DEPARTMENT OF COMPUTER SCIENCE

The Impact of Social Media on Political Radicalisation

A Semi-Systematic Literature Review

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Master of Engineering in the Faculty of Engineering.

School of Computer Science, Electrical and Electronic Engineering, and Engineering Maths (SCEEM)

Thursday 4th May, 2023

Word Count: 13,680

Abstract

As social media revolutionised the way people communicate with each other and interact with information, it has also changed how terrorists and other radical groups recruit and mobilise. This dissertation wants to explore this phenomenon by understanding what impact social media has on political radicalisation through a semi-systematic review of current academic literature. To do this, I tried to determine who was most at risk of radicalisation through social media, and I found that there is no academic consensus on the most affected demographic other than relative youth. I investigated which social media features and platforms were most vulnerable to radical exploitation, and I determined that *Facebook* and *Telegram* were the most mentioned platform, while the most exploited features were user-generated content, measurable engagement metrics, anonymity, and overall accessibility. When I looked for areas of academic consensus, I found significant overlap and agreement regarding the most common perpetrators of radicalisation (the far-right and Islamic) and examples of common types of radical content (conspiracy theories, hate speech, dog whistles, and extremist propaganda). I also discovered areas of contradiction with the literature, specifically regarding the most vulnerable demographic, the existence and efficacy of echo chambers, and the extent to which social media is a driving factor of radicalisation. Lastly, I identified currently unresearched aspects of the overlap between radicalisation and social media, for future academia to explore.

Table of Contents

List of Figures

List of Tables

Ethics Statement

"This project did not require ethical review, as determined by my supervisor, Dr. Ola Michalec.

Supporting Technologies

- I used *Endnote* in the literature screening process to keep track of the papers I was reading and eventually including in as sample papers.
- *Notion* was used for the initial notes and to record all relevant quotes in a searchable manner.
- I used *Microsoft Excel* to systematise my notes and create all figures and tables.

Chapter 1: Introduction

Social media revolutionised how we consume, share, and view information. Thanks to social media platforms, users can interact with one another from across the globe, create and share their own content with like-minded individuals, and access information that they could not access before. And it is clear that more and more people are taking advantage of it. In 2021, 4.48 billion people used social media, and that number continues to increase, (since 2015 at an average rate of 12.5% [31]).

Unfortunately, social media platforms are not used exclusively by well-meaning individuals, and they can be invaluable resources for the ill-intentioned. This is apparent when considering events like the Christchurch shootings in 2019, where a gunman opened fire in a mosque and murdered 50 people, injuring 50 more. He live-streamed the attack on Facebook and later Twitter and YouTube when the original was taken down. The Facebsook account even had a violent manifesto posted on it an hour before the attack [32]. Facebook did remove both the live stream and the manifesto, but they were reposted repeatedly and easily accessible online, despite being hateful and violent propaganda. As a result, in a public letter two weeks after the shooting [33] Facebook vowed to take stronger steps to ban hate speech on their platform and use their AI to identify hate groups. Their attempts have not been successful, as over half of the white supremacist groups on Facebook's own watchlist remained on the platform as of 2021 [34], and thus this remains a pressing issue to this day.

While the Christchurch shooting was what first attracted my attention to the uses of social media in radicalisation, it is merely an example and warning of the possible ramifications of social media platforms failing to stifle extremist content. A significant portion of terrorists are already using social media. A study from 2017 [35] shows that 76% of U.K. terrorists used online resources to plan or execute their radical behaviour. Another about IS in the United States found that 80% of terrorists studied used social media in some part of their plans [36]. And social media has been proven to have an active and increasing role in the radicalisation and mobilisation process of these extremists too, with 55.3% of terrorists in the PRIUS database between 2005-2016 having social media as either a primary or secondary mode of radicalisation [37].

Political radicalisation on social media manifests beyond live-streamed terror attacks and white supremacy. Terrorists, hate groups, and politicians alike have begun using social media to garner and mobilise sympathisers en masse. They've been exploiting the nature and affordances of the platforms to spread their propaganda through fake engagement, bots, algorithm exploitation, and fake news. In the current technological climate, where getting news from social media is common practice—nearly half of all Americans reported getting digital news from social media at least sometimes [38]—convincing propaganda of this sort can easily find a receptive audience if left alone.

Even beyond that, within the papers analysed for this dissertation, other modes of radicalisation were found on social media, including conspiracy theory blogs that promote political agendas, activist accounts that hinge on bigoted ideals, and moderated groups to indulge in and spread hate speech and political misinformation. The exact effectiveness or validity of any of these radicalisation methods is not universally agreed upon, but they all exist and can become popular within their niche, thus they are worth investigating.

While Facebook was fast and removing the Christchurch live stream and their AI system removes some extreme content, the most effective way of minimising tragedy and violence is to intervene before it

occurs, targeting the earlier stages of political radicalisation. To do this, the entire process needs to be understood. Thus, any social media platform has an ethical responsibility to at least attempt to understand how their platform is used to facilitate radicalisation, in order to help limit it. Ultimately, understanding the extent to which social media platforms are an effective means of recruitment and radicalisation, how the process itself operates online, and the platform-specific features that allow it to happen, is essential to create frameworks and algorithms to better detect and prevent it.

In this dissertation, I will explore this issue through a semi-systematic literature review. While the detection and prevention tools are beyond the scope of the dissertation, the findings of the studies will provide an academic basis for our current understanding of social media's impact on radicalisation, on which these tools can be based upon.

1.1 Research Question and Aims

This study wants to explore the current academic understanding of the overlap between social media and political radicalisation though a semi-systematic literature review, answering the research question:

RQ: What impact does social media have on political radicalisation?

This paper aims to

i. Determine who is most likely to be affected by radicalisation through social media.
ii. Understand how radicalised individuals and extremist recruiters use social media to radicalise others and themselves.
iii. Identify academic consensus and possible areas of contention or contradiction on what impact social media has on radicalisation.
iv. Find any gaps in the current research on the topic and possible new areas to explore.

1.2 Structure of this Paper

Since this paper concerns itself with the area where social media meets radicalisation, in the *Contextual Background* section I will be defining both terms and the main theories and frameworks that surround them independently. Then, in Section 3, I will explain the methodology of this dissertation; this includes the definition of a semi-systematic literature review, the exclusion and inclusion principles that yielded the sample, and how it was analysed. The *Results and Discussion* Section will apply the method and attempt to explain the outcome. In Section 6, I will evaluate the limitations of the papers reviewed and the review itself, and I will offer possible expansions to this dissertation. Lastly, I will summarise the findings, and answer the aims and research question.

Chapter 2: Contextual Background

2.1 Social Media

Although social media has become a substantial area of research in both computer science and the social sciences, there is no unique, formal, and mutually agreed-upon definition that holds between disciplines [39]. However, a review on the topic [40] found some general consensus. The role of social media has always been the enabling of human interaction and connection with other users, and modern definitions (post-2010) mostly focus on the creation and sharing of user-generated content as the defining feature. However, it is important to understand the minutiae of social media beyond this, how it evolved in time, and what definitions are most used in papers, to understand how their results interact with one another and what definition is most suited for this dissertation.

Because of the speed with which technology and social media specifically are evolving, our understanding of social media varies widely within just a few years. The first formal definition (according to the review mentioned above [40]) was from a paper published in 1996, and it defined social media—which they called computer-supported social networks—as "computer networks that link people as well as machines" [41]. By this definition, *Email* would be the first social media platform in 1971. However, the modern understanding of social media argues against this claim because *Email* has a distribution mechanism rather than a collective one, and it ensures communication back and forth rather than genuine connectivity [42].

A popular paper from 2007 [43] defines social media—which they call social network sites—as platforms that let users (1) construct public or semi-public profiles, (2) connect with a list of other users, and (3) navigate both their list of connections and other people's, thus making their social network visible. Using this definition, they found that *SixDegrees.com* was the first social media platform in 1997. This perception of social media was typical of the early 2000s but does not truly encapsulate current social media platforms and their purpose, which is more focused on the sharing and creation of content than friends or follower lists.

Modern understanding of social media is broader in terms of medium—as the number of social media platforms and features increases, the definition must also grow and develop to accommodate. It also adds user-generated content as one of the defining traits of social media. This has become the trend as of 2010, with the first paper being [44]. A more exhaustive attempt at a formal definition of modern social media was found in a paper from 2012 [45], which determined that,

> "social media consists of (a) the information infrastructure and tools used to produce and distribute content that has individual value but reflects shared values; (b) the content that takes the digital form of personal messages, news, ideas, that becomes cultural products; and (c) the people, organizations, and industries that produce and consume both the tools and the content."

This definition uses 'social media,' and the platforms described are more representative of the modern colloquial understanding of the term. This definition will be used for this review. And though a few papers offer their own definitions, they are all similar or adjacent to it, as is expected considering they were all published after 2010.

This modern kind of social media platform first became popular in 2003-2004 with *MySpace* and *Facebook* [44]. It has grown at a staggering rate since then. For example, the percentage of US adults using social media went up from 5% in 2005 to 79% in 2019 [46]. This number is set to increase further,

as social media is even more popular with the younger generation (90% of American young adults used social media in 2015, a percentage which is far higher than that of the general population (65%)) [47]. Thus, it is likely that the number of studies on the topic will also continue to grow. Since social media is not going anywhere, the need to study how it impacts various human phenomena—like radicalisation—becomes more relevant with each passing year.

2.2 Radicalisation

'Radicalisation' is a relatively new concept, but it has become the main framework for understanding micro-level transitions towards violence [48]. It first emerged after 2001—post 9/11—when the discourse regarding the root causes of terrorism that was prominent in academia since the 70s became distasteful—as it was deemed an attempt to justify the killing of innocents—and the term radicalisation was coined instead to describe 'what goes on before the bomb goes off' [49]. It rose to prominence in 2005-2007 with the 'home-grown' terrorist phenomenon in Western Europe, when programs that attempted to counter and prevent radicalisation began receiving significant funding [50], and has remained prominent in conversations around terrorism, extreme violence, and political extremism ever since.

