the US only (Cavazos 2019)

As seen in **Figure 1**, the total cost of misinformation across multiple sectors is significant, highlighting the threat misinformation poses to society, in terms of financial costs (amongst other issues). This observation is shared by the World Economic Forum (WEF), which ranked the spread of misinformation amongst top risks to the modern globalised and interconnected world in its 'Freedom on the Net' report of 2017 (Kelly 2017). The above data from 2019 does not take into consideration the impact the global Covid-19 pandemic of 2020, which has seen an increase in misinformation and disinformation globally, and which potentially increases future financial losses. Moreover, misinformation and disinformation pose real risks on social media, which can be evidenced particularly during the Covid-19 pandemic.

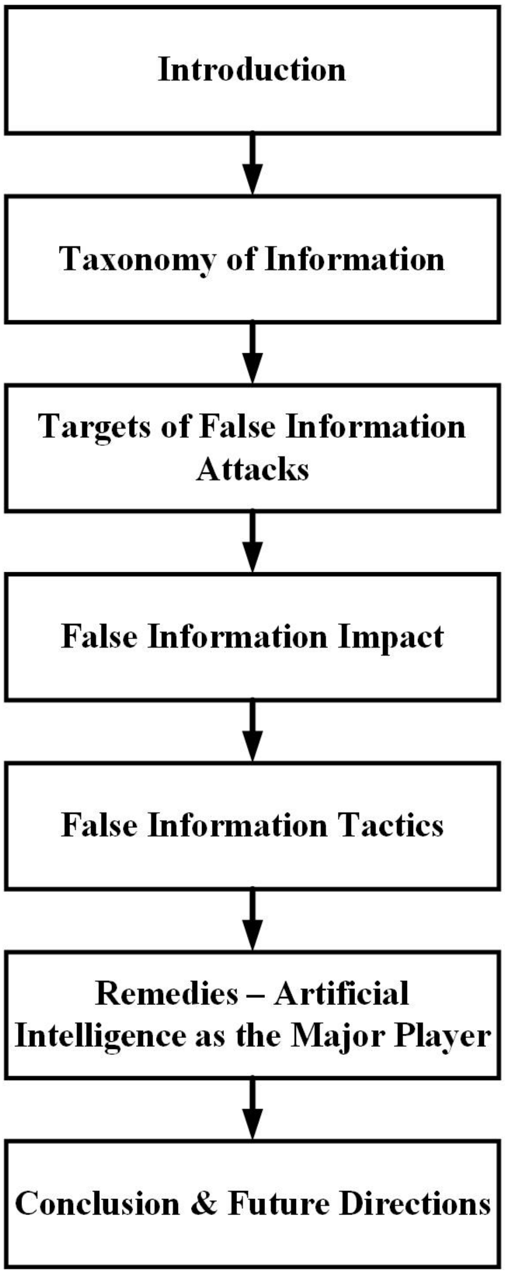
Introduction
Taxonomy of Information
Targets of False Information Attacks
False Information Impact
False Information Tactics
Remedies – Artificial Intelligence as the Major Player
Conclusion & Future Directions

Figure 2: False Information overview and structure of article

This article examines misinformation to highlight the need for countermeasures and misinformation detection systems. Attention then moves to currently available misinformation detection techniques, critically evaluating available models to determine the effectiveness and efficiency when compared to other current solutions. This paper aims to make suggestions as to what (at the time of writing) are the most effective countermeasures and how they should be implemented to reduce the impact misinformation can have. Continuing from this, suggestions are made as to how improvements on current available systems may be achieved, as well as highlighting potential weaknesses. The findings highlight gaps within current knowledge and focus on areas that may need more attention. This tangible outcome can be used to improve and focus the field on current identified issues, serving as somewhat of a homing beacon for other studies. Understanding gaps in knowledge is vital for any researcher, thus the importance and benefit of this paper.

Taxonomy of Information

Information is known to be the fundamental inherently true notation, being the basic ingredient for developing knowledge (Søe 2019). With information being the true part of knowledge (Søe 2019), false information (or the dark side) pertains to the falsification of said information, the counterpart to the good and true information we know of (Søe 2019). False information can only be effective if it is not easily identified as such, else it would just be discarded (Kumar & Shah 2018). With further improvements in how information is shared, humans have become less accurate in identifying false information: at present only accuracies of 53-78% in detection are achieved, particularly relevant in regard to false information such as hoaxes, fake reviews, and fake news (Kumar & Shah 2018). Both trained and casual readers can be fooled easily when content is well-written, substantial, and referenced (Kumar & Shah 2018).

False information can be categorized into two parts: misinformation and disinformation. Misinformation is created and used without the intent to mislead; it is simply inaccurate information (Søe 2018); disinformation, on the other hand, is created with the intent to deceive (Kumar & Shah 2018; Søe 2018). The difference between the two forms of fake information is distinguished through intent and the prospective target. Both of these forms have negative consequences; however, disinformation is considered to be more dangerous due to underlying malicious intent (Kumar & Shah 2018).

These two main types of false information can be further broken down into different names, by their level of veracity and deception (Babcock, Beskow & Carley 2019). False information can be further categorized into the following: false information that intends to inform (satire, parable, storytelling) (Babcock, Beskow & Carley 2019), false information that intends to mislead and misinform (propaganda, hoax, clickbait) (Ghanem, Rosso & Rangel 2020), and false information that tests one’s understanding of common thinking (jokes) (Babcock, Beskow & Carley 2019). A further understanding of these different categories is key.