For this review, 'what happens before the bombs go off' is not a sufficiently specific definition, so other more modern ones were explored. The Oxford English Dictionary provides a colloquial definition of the term radicalisation as: "the action or process of making or becoming radical, esp. in political outlook" [51], with radical being defined as, "advocating thorough or far-reaching political or social reform; representing or supporting an extreme section of a party" [52]. Interestingly, this definition does not mention violence as a requirement for radicalisation and acknowledges established political parties as a possible radical cause to support. It is also too broad for both academic and legislative purposes, as terms like 'far-reaching' are subjective and 'advocating' can mean both peaceful and legal protests and violent terrorism.

Various governments have provided their own definitions. For example, the US government defined (violent) radicalisation as

> "the process of adopting or promoting an extremist belief system for the purpose of facilitating ideologically based violence to advance political, religious, or social change" [53].

The UK home office said,

> "radicalisation means someone is being encouraged to develop extreme views or beliefs in support of terrorist groups and activities" [54].

And the Netherlands' AIVD defined radicalism as,

> "the (active) pursuit of and/or support to far-reaching changes in society which may constitute a danger to (the continued existence of) the democratic legal order (aim), which may involve the use of undemocratic methods (means) that may harm the functioning of the democratic legal order (effect)" [55].

The general message within these definitions is similar, but there are some significant differences between each which may make deciding what is included in radicalisation difficult. For example, 'terrorist groups and activities' do not—generally speaking—include riots, though they may be considered 'violence to advance political change' and a 'danger to the democratic legal order.' The UK definition implies an active agent that is encouraging radicalisation, while the US definition implies individual responsibility. Both the US and UK definitions imply that the development of radical ideas

is part of the radicalisation process, while the Dutch definition implies that radicalisation starts with active support. And 'Undemocratic methods' does not necessarily mean violent methods either.

Attempts to find a unique academic definition for radicalisation have also been futile, as the term is used in multiple different ways and contexts [50], and is overall ambiguous [56], or far too specific and thus only applicable to the context the specific definition was coined [57]. One especially relevant point of contention is the split between studying radicalisation with emphasis on extremist beliefs—known as cognitive radicalization—and studying it in terms of extremist behaviour—known as behavioural radicalization [56]. It is, however, far from the only point of contention, as some academics have gone so far as to argue that radicalisation does not exist, and rather that it is a myth propagated by governments and the media to legitimise policies and support the 'war on terror' [58]. While this is an acknowledged position, for the purposes of this dissertation, I will be expanding on some existing radicalisation models and definitions assuming radicalisation does exist.

One popular paper from 2012 [59] analysed the literature on radicalisation prior to its publication and found that the conceptualisation of radicalisation could be divided into 4 categories: a cultural-psychology disposition, a theological process, a theological-psychological process, and as policing tools. The first argues that the macro-context is irrelevant, there is no root cause but a predisposition that is either cultural or psychological or a combination of both. The second identified indicators specifically related to Islam that are supposedly present in people before they become radicalised. Both have become less popular and influential in recent years, as the focus shifted away from solely focusing on Jihadist radicalisation, and other models were developed with more factors on all contextual scales.

The third model of radicalisation proposed in 2012 [59] argued that it is not the religious beliefs by themselves that radicalise people, instead, the trigger is either a 'cognitive opening,' an identity crisis, or a group bonding process. This concept of a catalyst or cognitive opening is still influential in many studies. It has been defined as a mental aptitude where someone is able to embrace a new ideological system, and it has been split into categories to better identify its scope. 4 types of triggers were found: (1) economic difficulty like recent unemployment, (2) political motivators like wars, (3) social factors like real and perceived alienation or discrimination, and (4) personal loss [60].

The final model mentioned in 2012 [59] was that of radicalisation as policing tools which is named as such because it was spearheaded by governments. It was mostly an attempt to systematise and build upon the previously mentioned models in a way that would be useful for policy-making and de-radicalisation efforts. This model defined radicalisation as a four-step process: (1) pre-radicalisation, (2) self-identification—the exploration of 'Salafi Islam' as a result of a cognitive opening and a breakdown of one's current identity— (3) indoctrination, where—through group association—the beliefs become more and more radical until the ideology is fully adopted, and (4) jihadization, where the radical individual embraces the need for Jihad. This also has a sole focus on Islam that makes it unusable as a modern model for all radicalisation, but it has impacted future theories.

An example of how this was expanded beyond Islamic radicalisation can be found in a proposed model from 2016 [61]. It analysed multiple forms of radicalisation, which they separate as: far-right, far-left, separatist groups, single issues, and religious extremists. This model proposes a three-step process: sensitivity, group membership, and action, which correspond roughly to steps 2-4 of the above. The sensitivity step includes at the micro-level, feelings of uncertainty, lack of belonging and insignificance; at a meso-level, social factors such as already radical loved ones or perceived discrimination against one's social group; and at a macro-level, the wider societal context, global conflicts, and globalisation.

The second step occurs when individuals join a radical group and follow their norms and values to ingrate themselves, go through initiation rituals and training, and cut ties with those outside of the group, virtually fusing the individual with the radical group. The action phase is when the individual turns to violence, a step which may be aided by personal tragedy, dehumanisation of the enemy, and the use of authority and manipulation to make the violence easier to enact.

Another interesting radicalisation framework is offered by one of the most influential books on the topic, titled: *Friction: How Radicalization Happens to Them and Us* [62], which found that radicalisation was a complicated phenomenon—with 12 main mechanisms—that operated on multiple scales [60]. They divided these mechanisms into individual, group, and mass radicalisation. Those identified for individual radicalisation were: (1) personal grievances, usually perceived injustices against either the radicalised individual or their loved ones; (2) group grievance, where actors like "lone wolves" identified with a group because of shared political and personal ideas; (3) slippery slopes, manifesting as repeated engagement in small actions that increase commitment; (4) affective relationships with already radical individuals; (5) the evolutionary perception that engaging in risky behaviour increases status; and (6) unfreezing which is the external removal of social and personal barriers (like family). Group radicalisation had the mechanisms of (7) group polarization—where political discussion in radical groups leads to an increased agreement about radical opinions— (8) group competition and the 'us vs them mentality,' and (9) group isolation, where actions are performed according to group consensus rather than personal opinions, and members are coerced to comply. Lastly, mass radicalisation included (10) jujitsu politics, where terrorists purposely incite violent counter-responses to generate sympathy; (11) hatred for the opposition and viewing them as inherently evil, and (12) the ideal of martyrdom as an incentive to sacrifice among group members [63].

The final theory of radicalisation I want to explore here is one of the most recent popular models: The Motivational Imbalance Theory, first coined in 2019 in *The Three Pillars of Radicalization* [64]. It is a psychological theory that hinges on the notion that individuals have basic needs and that when one of these needs becomes dominant, a motivational imbalance can occur. This imbalance leads to individuals taking more radical actions to satisfy the dominant need and may eventually result in extremism—which may manifest in multiple forms one of which is violent extremism. Terrorists and other radical groups play a vital role in inducing motivational imbalance. They do this through the narratives spread in their propaganda, which appeal to values individuals hold sacred and thus trigger the need for respect, and through the social networks they provide, which allow people to embrace radical ideologies with like-minded individuals and eventually collaborate in extremist behaviour. Many of the triggers for the 'quest of significance' are common causes of radical extremism identified in other models—like feeling personally rejected by society, perceived discrimination, or the hero narrative of those engaging in radical violence—but what makes this model unique is the more complex psychological backing, and its main focus being on the ideological and cognitive aspects of radicalisation [60].

There are many more frameworks for radicalisation and probably more to come as there is still significant disagreement on the topic, as evidence from the sheer variety within the few models explored here. However, attempting to find a common definition for radicalisation is beyond the scope of this dissertation. Thus, to accommodate for the vast differences in understanding of the term, I decided not to pick or create a specific definition for the review, instead including multiple types of radicalisation.

This ended up including the typical behavioural radicalisation that leads to extremism, radicalisation that focused solely on the cognitive or ideological aspect, radicalisation that resulted in hate speech, and even extremism which operated on a governmental level. The things I deemed necessary to classify

the papers as about radicalisation—other than the search term being present—were relative extremism in some form and the acknowledgement of a process that involved human actors being radicalised.

Chapter 3: Methodology

3.1 Research Method: Semi-Systematic Literature Review

Literature reviews collect and synthesise academic research on a topic or discipline [65]. They are useful to provide a wider overview of a specific issue than a single study could, because they can draw from multiple trustworthy sources with different methodologies and scopes.

Historically, they were most common in healthcare, and they were used because the body of research became too large to succinctly summarise within the literature review section of other papers. However, until the 1980s they lacked a formal definition or framework, which lead to significant bias and loss of information [66]. This changed as the volume of literature grew further and literature reviews became popular in other sciences [67], making the need for reliable and replicable reviews increasingly urgent.

Different frameworks, techniques and categorisations were thus created to limit bias and offer consistent structures for authors. Three main categories were proposed by one of the most cited papers on literature reviews themselves [65]: Systematic, semi-systematic and integrative reviews.

Integrative reviews—also known as critical reviews—analyse and research a topic as a means to provide and explore new theoretical frameworks [65]. They require expertise in the entire topic of study, to ensure that any created hypothesis is illustrative of the entirety of the research it is trying to represent [68]. This cannot be feasibly achieved in this dissertation.