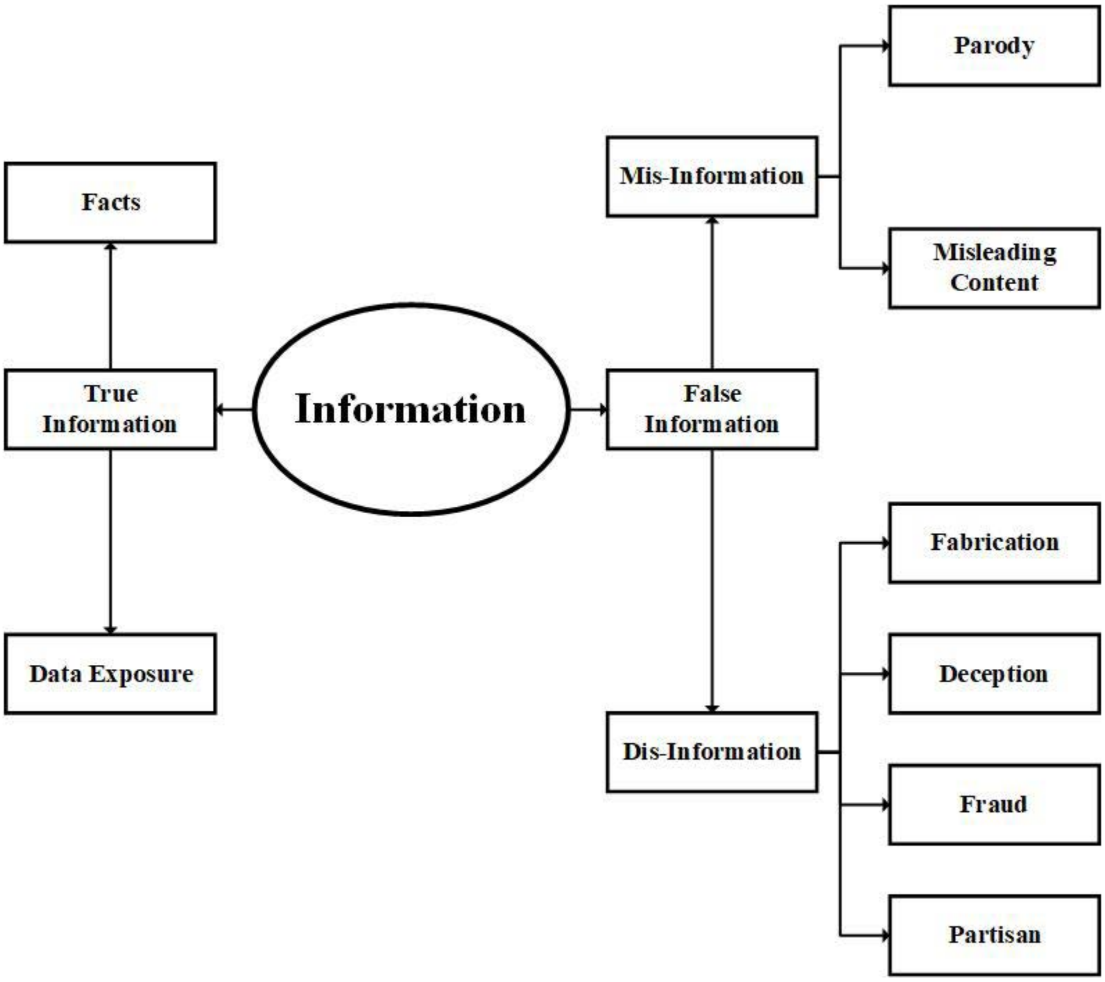

Figure 3: Taxonomy of information

The taxonomy in **Figure 3** reflects how information can be seen in the context of false information. True information is dependent on facts and exposures, whereas the false information has many areas as shown above.

Targets of False Information Attacks

Online platforms

With online content allowing people to form and to express personal views and opinions, its openness has promoted the use of false information (Qi *et al.* 2019). False information has become more prevalent, with the average person having more possibility to view stories and leaving them at greater risk of exposure to information tampering (Qi *et al.* 2019). With false information more likely to be shared than true information (Singh, Ghosh & Sonagara 2021), an understanding of how said information has been tampered with is an ongoing concern.

Online platforms such as web sites, blogs, and social media have become affected and tampered with and have become a main platform for false information dissemination (Rodríguez *et al.* 2020), accredited especially to the progression of replacing expert advice with misleading content (Lewandowsky *et al.* 2012). Online platforms (such as Facebook and Twitter) allow information to circulate freely and (mostly) uncontested. On these platforms, fake information is being viewed as more novel and likely to be shared more often than true information, often becoming associated with events and emergencies, displaying an ability to appear during emerging issues or events (Rodríguez *et al.* 2020). The reliance on the Internet as a source of information has become a risk, due to the reliability of information being highly variable with content often severely lacking in accuracy. Social media platforms present a unique effective medium due to their popularity, spreading false information quickly and efficiently (Lewandowsky *et al.* 2012).