Systematic literature reviews are the most rigorous ones. They aim to find all available empirical evidence to answer a precise research question through systematic and explicitly stated methods [65]. They typically require multiple reviewers and are conducted over an extended time frame—sometimes spanning years— [69], which is beyond the scope and timeframe of this dissertation. Additionally, my research question and topic are both multidisciplinary and broad in nature, which would be problematic for strictly defined metanalysis, as different disciplines and methodologies require different methods of analysis [67]. Thus, a systematic literature review is not suitable for this project.

The chosen style of literature review is the semi-systematic—or narrative—approach. Unlike systematic reviews, these reviews are most useful for topics that span different disciplines. A semi-systematic literature review aims to provide an overview of the research on a topic where there are multiple research methods and traditions, exploring the meta-narrative in a way that is still explicitly explained—as to enable the reader to assess for themselves the reasonability of the conclusions—but less numerical and rigorous, to allow for the complexity of the topic and a broader research question [65].

3.2 Sample

3.2.1 Exclusion and Inclusion Criteria

The first step in selecting a sample for the review is to determine the key words and the databases. Because I wanted to explore the overlap between political radicalisation and social media, the key words selected were *"social media"* in quotes to ensure it considers only instances where the words are together, *political,* and *radicali?ation* to include both the British and American English spellings of the word. Then to ensure that the articles given by the search focused on the process of radicalisation and its causes within social media rather than proposed solutions, the keywords *prevent*, detect*, and reduc** were excluded from the search. Thus, the final search term inputted in the databases was:

(TITLE-ABS-KEY ("social media" political AND radicali?ation) AND NOT TITLE-ABS-KEY (prevent OR detect* OR reduc*)*

Initially, I attempted to search Google Scholar but that yielded tens of thousands of results, which is beyond what can be feasibly read in the timeframe of this dissertation. So, I researched other academic databases to use and eventually chose Scopus and Web of Science, since they are well-known databases that cover a wide range of social sciences.

I did not filter by year or article type, but I did restrict the results to only English papers. I ran the search on both databases on the 9th of December 2022, and it returned 118 total results to begin the screening process.

3.2.2 Literature Screening Process

The initial search of the databases yielded 56 papers on Scopus and 62 papers on Web of Science, giving a total of 118 articles. These were exported into EndNote to organise and sort through. Endnote's automatic duplicate finder found 13 duplicate articles. Then, I manually scanned the titles, abstracts and authors, and if there was at least a 90% match, I deleted one of the articles as a duplicate. I did this to account for possible publication of the same studies in different journals and errors or bugs in the exporting of the information into Endnote that made the texts not identical and thus not detectible automatically. This, combined with the removal of any articles that were inaccessible both through the internet and the university library left 89 papers.

Once I had a list of unique articles, I read all abstracts and excluded those that were clearly irrelevant to the research question, a process that resulted in 47 articles. From there, I read each paper and annotated them systematically, excluding any paper that didn't help answer any part of my research question or aims. An example of an excluded paper at this stage of the process would be one that only mentioned social media is passing without acknowledging any social media specific features [70]. The literature screening process (graphically summarised in *Figure 3.1*) ended with a total of 30 eligible and relevant papers.

Since they covered a range of disciplines and research methods and offered insights into all aspects I wanted to investigate, there was no need to supplement the sample with any additional papers, so the review was conducted on these 30 articles which will from now on also be referred to as the sample of this review (a full list of which can be seen in Appendix A). When referenced these papers will be in blue, to distinguish papers that function as sample points for the literature review from normal references.

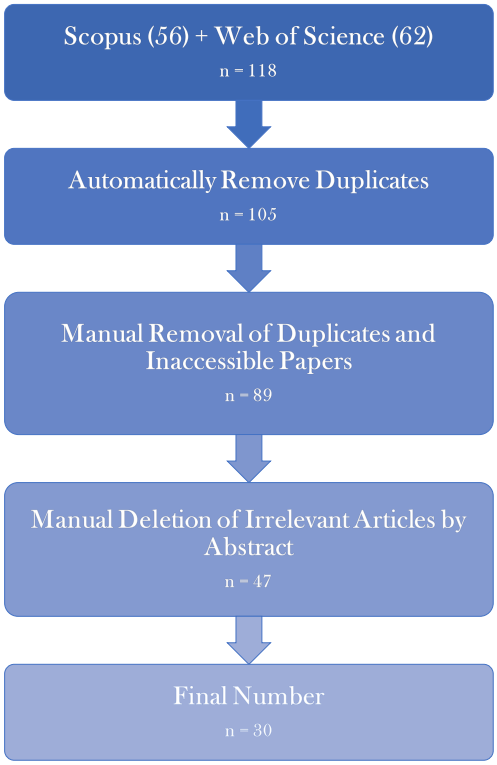

Figure 3.1: Flow Chart of the Literature Screening Process

3.3 Analysis Categories

To analyse the papers in a way that addresses the aims and research question I created different analytical categories and systematically took notes in an Excel sheet. These were either made numerical—for quantitative analysis—or collected to discuss critically in the second half of the results, depending on the type of information collected.

The quantitative aspect of the analysis includes the basic metadata and demographics of the papers. I tracked the publishing year, discipline, and geographic location (both that of the authors and the one discussed within the paper). I also tracked the type of data analysed, the type of research the paper conducted, how often keywords were used, and the analysis techniques mentioned by name. This was done to better understand the sample scope and communicate the methods that yielded the results I will use for the later—more narrative—aspects of the review.

Some of this collected data required a muti-stepped approach, to account for ambiguity.

To determine the disciplines of each paper, I first checked whether the paper itself claimed to belong to a specific discipline, then whether the journal it was published in specified and, if they did, I assigned accordingly. However, most papers did not specify, and most journals were multidisciplinary. In that case, I used the author's public academic profiles, prioritising their university ones but using others whenever necessary. If multiple disciplines were listed the ones most pertinent to the paper were recorded. Sometimes this meant that more than one discipline was listed for a paper. Whenever there was more than one author, I selected the disciplines that had majority overlap or—if no overlap was found—I chose from each author individually like I did for papers with a single author.

For the geographic locations, whenever there was more than one location, all were recorded, but each country was only added once per paper and tracking category. For the studied location there was an additional level of complexity, since they were only added when the location itself was relevant to the paper's content. For example, in the study of the *Tumblr* blog belonging to a woman who joined ISIS [24], despite there only being one case study where the location could be easily determined (she was originally in Scotland and travelled to Syria), because the study focused exclusively on the blog and explicitly stated that her radicalisation occurred online, neither the UK nor Syria were recorded. This choice ensures that the geographical locations represented are actually relevant to the research topic.

Traditionally, papers are divided into quantitative and qualitative research, but I wanted to represent some specific trends I noticed within the sample, so the qualitative studies were broken down into more categories. The final list of types chosen to distinguish and represent the entire sample are: quantitative research, case study, exploratory paper, explanatory paper, literature review, and 'other qualitative research.' Quantitative papers were research papers that collected multiple data points and coded them numerically, usually through a number scale in surveys or a quantity of posts. Case studies were papers that focused few specific data points (between 1 and 3) and used them to make a bigger point about the topic. Papers that investigated a new aspect of radicalisation, a new social media platform, or a new feature, were called exploratory. Research that observed a phenomenon and applied theory to it in an attempt to explain it was labelled explanatory. Literature reviews were papers that used other academic papers as data. The 'other qualitative research' included category included interviews and other types of qualitative studies that did not fit in any of the above categorisations.

When discussing the types of analysis, I only recorded those mentioned explicitly. However, to account for slightly different techniques with the same kind of results and different naming practices, I grouped them by common analysis kind. For example, binary regression and logistic regression were both categorised as correlation analysis, and sentiment analysis was counted as natural language processing.

For the more narrative aspect of the review, I had a specific column for any information regarding affected demographics, identified contradictions, recurring themes, and social media affordances, where I noted any claim made in the papers to fully address everything that relates to the research aims.

Chapter 4: Results and Discussion

4.1 Sample Demographics

This section will explore the basic demographic and publishing details of the papers in this review, which are important to understand the size and relevance of the topic as well as the origin of the conclusions and statements that will be made further into the review. In this section the data tracked includes: the publication year, the discipline, the geographical location analysed in the paper, and the geographical location of the author.

4.1.1 Publication Year

Figure 4.1: Distribution of Papers by Their Publication Year

Despite not filtering by year, all the papers were within the last decade, which makes them more likely to still be relevant to the current social media landscape and emphasises the relatively new nature of the topic. When looking at the bar chart, there is also a noticeable upward trend, with 2018 and 2022 being exceptions. There is not enough data to establish that the trend is representative of the entirety of academia on the topic, but it does suggest that the topic is still relevant and being explored.

The reasons for these exceptions are unknown. The smaller number of papers in 2022 may be explained by the sampling being taken in early December, meaning an entire month of articles would not have been published yet, and that some published papers in the later months of the year might not have been added into the databases yet, thus making them not part of my sample.

I speculate that 2018 comprises nearly a third of the total sample partially due to a combination of the Cambridge Analytica scandal and Trump's presidency. In March 2018 it was revealed that the data collected by Cambridge Analytica was allegedly used to influence the 2016 US elections [71], bringing the issues of social media influencing the political sphere and far-right radicalisation into the mainstream—and thus possibly to researchers' attention—for the second half of 2018. Additionally, Trump's presidential campaign in 2016 saw an insurgence of alt-right content, including online spaces like the subreddit *r/the_Donald* which legitimised the far-right discourse and eventually bled into real-world violence in August 2017 with the Charlottesville March [72]. This shift from online extremism to offline violence made the need for research on the topic more evident—the few months' delay is reasonable considering the time taken to conduct and publish research.

4.1.2 Disciplines

I found 22 different disciplines within my sample. This emphasises the multidisciplinary nature of the topic analysed, even though some disciplines clearly share similarities—like political science and political studies. The results are summarised below.