Images

Images have become targets of tampering to spread false information, manipulating public opinion, (Gaborini *et al.* 2014) where modifying images for malicious purpose or using an original image in the wrong context becomes commonplace (Maigrot *et al.* 2017). Through tampering/modifying images to create fake information, malicious actors can use different methods such as: duplication of parts of images (copy-move attack), inserting a region in another image (copy-paste/splicing attack), or deleting a region from the image (painting or seam carving) (Maigrot *et al.* 2017). The use of altering media can have multiple implications on social and legal aspects, especially as digital images are easily altered to create realistic images to fool the human eye (Gaborini *et al.* 2014).

Electronic health records

Within the healthcare industry, clinical laboratories aim to achieve a marginal error of only 5%, translating to only 1 in 20 false results in clinical tests (Burnum 1989). Even with filing and typing of charts and other important medical documents, it is not uncommon for information to be incorrect due to the reality of medical life; thus, misinformation does occur with respect to patients and their medical records (Burnum 1989). Misinformation and disinformation in medical records can arise for many reasons: errors on what kind of operations patients have had from lack of understanding of the procedure, a block from memory if they had a certain disease (cancer), or deliberately withholding important facts and information such as drugs, alcohol, and sexual practices (Burnum 1989). Medical records are also vulnerable to tampering from external attackers. External parties can tamper with the information of medical images from volumetric (3D) medical scans through deep-learning practices (Elovici 2019). From these practices, attackers can add or remove evidence of medical conditions to change the diagnosis of a patient for a range of motives (political, fraud, terrorism) (Elovici 2019).

Government and politics

Governments and politicians are often the targets of false information campaigns, in connection with election campaigns, or due to policies and agendas associated with these targets (Lewandowsky *et al.* 2012). The tampering of information by politicians can cause scenarios where fake information is produced to push agendas and/or to win an election. Such examples include the Weapons of Mass Destruction (WMD) ‘evidence’ in Iraq, which was used to justify the U.S. led invasion in 2003 (Lewandowsky, Ecker & Cook 2017) and the U.S. public health care debate surrounding The Affordable Care Act (Obama Care) (Lewandowsky *et al.* 2012). Other types of political tampering can include the production of false information during an election campaign to be used against an opponent as a form of offensive information warfare (Lewandowsky *et al.* 2012). Tampering can also be used by corporations that may become invested in politics for many reasons and may attempt to influence public debate or opinion, creating false information on government policies, such as the removal of regulation burdens on industries (tobacco or fossil fuel) (Lewandowsky *et al.* 2012). It is often difficult to identify specific information tampering due to the difficulty in determining the difference between what is true or false (Lewandowsky *et al.* 2012).

False Information Impact

The result of the spread of false information in the new digital age of online media has changed the way people are getting information. It was identified by the World Economic Forum (WEF) as one of the main threats to modern society (Del Vicario *et al.* 2016) and as such the consequences of false information can be potentially disastrous for modern democratic societies (Kyza *et al.* 2020). Consequences arising from false information can potentially impact many different aspects of modern life. They range from people's lifestyles to a state's global relations (Agarwal *et al.* 2020), a problem which is being amplified as misinformation often spreads faster than corresponding true information (Kyza *et al.* 2020).

The rise of social media has already undermined the public's reliance and confidence in traditional media and governments (Johnson & Kaye 2015) as the public seems to increasingly trust new mediums, thus creating over-reliance on such media forms as people tend to engage with political groups and webpages on an increased frequency (Johnson & Kaye 2015). With the rise in social media, false information can spread rapidly and such spread is often unchecked (Wang *et al.* 2019) with users sharing false stories with their followers both intentionally or unintentionally (Shao *et al.* 2018). Since the beginning of the COVID-19 pandemic, this has lead to false information spreading 'faster than the virus itself' (Cuan-Baltazar *et al.* 2020), an alarming prospect for national and global health and pandemic resilience.

Panic buying

During the Covid-19 pandemic, panic buying became more widespread with items such as toilet paper and other 'emergency' supplies falling into short supply (Taylor 2021). These panic buying trends are a direct consequence of the circulation of fake news and misinformation on social media and traditional media portals (Hossain, Ferdous & Siddiqee 2020). The Covid-19 pandemic was not the only event that has caused panic buying in history, as evidenced in the 2003 SARS outbreak, which also resulted in panic buying of salt, rice, and other perishable supplies in Hong Kong and China (Leung *et al.* 2021), demonstrating that multiple events can develop into panic buying situations (Naeem 2021).

But what exactly is panic buying? Panic buying is the process of purchasing food, supplies, and other necessities from vendors in large quantities resulting in a limitation or elimination of availability (Loxton *et al.* 2020), creating product scarcity (Barnes, Diaz & Arnaboldi 2021). Panic buying differs from hoarding and compulsive buying, as hoarding is the excessive accumulation of items over several years, with difficulty discarding said items (Taylor 2021) and compulsive buying differs again, being the urge to buy particular items with regret after (Taylor 2021).

With the events during and leading up to lockdown, acts of panic buying became exacerbated partly due to misinformation and rumours being spread about shortages over digital communication media (Martin-Neuninger & Ruby 2020). These rumours resulted in panic buying which progressed into hoarding behaviour due to uncertainty, allowing people to regain the feeling of control (Usher, Durkin & Bhullar 2020) and such an exacerbation influenced others in their behaviour during uncertain times (Martin-Neuninger & Ruby 2020) creating an amplifying effect.

National security

The spread of misinformation within a nation can match the spread of the Covid-19 virus itself. This is achieved through multiple communication channels through which humans can interact (Pang & Ng 2017), development of technologies and ways users search for news, overall increasing its prevalence (Mills & Robson 2019).

To combat the spread of misinformation on social media platforms, governments have proposed and have subsequently passed new laws (Rodrigues & Xu 2020). These laws give the government more power to hold technology companies and individuals more accountable when using/ spreading misinformation (Haciyakupoglu *et al.* 2018). Governments create countermeasures as a consequence of misinformation impacting national security, ranging from laws to media literacy campaigns, podcasts, and handbooks (Norri-Sederholm *et al.*2020).

These laws, used to criminalise the spread of misinformation, have other consequences with civil liberties, creating concerns surrounding the abuse of said regulations on journalists and detractors within the banner of ‘regulation’ (Rodrigues & Xu 2020). This includes countering the impact of automated media accounts (bots) (Haciyakupoglu *et al.* 2018). In contrast, other countries have gone to different extremes, sanctioning and ‘turning off’ the Internet completely (Norri-Sederholm *et al.* 2020) to control the flow of the information (as seen in India during the COVID-19 pandemic) (Rodrigues & Xu 2020), thus reducing the spread of misinformation but also health information at the same time, an unfortunate reality (Rodrigues & Xu 2020).

This can allow policies and opinions to influence the population through other government actors. This can be seen in Russia, where operatives used/uploaded social and political content onto social media services during the U.S. 2016 election between Hillary Clinton and Donald Trump (Haciyakupoglu *et al.* 2018), an action which is believed to have influenced the result of said election (Girgis, Amer & Gadallah 2018). Other instances have inflamed racism and anti-immigrant sentiments (Liu 2019), causing serious harm. This extends into areas that include people’s democracy and free speech (Katsirea 2018).

Such scenarios create a lack of general trust in a nation’s security (Norri-Sederholm *et al.* 2020) and potentially influences voters and community leaders who are susceptible (Gallacher *et al.* 2018). Such betrayal of trust and confidence can lead to public unrest, violence, and disruption to the economy and social fabrics, resulting in damage to the collective identity of a nation (Neo 2020). Negative societal impacts of misinformation on national security can be (but are not limited to) political polarization, loss of trust in government, mainstream media, and future elections and the discreditation of news organizations (Bellutta, King & Carley 2021). This however does allow future generations to strengthen their resistance to misinformation as they have seen its impact first-hand, making them potentially less prone to misinformation going forward (Norri-Sederholm *et al.* 2020).

False Information Tactics

Information is at the epicentre of both human interaction and societal growth. However, it is hard in the modern interconnected multi-media world to become numbed and blind to various sources and forms of tampered and fake content. As the world progresses and advances technologically in every aspect of life, the ability to deceive is also progressing. The adaptation of technology for nefarious means has impacted information on both ends of the spectrum, whether technology adaptions in the social engineering realm of fake information (phishing attacks, spear phishing, and cyber fraud adaptions of this) are concerned all the way to far more technological heavy adaptions of Deepfakes and algorithm-generated human writings. Thus, it is imperative to understand how information is tampered with and turned into fake information.

Generative Adversarial Networks (GANs)

GANs provide an innovation in machine learning that allow for the simultaneous training of two competing models. These models are composed of two generators, one working to create a generator model, and one input with training data; both of these generators then feed into one single discriminator (Kwok & Koh 2020; Wang *et al.* 2017) This allows for GANs to be involved in the process of altering data that gets introduced to it, or alternatively acting as a system for checking authenticity of data through complex checks and balances set by the user. This altering may take many forms, with many unique practical and legitimate applications; unfortunately, it also allows nefarious uses (Kietzmann *et al.* 2020). GANs provide a unique and powerful tool, by allowing a high level of complexity within data to be mapped and vectors from the latent space to be identified and transitioned to the product (Creswell *et al.* 2018). This innovation could lead to greater benefit in modern society; however, it is incredibly dangerous in the hands of cyber criminals.

Deepfakes

Deepfake technology can be leveraged to create falsified videos, transforming elements and identities within video footage. The production of these videos utilises GANs; however, interplay between generator and discriminator takes a slightly different form. Within the creation of video deepfakes, the original image and the target fake are trained into two separate encoders, with production taking place in one central decoder (Guera & Delp 2018; Verdoliva 2020). Video-based deepfakes began emerging in 2019, as they appeared in an open-source form on Reddit. Initially, this software was used to create fake celebrity-based pornography, but it has since moved to become a concern from grassroot levels to a tool to influence and create misinformation at nation-state levels. The creation of video-based deepfakes takes four main forms: Entire Face Synthesis, Identity Swap, Attribution Manipulation, and Expression Swap.