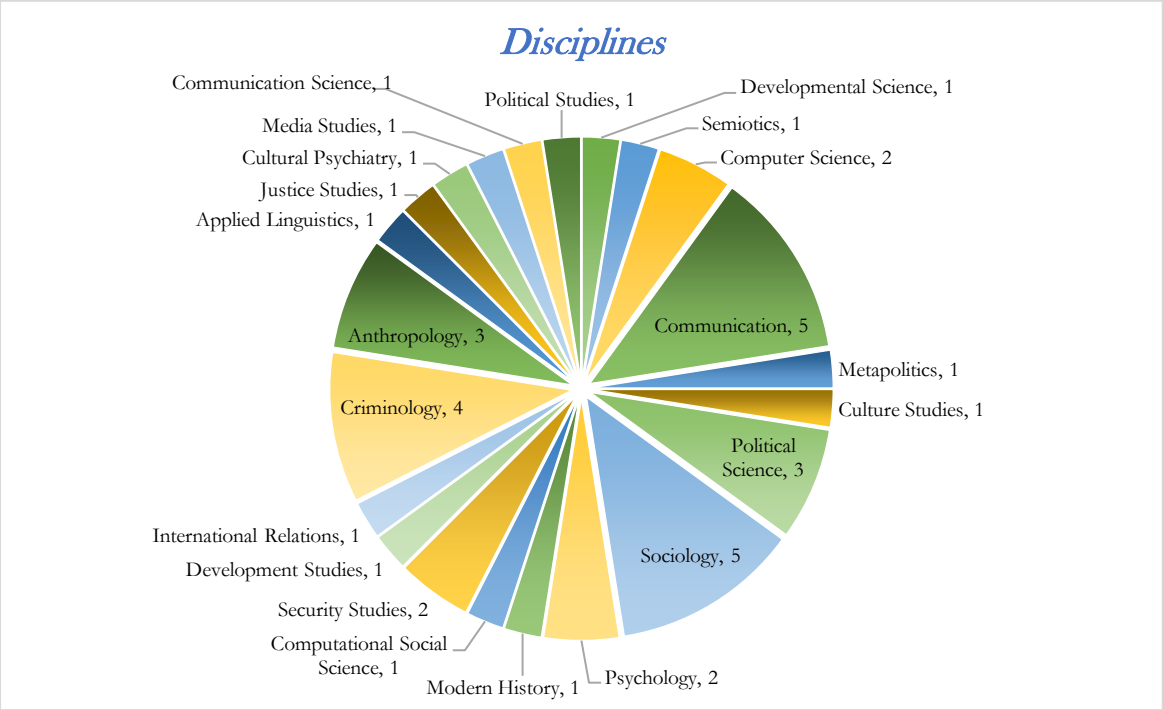

Figure 4.2: Disciplines Pie Chart

4.1.3 Geographical Locations

Two different geographical locations were tracked by country: the author's university and what area the study itself examined. The universities involved in the sample were located in 19 different countries, and 19 unique countries were studied by the papers, but not all authors wrote about their own countries, so the overlap isn't exact. In fact, there were 6 papers which had no authors domestic to the specific radicalisation context explored. 3 papers [1][27][7] were international collaborations, with two of them studying radicalisation without a geographical context and the third analysing protests in the United States and the United Kingdom, while having authors from the UK and the Netherlands.

Because I only tracked studied areas if they were relevant to the radicalisation process in some way, 6 of the papers did not have any relevant geographical location and were thus excluded from the 'area studied' data points. Additionally, Syria did not have any data points despite being mentioned in multiple papers, as it was never where the radicalisation itself occurred.

The United States were the most represented country, both by authors, with 9 papers covering almost 1/3 of the sample size, and by examined location, with 5 studies looking at radicalisation in the US. I speculate that this is partially because the US has the biggest population of English speakers [73], and because of how mainstream the far-right became in the United States following the 2016 elections. It was followed by Germany with 2 universities and 4 studies examining German radicalisation.

The results for each country are summarised in *Table 3.1*.

Country	Authors	Studied Location	Studies From Authors About Their Own Country	Studies from Foreign Authors
Australia	2	1	1	-
Bangladesh	-	1	-	1
Belgium	1	1	1	-
Brazil	1	1	1	-
Canada	2	-	-	-
China	1	2	1	1
Croatia	1	-	-	-
Estonia	1	-	-	-
France	2	1	-	1
Germany	3	4	2	2
Hungary	1	-	-	-
India	1	1	1	-
Indonesia	1	2	1	1
Ireland	-	1	-	1
Italy	1	1	1	-
Japan	-	1	-	1
Kenya	1	1	1	-
Netherlands	2	-	-	-
Norway	1	1	1	-
Philippines	-	1	-	1
Poland	1	1	1	-
South Korea	-	1	-	1
United Kingdom	2	2	2	-
United States	9	5	4	1

Table 4.1: Geographical Location of the Studies

The final key consideration about the geographical distribution of the studied areas is how they compare to the distribution of expressions of radicalisation, like terrorism. To compare these values, I use Global Terrorism Index [74], because it has multiple indicators (incidents, fatalities, injuries, and hostages) and can thus account for multiple dimensions of extremism, but it is not the only measure of radicalisation or terrorism. It is merely a useful tool for comparison.

When comparing these values to the some clearly under-represented areas appear. The United States and Germany, rank 30th and 35th respectively despite being the most represented in the sample, while the top 12 countries in the index do not appear in the sample once.

This is not to say that authors should stop studying either country. With so many authors being American it is natural for them to be most familiar and interested with the US's context. Terrorism is only one example of radicalisation, and terrorism occurring in a specific place doesn't necessarily imply that the radicalisation occurs there—as evident from Syria being a hotspot for radicalised immigration. Other factors may also impact which countries are most studied in papers with the keywords chosen in this dissertation—the most obvious one being how much social media is used within any given country. However, the dissonance between areas studied and areas affected by terrorism may be worth investigating. It is most easily viewed by comparing the heat maps.

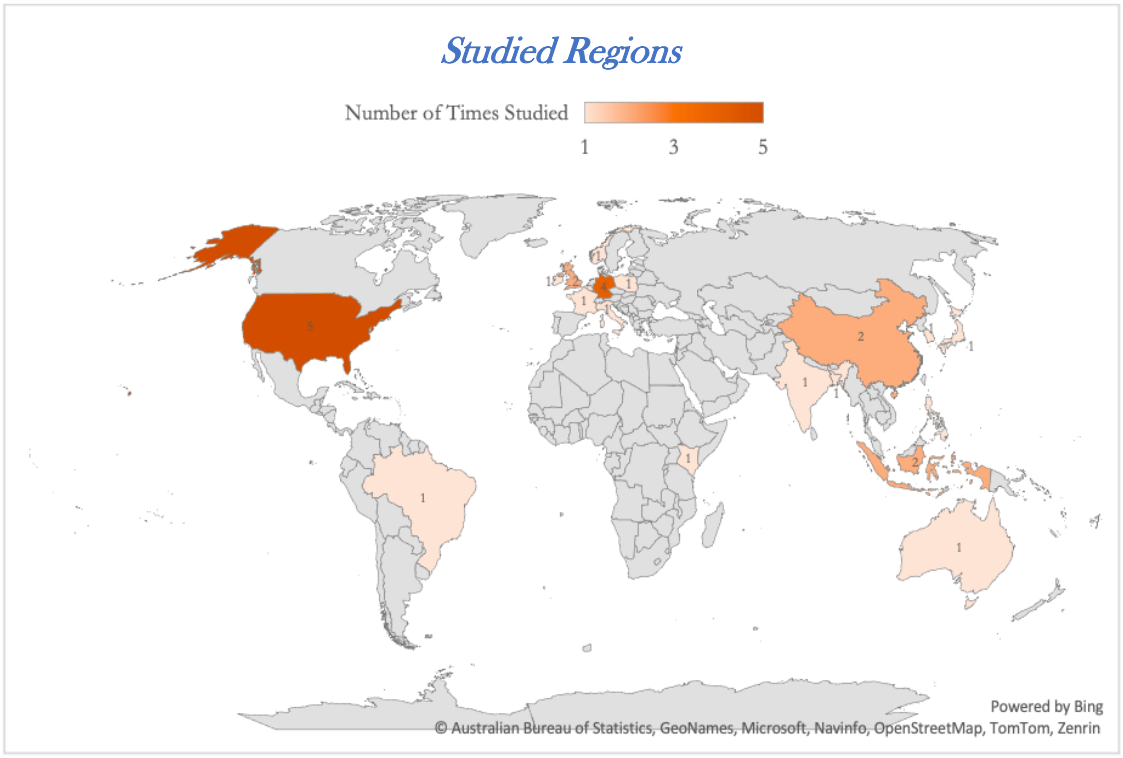

Figure 4.3: Geographical Location Studied Heat Map

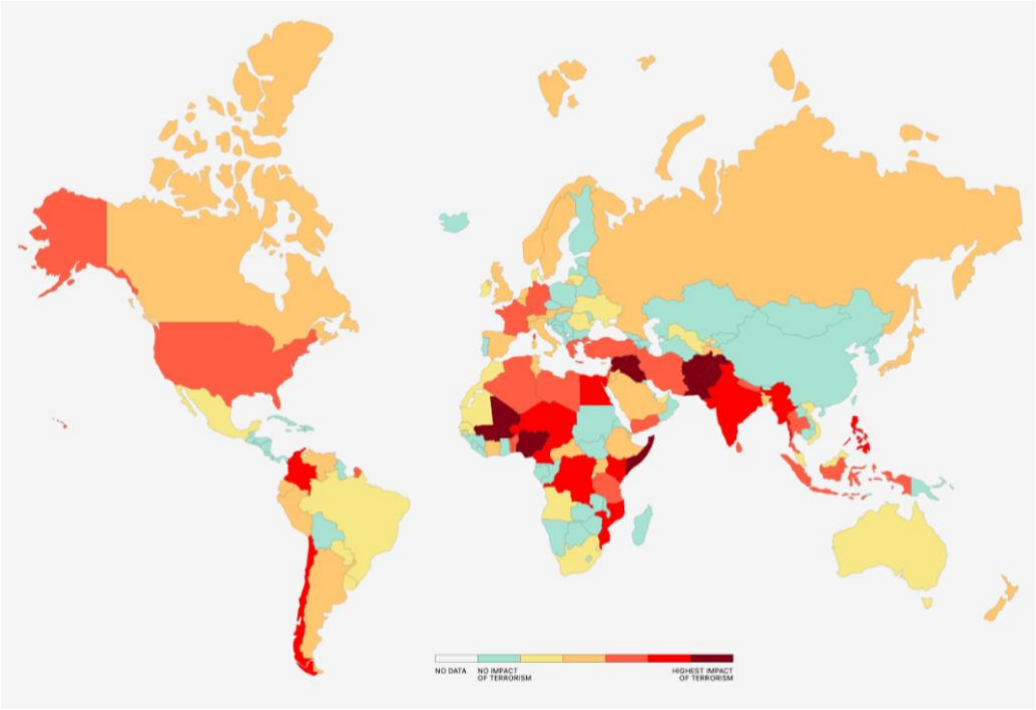

Figure 4.4: Global Terrorist Index Heat Map [74]