Remedies—Artificial Intelligence as Key

If misinformation can pervade society without any control, the effects could be cataclysmic. In 2016, the lead up to the U.S. Presidential election saw as many as 529 different rumours spreading through Twitter (Guo *et al.* 2019), while, in 2018, fake news stories were used to contribute and to create inter-ethnic tensions in Africa, where the circulation of a fake video entitled "Somalis pushed into shallow grave Ethiopia" (BBC 2018) led to violent clashes between two ethnic groups in Ethiopia. These are just two examples that highlight the necessity for prompt and swift detection of malicious digital information and the need to be able to separate disinformation, fake news, rumour, urban legend, spam, and troll information (Wu *et al.* 2019) from the truth, in order to take action.

It is important to briefly touch on current remedies and detection systems, to assess what is effective and what could be used to counter misinformation. Current supervised learning systems have their shortcomings, where response time remains slow as systems are still reliant on final human decision making (Jain, Sharma & Kaushal 2016); however, transitioning to more automated deep learning-based systems requires the availability of vast quantities of verifiable and authentic training databases (Guo *et al.* 2019).

Supervised artificial intelligence algorithms have been explored for an application with respect to online social media as recently as 2020 (Ozbay & Alatas 2020). Such an application can be incredibly efficient when properly supervised and when parameters are clearly set. Yet, the scope and framework of misinformation are ever-changing in the fast-paced modern world: any effort vested into detection, an equally matched effort is placed upon creating misinformation. This leads to an issue where the system requires constant input from a human, while also needing human validation as a final step in the detection chain. As misinformation can spread at an uncontrollable rate, it is a necessity to not only be able to identify misinformation, but to be able to do it in a timely way which stops misinformation in its tracks, without any lag phase. Another drawback of artificial intelligence for misinformation detection is that the technologies themselves are more reliable and operate at a fast enough speed to be suited to text-based content. With the increasing appearance of deepfakes and visual-based misleading content, these systems will need to reflect on the pace and scope of modern media.

Machine Learning (ML) models provide a subset of the field of Artificial Intelligence, and while they have been shown to have similar or even better accuracy in their identification of misinformation (Ahmed, Traore & Saad 2017; Reis *et al.* 2019; Aphiwongsophon & Chongstitvatana 2018), they have also shown to follow similar drawbacks to AI approaches of misinformation detection. Given that these processes are very reactionary in their approach in identifying misinformation, both require human intervention to make the final decision while also defining the characteristics of the algorithm and the information they are seeking to identify. Machine learning is not suitable when applied to video-based content; due to ML being more text-based, it allows for much more logical based arguments.

Dataset	Number of Instances	References
BuzzfeedPolitical	120	(Silverman *et al.* 2016)
LIAR	12.8k	(Wang 2017)
CREDBANK	4856	(Mitra 2015)
FakeNewsNet	422	(Shu, Wang & Liu 2019)
Twitter	1111	(Ma *et al.* 2016)
PHEME	6425	(Aiello *et al.* 2013)
NewsFN-2014	221	(Nan 2015)
Politifact	488	(Bathla, Rani & Aggarwal 2018)
Weibo	816	(Ma & Wong 2018)
YelpChi	67k	(Mukherjee, Liu & Glance 2013)
YelpNYC	359k	(Rayana 2015)
YelpZip	608k	(Rayana 2015)
Twitter	5.5M	(Concone, Morana & Ruocco 2019)

Table 1: Currently available datasets that could be used for the development of learning models

The datasets featured in **Table 1**, above, can be used to program misinformation detection systems and to highlight distinct discrepancies between the number of subjects between the included datasets. This highlights the importance of a more defined and validated singular dataset source. Such a range of datasets highlights the difficulty of comparing detection systems, as it does not allow for direct comparisons. These datasets could be used as the basis for further expansion to form more reliable sets that are applicable to more scenarios. Furthermore, this section provides an original source of currently available datasets, which could be used in the future.

Conclusion and Future Directions in Research

The purpose of this article was to investigate and evaluate the effectiveness of false information operations with some recent examples and to highlight some current remedies to the threat. Based on this research, the authors can conclude that false information has grown over time with new and improved methods consistently developing, leading to a greater spread and increase in overall effectiveness. Different types of false information are subject to modification and the manipulation of information is constantly evolving.

The rise of false information has created many challenges as pointed out in this paper. Consequently, there is a lack of up-to-date conclusive databases of fake content available and the available sources are limited. The discussed datasets are not conclusive in terms of proper data and fake data with the use of them being inconsistent resulting in no validation of results between detection systems. Without a universally accepted model for validating results, the only way false information can be found is through detection. Without any validation, results from the detection of false information will produce false positives and negatives, a common consequence without the missing validation step. Given how the world currently uses human-centered computing (where detection results in system flags which are being evaluated by humans), imperfections and false detections will always exist due to human error and thus this is not a reliable system.

With the focus of research on false information attacks in cyber space, further studies are needed in the researching of different remedies to identify and combat the spread of false information, with the added difficulty that technology that is to be used to detect information must be improved vastly. The creation of misinformation bots without proper analysis and human oversight is not going to change the current shortcomings in countering false information. Overall, false information is not something that has an easy fix or will go anyway anytime in the future. Therefore, new datasets must be created to be used with currently available solutions, allowing them to expand and become more effective moving forward. A detailed application is needed to successfully cope with what is becoming a large-scale, expanding problem that will only continue to worsen in the future if new combative technologies are not developed. A comprehensive collaborative approach to counter false information and involving an interagency approach is needed.