Even when removing countries that have no significant social media usage, a geographical research gap is still visible. This has been a trend observed in more general literature on radicalisation too and has been mentioned in multiple papers within my sample, with one paper [13] emphasising how "the decoupling of 'research from the US context'" is a valuable contribution to the field, and another [15] stating that "social media's polarizing potential has been predominantly studied in western societies

and focused on Euro-American-centric platforms." Expanding beyond this context is important to ensure that conclusions drawn are generalisable to other cultures too.

Using the heat maps, North Africa and South Asia are the most underexplored areas with affected by radicalisation. A significant portion of the population of North Africa (56% [75]) uses social media so it is a natural area to consider expanding the research into. When considering the index country by country, the country with the highest Global Terrorism Index is Afghanistan, which has not been mentioned in any of the papers, but only about 23% [76] of the population uses the internet, which is a significantly lower portion when compared the more analysed West, so social media usage might not be a factor impacting this terrorism. However, within the top ten countries in the Index, Iraq [77] and Myanmar [78] have significant social media usage and may also be worth investigating as a result.

4.2 Sample Methods and Types

All data types and analysis methods have their own qualities, contexts, and limitations. One of the main advantages of literature reviews over other studies is that by combining many papers with their own methods and data, they can compensate for weaknesses in each other's methods. For example, surveys often suffer from desirability bias, which may make people appear less radical than they are, but studies that focus on analysing fights on social media are likely to study views that appear more radical, because of the nature of emotional conflicts and troll accounts. If both studies agree on something, they form a sort of upper and lower bound that makes the shared conclusion more reliable. This is at its most effective when the variety of study and analysis is wide, which is why knowing the study types, data types and analysis types of the studies in this review—and how they are distributed—is important.

4.2.1 Types of Study

When reading the papers I noticed some clear trends in approach within the sample, which I recorded by assigning each paper to a *type of study*. The types chosen to distinguish and represent the entire sample are: quantitative, case study, exploratory, explanatory, literature review, and 'other qualitative research.'

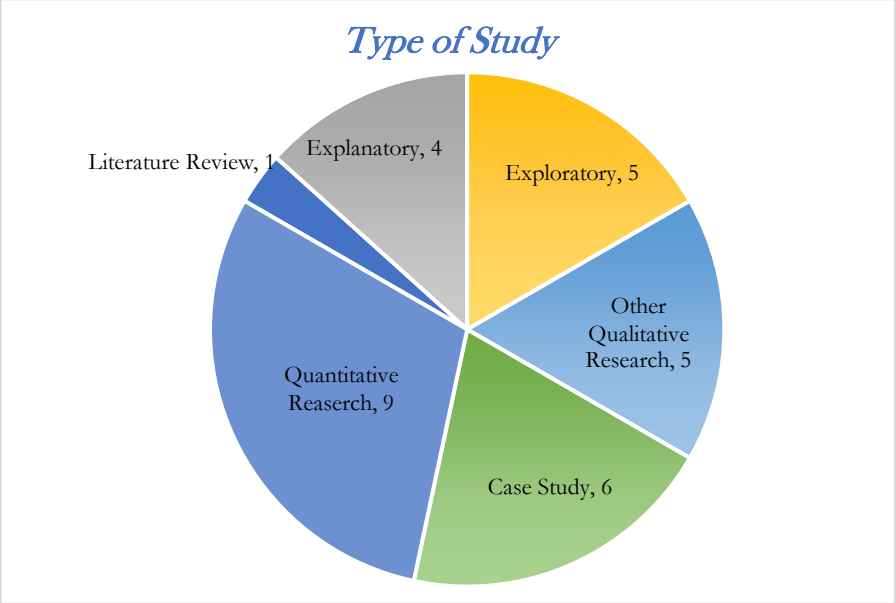

Figure 4.5: Type of Study Pie Chart

4.2.2 Types of Data and Data Sources

Data sources are traditionally divided in Primary and Secondary. However, to understand exactly what type of limitations and features each source has the exact kind is also important, so I recorded both the source type and what the sources actually where. For any study that used more than one source I recorded multiple. The results are represented in *Figures 3.6* and *3.7*.

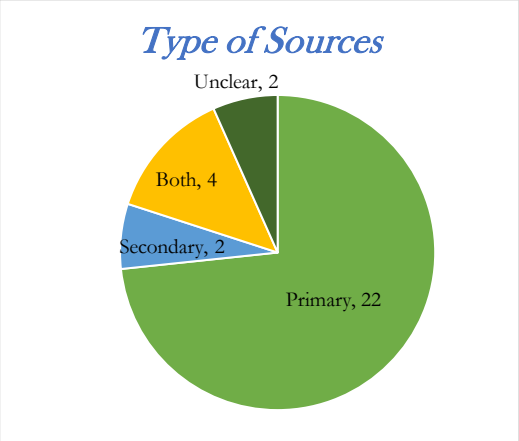

Figure 4.6: Type Sources Pie Chart

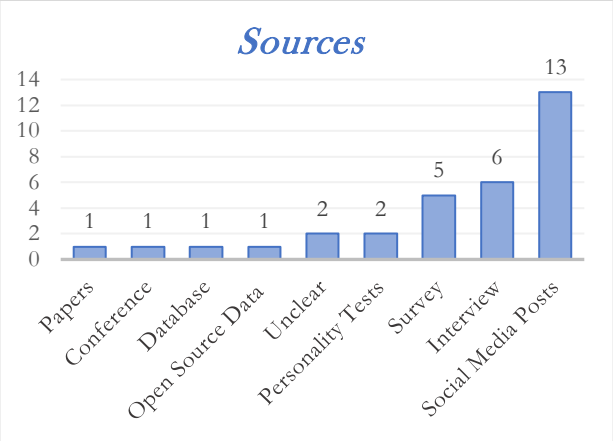

Figure 4.7: Sources Used

4.2.2 Type of Analysis

I recorded any analysis method the papers mentioned explicitly, grouping them whenever possible. The results are summarised in the bar chart below.

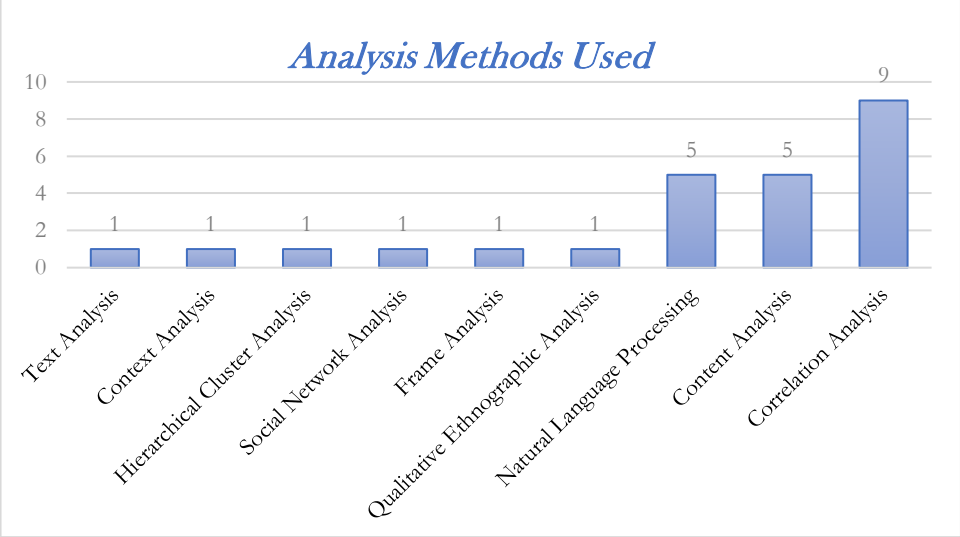

Figure 4.8: Analysis Methods Used Within Sample

4.3 Sample Content

4.3.1 Radicalisation Perpetrators

There was a clear grouping of the perpetrators of radicalisation explored, with 25 of the 30 papers discussing either Islamic radicalisation or far-right radicalisation. Only 2 papers had a different topic, a paper on the religious divide between children in India [28] and one on the war for public opinion over the Syrian conflict [29].

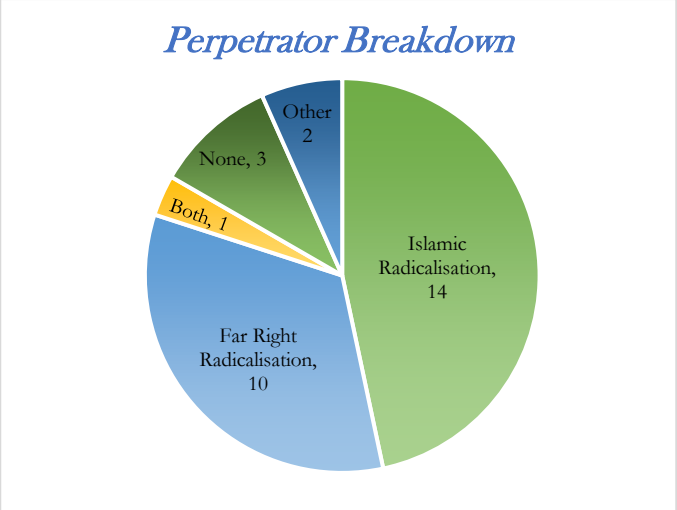

Figure 4.9: Perpetrators of Radicalisation

Islamic radicalisation is studied in 50% of the papers in the sample and 'Islam' is the most used keyword other than the search terms (for a comprehensive list of all keywords see Appendix B). This is probably because the research is Euro-American-centric—as noted in the geographical section—and for the last decade the deadliest form of terrorism in the West has been Islamic terrorism [79], a number which is exacerbated by the over-representation of terrorist attacks perpetrated by Muslims in the media [80]. However, this may also be a symptom of the link that has been created in the collective mind between Islam and terrorism and an effort could be made to ensure research does not equate all radicalisation with Islamic radicalisation by adding more topics.

The other major perpetrator is far-right radicalisation, which also has some statistical justification, with 31% of fatalities in the West since 2007 being a consequence of far-right ideological terrorism [79], and the far right outpacing all other perpetrators of terror attacks in recent years [81]. Additionally, the papers themselves span multiple types of right-wing radicalisation beyond terrorists, like the elected party of Brazil's 2018 elections [14].

Interestingly, despite causing significantly fewer deaths, far-left terrorist attacks are more common—though it is important to note that this categorisation is becoming outdated as more groups stop identifying with a specific ideology [79]. Overall, radicalisation on the opposite end of the political spectrum may be an interesting direction for future research, especially considering how both the far right and radical Islam tend to favour fundamentalist conservative rhetoric [30][13] and bigotry [14][5] that have become somewhat synonymous with radical discourse but have not been assessed as thoroughly against radicalisation on the traditionally progressive end of the spectrum. That being said, one paper [82] was generated from my keyword search about the traditional radical left and was excluded because it established that the movement is not interested in radicalising others and did not have a significant social media presence, so research would need to find new movements.

4.3.2 Social Media Platforms Studied

I recorded both all social media platforms used to source data for the studies and all social media platforms mentioned as examples or in passing, a summary of which can be seen in *Figures 3.10* and *3.11* respectively.

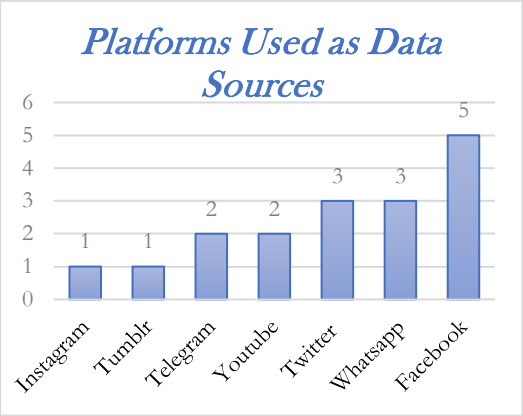

Figure 4.10: Social Media Platforms That Were Explicitly Used as Data Sources

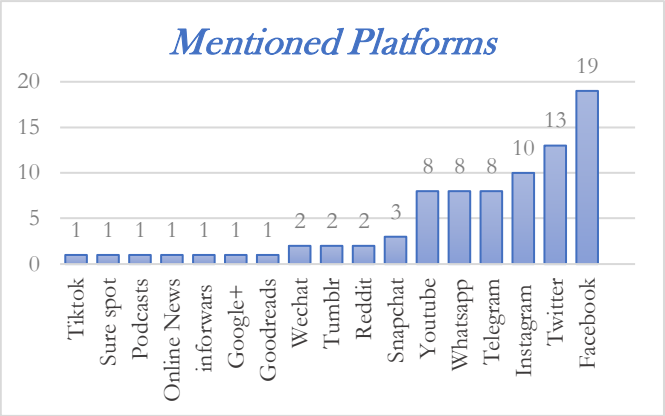

Figure 4.11: Social Media Platforms Mentioned

The most mentioned platform in both categories is Facebook, which is predictable considering it has been the lead social media platform for over a decade [46]. Telegram is the most over-represented platform when considering global usage, but it has become the app of choice for terrorists [6] as the mainstream ones like Facebook, YouTube, and Twitter began monitoring and removing radical content more aggressively [83]. If anything, considering that IS itself favours Telegram [83] and that this shift is still somewhat recent, the platform is still under-researched. According to an exploratory study on the matter [6], this may be because ethical study practices dictate that researchers do not share or otherwise engage with the extremist content they study, making investigating the groups difficult both from an academic perspective—due to the need of research approval—and from a practical one—since groups tend to kick out and blacklist accounts that do not interact.

Another platform that may warrant further research, especially when discussing the more cognitive aspect and early stages of radicalisation, is TikTok, which has only been mentioned once within the entire sample. This is because TikTok primarily targets the younger generation [84], and multiple studies have determined that younger people are more likely to be affected by radicalisation (see *Section 4.4.1* for more details). Beyond that, TikTok is growing quickly [85], meaning it is likely to become even more relevant soon, and has been known to successfully popularise and spread misinformation already [86], making it feasible for conspiracy theories and fake news with a specific radicalising agenda to also go viral on the platform.

4.3.3 Most Radicalising Features of Social Media Platforms

Part of understanding how social media impacts political radicalisation is understanding which features are used to radicalised others and as means to self-radicalise. So, I extrapolated the features from all the methods of radicalisation mentioned in studies and the ones named explicitly and recorded them.

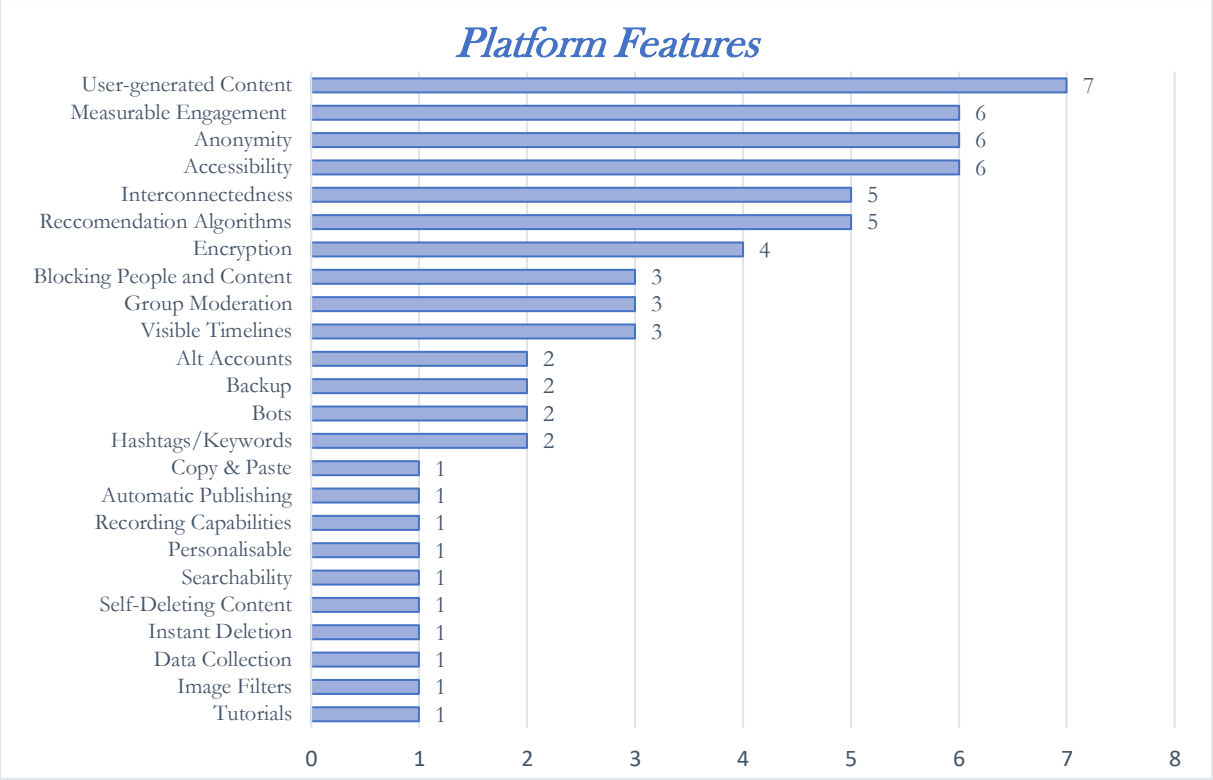

Figure 4.12: Mentioned Platform Features Used for Radicalisation

This is a nuanced topic, with each feature being usable in many different ways, and thus necessitates a more thorough explanation of how some of the most mentioned features were found to aid radicalisation.

The user-generated nature of social media is the most mentioned feature. It fosters a sense of community and belonging within radical groups, which appeals to vulnerable and isolated individuals [16]. This content also appears more trustworthy than produced or scripted content would, leading to distrust in mainstream media whenever what is reported is different from what friendly users say [26]. User-generated content is easier to maintain for radical groups because it gives them control over the narrative presented on their pages [2], and sympathisers repost it—and create adjacent content—providing free visibility [18]. This ties into the interconnected aspect of social media—the network-like social connections determining who follows and interacts with whom—which connects sympathisers with extremists and recruiters, to create a leaderless system that enables radical behaviour, facilitates mobilisation and is harder to stop [9].