References

Agarwal, A, Mittal, M, Pathak, A & Goyal, LM 2020, 'Fake news detection using a blend of neural networks: An application of deep learning', *SN Computer Science*, vol. 1, no. 3, doi:10.1007/s42979-020-00165-4.

Ahmed, H, Traore, I & Saad, S 2017, 'Detection of online fake news using n-gram analysis and machine learning techniques', *Lecture Notes in Computer Science*, pp. 127-38, Springer International Publishing, Cham, Switzerland.

Aiello, LMPG, Martin, C, Corney, D, Papadopoulos, S, Skraba, R, Goker, A, Kompatsiaris, Y & Jaimes, A 2013, 'Sensing trending topics in Twitter', *IEEE Trans Multimedia*, vol. 15, pp. 1262-82.

Allen, J, Howland, B, Mobius, M, Rothschild, D & Watts, DJ 2020, 'Evaluating the fake news problem at the scale of the information ecosystem', *Science Advances*, vol. 6, no. 14, eaay3539, doi:10.1126/sciadv.aay3539.

Babcock, M, Beskow, DM & Carley, KM 2019, 'Different faces of false', *Journal of Data and Information Quality*, vol. 11, no. 4, pp. 1-15, doi:10.1145/3339468.

Barnes, SJ, Diaz, M & Arnaboldi, M 2021, 'Understanding panic buying during COVID-19: A text analytics approach', *Expert Systems with Applications*, vol. 169, 114360, doi:10.1016/j.eswa.2020.114360.

Bathla, G, Rani, R & Aggarwal, H 2018, 'Improving recommendation techniques by deep learning and large-scale graph partitioning', *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, doi:10.14569/ijacsa.2018.091049.

BBC 2018, *A year in fake news in Africa*, viewed 27 October 2021, <https://www.bbc.com/news/world-africa-46127868>.

Bellutta, D, King, C & Carley, KM 2021, 'Deceptive accusations and concealed identities as misinformation campaign strategies', *Computational and Mathematical Organization Theory*, doi:10.1007/s10588-021-09328-x.

Burkhardt, J 2017, 'Combating fake news in the digital age', *ALA TechSource*, vol. 53.

Burnum, JF 1989, The misinformation era: The fall of the medical record', *Annals of Internal Medicine*, vol. 110, no. 6, pp. 482-4, doi:10.7326/0003-4819-110-6-482.

Concone, FLRG, Morana, M & Ruocco, C 2019, 'Twitter spam account detection by effective labeling, *Proceeding of the Third Italian Conference on Cyber Security*, Pisa, Italy.

Creswell, A, White, T, Dumoulin, V, Arulkumaran, K, Sengupta, B & Bharath, AA 2018, 'Generative adversarial networks: An overview', *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 53-65, doi:10.1109/msp.2017.2765202.

Cuan-Baltazar, JY, Muñoz-Perez, MJ, Robledo-Vega, C, Pérez-Zepeda, MF & Soto-Vega, E 2020, 'Misinformation of COVID-19 on the Internet: Infodemiology study', *JMIR Public Health and Surveillance*, vol. 6, no. 2, e18444, doi:10.2196/18444.

Del Vicario, M, Bessi, A, Zollo, F, Petroni, F, Scala, A, Caldarelli, G, Stanley, HE, Quattrociocchi, W 2016, 'The spreading of misinformation online', *Proceedings of the National Academy of Sciences*, vol. 113, no. 3, pp. 554-9, doi:10.1073/pnas.1517441113.

Gaborini, L, Bestagini, P, Milani, S, Tagliasacchi, M & Tubaro, S 2014, 'Multi-clue image tampering localization', *2014 IEEE International Workshop on Information Forensics and Security (WIFS)*, Atlanta, GA, US.

Gallacher, JD, Barash, V, Howard, PN & Kelly, J 2018, ‘Junk news on military affairs and national security: Social media disinformation campaigns against US military personnel and veterans’, *arXiv [cs.SI]*, <http://arxiv.org/abs/1802.03572>.

Ghanem, B, Rosso, P & Rangel, F 2020, ‘An emotional analysis of false information in social media and news articles’, *ACM Transactions on Internet Technology*, vol. 20, no. 2, pp. 1-18, doi:10.1145/3381750.

Girgis, S, Amer, E & Gadallah, M 2018, ‘Deep learning algorithms for detecting fake news in online text’, *2018 13th International Conference on Computer Engineering and Systems (ICCES),* Cairo, Egypt.

Guera, D & Delp, EJ 2018, ‘Deepfake video detection using recurrent neural networks,' *15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS),* Auckland, New Zealand.

Guo, B, Ding, Y, Yao, L, Liang, Y & Yu, Z 2019, 'The future of misinformation detection: New perspectives and trends’, *arXiv [cs.SI]*, <http://arxiv.org/abs/1909.03654>.

Haciyakupoglu, G, Hui, JY, Suguna, VS, Leong, S & Rahman, MFBA 2018, ‘Countering fake news—A survey of recent global initiatives’, viewed 27 October 2021, <https://think-asia.org/bitstream/handle/11540/8063/PR180307_Countering-Fake-News.pdf?sequence=1>.

Hossain, MS, Ferdous, S & Siddiqee, MH 2020, ‘Mass panic during Covid-19 outbreak: A perspective from Bangladesh as a high-risk country’, *Journal of Biomedical Analytics*, vol. 3, no. 2, pp. 1-3, doi:10.30577/jba.v3i2.40.