Measurable engagement refers to platform features like comments, likes, and shares that are easy for the audience to use and give a numerical value of how popular a post is. Since more extreme ideas usually yield higher engagement, what is expressed can at times be more radical than the individual's actual opinions, attracting even more radical people to their account and initiating a feedback loop. This cycle and the quantifiable approval may also become somewhat addictive, urging individuals to express even more radical ideas to obtain the desired response [19]. The engagement itself is also easy to perform—likes are a single click—this makes it a lower commitment than traditional conversation, so somewhat sympathetic or simply curious people are more likely to engage with radical content using these features. This raises the engagement values of propaganda and thus gives it an algorithmic boost [24]. Engagement can be further inflated with bots, buzzers, copy and pasted comments or like-farming

[30] simulating genuine engagement for both the algorithms and regular people, who upon seeing extreme content with thousands of likes and supportive comments may think the views expressed more mainstream than they are.

Many social media platforms do not require identifying information in the account creation process. This grants users a level of anonymity that lets them express their most violent and extreme opinions without the fear of getting recognised by peers [28], leading to hate speech and other similar radical behaviour. Being exposed to such inflammatory language makes for 'negative contact' between opposing factions, which in turn makes both sides more radical in their dislike for the opposition [7]. Anonymity also makes users harder to track, which makes those who share radical propaganda and misinformation—or use social media to mobilise—harder to find and stop [13]. This is even more effective when combined with end-to-end encryption, which makes the conversations themselves hard to find as well, which is one of the main reasons Telegram has become so popular among radical groups [6]. Anonymity is also essential for the creation of fake accounts which can be used to troll and inflate engagement as mentioned above.

I refer to the collection of elements that make social media more accessible as accessibility. The first of these elements is cost accessibility: social media offers a cheap means to communicate [1] and sometimes access is even offered for free, like in Brazil where the internet is often too expensive but phone companies offer free WhatsApp use [13] making people more likely to spend time on any given platform and less likely to spend money on other sources of information which may be less radical. Then there is geographical accessibility, where social media has the advantage over physical interactions because users can communicate between countries [7][18], which is especially relevant considering the issue of radicalised immigration and foreign fighter recruitment. Lastly, there is time accessibility. Social media provides instantaneous communication, which is invaluable when mobilising groups [4], and it also grants curious people 24/7 access to propaganda and groups of like-minded individuals [11]. Combined with self-deleting content or content that is often flagged and taken down, this 24/7 access can become all-encompassing and addictive [9], driving curious people further down the radicalisation process.

Other elements that are less mentioned are still used effectively. The ability to block and moderate groups allows for 'safe spaces' where people can express radical opinions and share propaganda and misinformation without meeting opposition, deepening the 'us' vs 'them' mentalities [3][28]. Bots and backups make it easy for taken down propaganda to be re-uploaded [6] and to flood people's timelines with misinformation and biased content [29].

4.3.4 Types of Radical Content Identified

There are multiple kinds of online content that contribute to radicalisation. The papers in my sample identified: conspiracy theories, hate speech, dog whistles, and extremist propaganda.

A trend of framing political events as conspiracy theories arose in recent years [21]. It led to the misrepresentation of political debates and incited paranoia [19] in already radical groups. One paper [22] studied German far-right narratives surrounding immigration and found multiple conspiracy theorists fearmongering about the UN global migration pact and the 'great replacement' of white people with immigrants, while also spreading narratives of mainstream media corruption to discredit legitimate media sources that disagree with them. Beyond supporting hateful narratives, conspiracy theories are dangerous because their believers are more likely to accept violence than sceptics are [14]. #Pizzagate

is a famous example of the possible repercussions, as narratives of supposed paedophile rings frequented by celebrities and left-wing politicians led to domestic terrorism, as the culprit when exposed to the theory believed it and shot up a restaurant in an attempt to free these fictional children [19].

The use of hate speech is positively correlated with how radical a space is [3]. Hate speech refers to social media posts making racist, sexist, or otherwise bigoted statements which play on pre-existing biases and are usually accompanied by a call to violent action. This speech is created to appeal to a conservative audience and has been shown to be effective even when done in an official capacity like political campaigns [13]. This is because the appeal hinges on strong emotions, like hatred and fear which have been acknowledged by multiple studies to be one of the strongest mobilising factors [7][11][27][28]. While engaging in hate speech, radical individuals effectively dehumanise [2][27], others [30] or emasculate the enemy [30][19], which serves to both make them less intimidating in a possible armed conflict and to justify possibly deadly actions.

Dog whistle in this context refers to content that appears politically neutral or moderate but has radical subtext. This area of research is mentioned but understudied, probably because its reliance on subtext makes an objective assessment of whether a piece of media is a dog whistle impossible. However, people are oftentimes exposed to radical content because they unsuspectingly clicked on something else [10], and one of the reviewed papers [5] studied this phenomenon exclusively, with a digital ethnography of the social media influencer Brittany Pettibone—whose content appears to be mostly apolitical with marriage and family-based narratives, but taps into radical conservative meta-politics through emphasis on traditional families and subtle nods at white supremacy—so I found it important to mention.

The final category of content is the broadest and most nuanced, extremist propaganda. This is radical by definition, but if the propaganda is stumbled upon without research or engagement it does not have a significant effect on radicalisation [26], with the most effect found being some desensitisation to violence when the propaganda is violent itself [1][6][20]. However, active engagement with the propaganda was found to be a major radicalising factor [20][23], with multiple studies [5][12][17] finding it to be a catalyst for the initial steps of self-radicalisation and even at times serving as the way for recruiters to find those sympathetic to their cause [4]. It was also found that subtle propaganda is more likely to be effective [14] with the more dramatic and violent one typically associated with terrorism actually being considered unappealing by youth [26]. Interestingly, exposure to propaganda from the opposite side of an ideological conflict can also serve as an agent of radicalisation [1] with one study [26] finding conservative media as a more effective radicaliser into radical Islam than jihadist propaganda.

4.4 Contradictions and Disagreements

When conducting literature reviews, contradictions, or discrepancies between the findings of different papers often arise. In this section, I will explore the three main contradictions I found within this review: which demographics are at risk of radicalisation, whether echo chambers are real and still relevant, and whether social media is actually an important factor in modern radicalisation.

4.4.1 Demographic Affected

To better counter and prevent radicalisation, it is important to understand who is most at risk of it, however, there does not seem to be much academic consensus on who the demographic most affected by radicalisation is. There are disagreements in nearly all demographical categories including: gender,

age, socio-economic context, education level, social context, mental health, religious affiliation, and immigration status.

Of the 30 papers, 16 either explicitly mentioned or implied the gender of the people most affected by the radicalisation studied. The implication was done either by describing the main appeal as hypermasculinity [2], offering a bride in a traditional wedding as a reward for joining [4], or generally making misogyny one of the main talking points [3]. 10 papers either only studied men or determined that men were the most affected, 3 either only studied women or determined that women were most affected, while the remaining papers had a more nuanced take on the subject. One [18] found that while men are more likely to be radicalised, Islamic terrorism is undergoing a shift that allows for more active participation and recruitment of women. Another [21] found that it depends on the radicalising idea, with women being more radical in humanitarian concepts and men being more radical in the rest. And the final paper [26] found that men are more likely to be radicalised in general, but that women are more affected by social media radicalisation specifically. Overall, men are more studied—understandable since men comprise the majority of arrested terrorists (91% in the UK for example [87])—but terrorism is not the only kind of radicalisation and between the actual demographical studies on the topic, and the shift in favour of female recruitment, future research should expand on the radicalisation of women, especially in computer science if results of this last study are corroborated and social media really is more impactful on women.

All papers that mentioned aged agreed that 'young' people are more vulnerable to radicalisation, but the definition of young is rarely given and different between papers. Only one paper studies pre-teens (ages 11-14) [28]. A few either explicitly mention teenagers [12][23] or conducted research on high school students (which implies teenagers) [20]. Most other papers either remain vague or seem to refer to people in their late teens and early twenties, with multiple mentions of university degrees and higher education [9][11][24].

When discussing socio-economic position, most papers [2][4][12][13][17][18] found that people experiencing financial hardship or in impoverished countries were more likely to be radicalised, but three disagreed. One paper [30] found that while the studied radical group (FPI) does recruit within poorer areas too, many participants were middle class and employed. Another [26] found that radicalised individuals had diverse socio-economic backgrounds. And the final paper [24] discussed the specific radicalisation of one girl through her *Tumblr* blog and found that she was fairly privileged.

The topic of education also has split results, with some papers finding that people with an education—usually in the form of university bachelor's or high school degrees—are more likely to be radical [11][30] and others finding that people without are [8][16]. Two studies [9][24] found people getting an education—specifically university students—to be most at risk of radicalisation. Studies do not agree on how results within formal education affect radicalisation either, as one paper [21] found that poor grades in school were an indicator of vulnerability to radicalisation, while another [23] found that most of the radicalised people analysed were considered academically gifted and even likely to go into higher education before their radicalisation (though they did drop out to become violent extremists).

Regarding individual's social lives, social isolation is deemed a factor by some papers [17][20], one [26] saying "lone wolves" are more vulnerable to radicalisation. This is consistent with assertions that a sense of collective belonging and group identity is part of the appeal of radical groups [11][25][27]. However, a paper on internet addiction as a means of political radicalisation [9] found that the radical subjects were not socially isolated, and other papers [23][24] supporting this stance with examples of

radicalised individuals that were close to their (not radical) families and did not appear to be socially excluded.