Jain, S, Sharma, V & Kaushal, R 2016’, ‘Towards automated real-time detection of misinformation on Twitter’, *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 2015-20, Jaipur, India, doi: 10.1109/ICACCI.2016.7732347.

Johnson, TJ & Kaye, BK 2015, ‘Site effects’, *Social Science Computer Review*, vol. 33, no. 2, pp. 127-44, doi:10.1177/0894439314537029.

Katsirea, I 2018, ‘“Fake news”: Reconsidering the value of untruthful expression in the face of regulatory uncertainty’, *Journal of Media Law*, vol. 10, no. 2, pp. 159-88, doi:10.1080/17577632.2019.1573569.

Kietzmann, J, Lee, LW, Mccarthy, IP & Kietzmann, TC 2020, ‘Deepfakes: Trick or treat?’, *Business Horizons*, vol. 63, no. 2, pp. 135-46, doi:10.1016/j.bushor.2019.11.006.

Kumar, S & Shah, N 2018, ‘False information on web and social media: A survey’, *arXiv [cs.SI],* <http://arxiv.org/abs/1804.08559>.

Kwok, AOJ & Koh, SGM 2020, ‘Deepfake: A social construction of technology perspective’, *Current Issues in Tourism*, pp. 1-5, doi:10.1080/13683500.2020.1738357.

Kyza, EA Varda, C, Panos, D, Karageorgiou, M, Komendantova, N, Perfumi, SC, Shah, SIH & Hosseini, AS 2020, 'Combating misinformation online: Re-imagining social media for policy-making', *Internet Policy Review*, vol. 9, no. 4, doi:10.14763/2020.4.1514.

Leung, J, Chung, JYC, Tisdale, C, Chiu, V, Lim, CCW & Chan, G 2021, 'Anxiety and panic buying behaviour during COVID-19 pandemic—A qualitative analysis of toilet paper hoarding contents on Twitter', *International Journal of Environmental Research and Public Health*, vol. 18, no. 3, p. 1127, doi:10.3390/ijerph18031127.

Lewandowsky, S, Ecker, UKH, Seifert, CM, Schwarz, N & Cook, J 2012, 'Misinformation and its correction', *Psychological Science in the Public Interest*, vol. 13, no. 3, pp. 106-31, doi:10.1177/1529100612451018.

—, — & Cook, J 2017, 'Beyond misinformation: Understanding and coping with the "post-truth" era', *Journal of Applied Research in Memory and Cognition*, vol. 6, no. 4, pp. 353-69, doi:10.1016/j.jarmac.2017.07.008.

Liu, H 2019, 'A location independent Machine Learning approach for early fake news detection', *2019 IEEE International Conference on Big Data*, Los Angeles, CA, US.

Loxton, M, Truskett, R, Scarf, B, Sindone, L, Baldry, G & Zhao, Y 2020', 'Consumer behaviour during crises: Preliminary research on how Coronavirus has manifested consumer panic buying, herd mentality, changing discretionary spending and the role of the media in influencing behaviour', *Journal of Risk and Financial Management*, vol. 13, no. 8, p. 166, doi:10.3390/jrfm13080166.

Ma, JGW, Mitra, P, Kwon, S, Jansen, BJ, Wong, KF & Cha, M 2016, 'Detecting rumours from microblogs with recurrent neural networks', *International Joint Conference on Artificial Intelligence*, New York, NY, US.

—& Wong, KF 2018, 'Rumour detection on Twitter with tree-structured recursive neural networks', *Proceedings of 56th Annual Meeting of the Association for Computational Linguistics,* vol. 1, Long Papers, Melbourne, Australia, bll 1980-1989, doi: 10.18653/v1/P18-1184.

Maigrot, C, Kijak, E, Sicre, R & Claveau, V 2017, 'Tampering detection and localization in images from social networks: A CBIR approach', *Image Analysis and Processing - ICIAP 2017*, pp. 750-61, Springer International Publishing, *ICIAP 2017 - International Conference on Image Analysis and Processing*, Catane, Italy, bll 1–11, <https://hal.inria.fr/hal-01623105>.

Martin-Neuninger, R & Ruby, MB 2020, 'What does food retail research tell us about the implications of Coronavirus (COVID-19) for grocery purchasing habits?', *Frontiers in Psychology,* vol. 11, doi:10.3389/fpsyg.2020.01448.

Mills, AJ & Robson, K 2019, 'Brand management in the era of fake news: Narrative response as a strategy to insulate brand value', *Journal of Product & Brand Management*, vol. 29, no. 2, pp. 159-67, doi:10.1108/JPBM-12-2018-2150.

Mirsky, Y, Mahler, T, Shelef, I & Elovici, Y 2019, 'CT-GAN: Malicious tampering of 3D medical imagery using deep learning', *28th {USENIX} Security Symposium ({USENIX} Security 19),* Santa Clara, CA, US, <https://www.usenix.org/conference/usenixsecurity19/presentation/mirsky>.

"Misinformation", Oxford English Dictionary, Oxford University Press, Oxford, UK.

Mitra, TGE 2015, 'CREDBANK: A large-scale social media corpus with associated credibility annotations', *Proceedings of the Ninth International AA AI Confernece on Web and Social Media*, vol. 9, no. 1, pp. 258-67, viewed 30 January 2022, <https://ojs.aaai.org/index.php/ICWSM/article/view/14625>.

Mukherjee, AVV, Liu, B & Glance, N 2013, 'What Yelp fake review filter might be doing?', *Seventh International AAAI Conference on Weblogs and Social Media*., vol. 7, no. 1, pp. 409-18, viewed 30 January 2022, <https://ojs.aaai.org/index.php/ICWSM/article/view/14389>.

Naeem, M 2021, 'Do social media platforms develop consumer panic buying during the fear of COVID-19 pandemic', *Journal of Retailing and Consumer Services*, vol. 58, p. 102226, doi:10.1016/j.jretconser.2020.102226.

Nan, CJ 2015, 'Social network analysis of TV drama characters via deep concept hierarchies', *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Paris, France.

Neo, R 2020, 'The securitisation of fake news in Singapore', *International Politics*, vol. 57, no. 4, pp. 724-40, doi:10.1057/s41311-019-00198-4.

Norri–Sederholm, T, Norvanto, E, Talvitie–Lamberg, K & Huhtinen, AM 2020, 'Misinformation and disinformation in social media as the pulse of Finnish national security', *Advanced Sciences and Technologies for Security Application*, pp. 207-25, Springer International Publishing, Cham, Switzerland.

Ozbay, FA & Alatas, B 2020, 'Fake news detection within online social media using supervised artificial intelligence algorithms', *Physica A: Statistical Mechanics and its Applications*, vol. 540, p. 123174, doi:10.1016/j.physa.2019.123174.

Pang, N & Ng, J 2017, 'Misinformation in a riot: A two-step flow view', *Online Information Review*, vol. 41, no. 4, pp. 438-53, doi:10.1108/OIR-09-2015-0297.

Qi, P, Cao, J, Yang, T, Guo, J & Li, J 2019, 'Exploiting multi-domain visual information for fake news detection', *2019 IEEE International Conference on Data Mining (ICDM)*, Beijing, China.

Rayana, SAL 2015, 'Collective opinion spam detection: Bridging-review networks and metadata', *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Sydney, Australia.

Reis, JCS, Correia, A, Murai, F, Veloso, A & Benevenuto, F 2019, 'Explainable machine learning for fake news detection', *Proceedings of the 10th ACM Conference on Web Science - WebSci '19*, Boston, MA, US.

Shu, K, Wang, S & Liu, H 2019, 'Beyond news contents', *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining,* New York, NY, US, Association for Computing Machinery (WSDM '19), bll 312-20, doi: 10.1145/3289600.3290994.

Silverman, CSL, Shaban, H, Hall, E & Singer-Vine, J 2016, *Hyperpartisan Facebook pages are publishing false and misleading information at an alarming rate*, viewed 30 January 2022, <https://www.buzzfeednews.com/article/craigsilverman/partisan-fb-pages-analysis>.

Singh, VK, Ghosh, I & Sonagara, D 2021, 'Detecting fake news stories via multimodal analysis', *Journal of the Association for Information Science and Technology*, vol. 72, no. 1, pp. 3-17, doi:https://doi.org/10.1002/asi.24359.

Søe, SO 2018, 'Algorithmic detection of misinformation and disinformation: Gricean perspectives', *Journal of Documentation*, vol. 74, no. 2, pp. 309-32, doi:10.1108/jd-05-2017-0075.

—2019. 'A unified account of information, misinformation, and disinformation', *Synthese,* doi:10.1007/s11229-019-02444-x.

Taylor, S 2021, 'Understanding and managing pandemic-related panic buying', *Journal of Anxiety Disorders*, vol. 78, p. 102364, doi:10.1016/j.janxdis.2021.102364.

Truong, A & Dove, C 2020, *Media content consumption survey,* viewed 27 October 2021, <https://www.communications.gov.au/file/51442/download?token=AMiuFZlr>.

Usher, K, Durkin, J & Bhullar, N 2020, 'The COVID-19 pandemic and mental health impacts', *International Journal of Mental Health Nursing*, vol. 29, no. 3, pp. 315-18, doi:10.1111/inm.12726.

Verdoliva, L 2020, 'Media forensics and DeepFakes: An overview', *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 5, pp. 910-32, doi:10.1109/jstsp.2020.3002101.

Wang, K, Gou, C, Duan, Y, Lin, Y, Zheng, X & Wang, FY 2017, 'Generative adversarial networks: Introduction and outlook', *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 4, pp. 588-98, doi:10.1109/jas.2017.7510583.

Wang, WY 2017, '"Liar, liar pants on fire": A new benchmark dataset for fake news detection', arXiv [cs.CL], <http://arxiv.org/abs/1705.00648>.

Wang, Y, Mckee, M, Torbica, A & Stuckler, D 2019, 'Systematic literature review on the spread of health-related misinformation on social media', *Social Science & Medicine*, vol. 240, p. 112552, doi:10.1016/j.socscimed.2019.112552.

Wu, L, Morstatter, F, Carley, KM & Liu, H 2019, 'Misinformation in social media', *ACM SIGKDD Explorations Newsletter*, vol. 21, no. 2, pp. 80-90, doi:10.1145/3373464.3373475.