Suffering from depression was found to make individuals more vulnerable to radicalisation and likely to endorse violence in some cases [9][10]. However, experts in cultural psychiatry [27] found that most individuals in organized terrorist groups do not have any psychiatric disorders, with mental illness only playing a role in extremists that act alone. One study that tracked mental illness diagnoses as one of its variables in radical youth in Australia [23] found that 42.4% of their sample was diagnosed as mentally ill before they were radicalised. This is more than the national average, since the government data on the 2017-2018 period [88]—time period chosen because it was the closest to the publishing date of the paper (2020) other than pandemic-related statistics, which would be irrelevant since the analysed people were radicalised before COVID—found that 30.0% of females and 21.3% of males between 15 and 24 years of age were diagnosed with mental or behavioural conditions. However, it still implies that 57.6% of those radicalised to violent extremism did not have mental illness diagnoses, which does not mean they were not mentally ill but still warrants consideration.

The religious status of those most at risk depends on the type of radicalisation, and even then, the extent to which it is a factor is not agreed upon. For Islamic radicalisation Muslims were at times deemed more at risk [4] but a paper specifically on who is most at risk of Jihadist radicalisation [21] found that once other variables are accounted for, being Muslim was not a factor. Not all those radicalised into Jihadist extremism were Muslim, as between 6% and 23% of those radicalised in the EU converted to Islam during their radicalisation [26]. For far-right radicalisation, religion is somewhat less relevant, but when it is mentioned, it is always Christianity [13][19]. A study on general radicalisation found that people with no religious affiliation at all were more likely to support violent extremism [27].

Being an immigrant was a factor for Islamic radicalisation [18][21]. A study on foreign fighters in Germany [16] found that 48 out of the 61 foreign fighters had an immigrant background. One paper [34] explains that first-generation immigrants may find themselves caught between cultures and disenfranchised with the country they live in, which is why their immigration status makes them more vulnerable to radicalisation into Jihadist extremism. Not being an immigrant makes people more likely to be susceptible to radicalisation into the far-right [3][22]. This makes sense considering far-right discourse often uses xenophobic and anti-immigration rhetoric.

With this level of disagreement in the demographic most at risk of radicalisation, all authors on the topic must specify what specific context of radicalisation they are studying and determine who is most affected for their specific study. De-radicalisation programs also need to accommodate form this wide array of people and ensure that their propagation efforts are not based on outdated demographic profiles that exclude significant portions of the population which might still be at risk.

4.4.2 Echo Chambers

When I say echo chambers, I refer to the popular theory that social media limits exposure to diverse perspectives and can instead create groups of like-minded radical users that don't interact with the opposing side but instead reinforce each other's worldview by echoing it back and forth [89]. With this theory's rise to notoriety, arguments against echo chambers' existence and their radicalising potential have also arisen. These critiques argue that selective exposure rarely translates into the real world, that most people are unlikely to be in political echo chambers because it requires consuming far more political content than the average person does, and that while most people will engage with content they

agree with more, they will not actively avoid disagreeing content enough for platforms to create echo chambers through algorithms alone [90]. The papers in my sample are also split on the existence and effectiveness of echo chambers.

There were various papers in favour of echo chambers' existence. Some approach the topic generally, saying that the internet itself acts as an echo chamber [1][11] or filter bubble [10] that feeds on confirmation bias, normalising radical ideas and thus enabling the self-radicalisation process. Or make a similar argument about social media platforms in general, claiming that they let people isolate themselves in like-minded ideological groups [20] who can express their opinions without any genuine dialogue or moderation [27]. This goes beyond the normalisation and sharing of overt radical propaganda, as the paper about conspiracy theories and their effect on mobilisation [14], found that conspiracy theories also thrive on self-validation in Facebook-created echo chambers to appear more widely believed and legitimate. Other papers [3][9] also agree with the existence of echo chambers and further focus on the how, arguing that echo chambers are a specific consequence of social media's collection of user data to inform the algorithmic recommendation process and the 'limitless' amount of information they can use to cater to specific interests, which discourages users with opposing views from interacting with one another.

However, this idea of echo chambers hinges on technological determinism, with the social media platforms themselves segregating people into groups, and some papers disagreed with this notion, determining that radicalisation will always centre around the human component. One [5] argues that social media affordances help creators construct an echo chamber of sorts, but that it is a conscious effort on the creators' part and a direct consequence of their media literacy and understanding of how to exploit these affordances. Similarly, a second paper [15] acknowledged echo chambers as an accepted theory but with its 'imagined audience' approach found that social media content creators create or at least reinforce the phenomena by framing every post based on expected audience reception, and only post things that they think will warrant the engagement they want, thus self-censoring and making the content expressed more homogeneous and echo chamber like.

Some papers that analysed social media groups and communities as a part of their method observed how the echo chambers were maintained. They also found a significant human component to the process, which spanned across different radicalisation kinds. The exploratory study of ISIS's telegram channels [6] found that moderators kicked out and blacklisted anyone they suspected of having different opinions, only allowing accounts that visibly interacted positively with their propaganda to stay and calling all others 'spies.' This method of blocking anyone critical is used to enforce radicalization in far-right spaces too, as evidence by the WhatsApp groups in favour of Brazil's far-right congressman Jair Bolsonaro [13] doing the same thing. Even a paper [28] outside of the two main perpetrators of radicalisation found a similar system, with Facebook and WhatsApp groups that blocked anyone belonging to the opposing religion, even friends from school, to create spaces where they could simulate their own society without the religious other or otherwise exclude and insult them.

Lastly, there is one paper [7] that found that echo chambers are not all-encompassing. They found that 32.0% of conversations they analysed were with people outside the radical groups and were often argumentative. Furthermore, they found that these conversations were a much larger contributor to radicalisation than their in-group counterparts. This does not necessarily imply that echo chambers are not a thing—68% of the conversations were between radical individuals—but it does challenge both the notion that echo chambers are absolute and the assumption that removing echo chambers is always a positive thing since interactions outside of them can be even more damaging.

4.4.3 Social Media as a Lead Factor of Radicalisation

The entire purpose of this review is to understand how social media interact with radicalisation. While all studies analysed this intersection, multiple considered other factors and the wider context as well, partially trying to address which variables were most to blame for the radicalisation and thus the extent to which social media is a factor. Different conclusions were reached. Some studies found that it was the primary mode of radicalisation for at least some of their sample. Others determined that social media was an important factor but had to be accompanied by others. Lastly, some studies found that it played a secondary, supportive role.

Some papers found that social media can be the primary factor of radicalisation. Three papers found that it was the primary factor in the radicalisation of at least part of the individuals they studied. These include the case study of the blog of the woman that self-radicalised into ISIS [24], 25% of the arrested radical youth in Australia [23], and Brazil when they elected Jair Bolsonaro [13]. All these papers acknowledged that contextual factors facilitated radicalisation but found that social media was the primary culprit. In less case-specific studies, social media was called an "incredibly effective tool for propaganda and group manipulation" [27]. One study [6] found that social media can be a primary factor in ideological radicalisation but not mobilisation, while another [7] found some correlation between more activity in radical social media groups—specifically Facebook and Twitter—and offline mobilisation and protests.

Other studies found that social media was one of multiple factors, though still effective and important. The opportunity for active engagement was what made it most effective, making social media surpass mainstream media in radicalisation potential [8][29]. Buzzers that directed dozens of accounts to keep propaganda viral [30] found that their job was important in the crafting of the community, keeping their message relevant and encouraging people to mobilise, but not effective in mobilisation itself. Another paper [20] found that youth favoured using social media to develop already radical ideas which formed offline. And a third study [10] found that social media was the easiest way to coordinate mobilisation and gain a reputation for an extremist organisation, though not to radicalise people into it.

Sometimes social media was the main catalyst but needed others to fully radicalised people. For example, social media was used to attract younger people to Irish republicanism [12]. Similarly, a study examining exposure to radical content [1] found that social media created more opportunities to become radicalised, acting like a catalyst, but required social contact in the real world to speed up and complete the process. Two studies found that social media was an efficient tool to find new recruits, by either luring them in with the promises of belonging and community [11] or simply recruiting from naturally occurring 'safe spaces' for radical discussions [18], but the radicalisation actually happened through peer-to-peer recruitment in real life that could only be enhanced by social media.

Finally, two papers found that social media only played a supportive role in their studied radicalisation. One article's [16] entire purpose was the determine if social media was a primary tool for radicalisation and it found "meagre evidence" to support the notion that it could radicalise anyone without traditional peer-to-peer recruitment. The second [26] found that social media propaganda failed to interest the target demographic completely, and social media only helped radicalisation by juxtaposing information from friends in Syria with the mainstream media narrative and thus stirring feelings of injustice due to the discrepancies found.

It is important to note that some papers either did not take a stance on the efficacy of social media as a means of radicalisation or warned against considering it in a vacuum. One [22] argued that decoupling the online and offline contexts is an "onerous task" and that modern-day life is too entrenched in the digital world for it to be feasible. Another [24] found that in their specific case study, social media was the main factor but that a mixed online-real world approach is essential to understanding radicalisation as a concept. This may be the main insight provided by this section. Social media clearly has some impact on modern radicalisation, but future studies should consider how variable this effect is and ensure that their method accounts for the general context or otherwise matches the level of impact social media had on their specific sample.

Chapter 5: Critical Evaluation

In this section, I will provide a critical outlook both on the papers used as samples for the review and on the review itself. This is important to evaluate the quality of the conclusions reached and where possible biases may lie.

5.1 Limitations of Papers Studied

When discussing the limitations of the papers analysed, I will not explore the methodologies paper by paper to determine the standard limitations associated with them, because the variety of different methods within the sample serves to compensate for one another. However, I did not conduct a bias or quality check during the literature screening process, and thus some papers have specific weaknesses that are not expected from the method type that ought to be acknowledged.

The most common of these limitations is that four papers [2][13][18][19] did not have an explicit methodology section or an explanation of how the study was conducted. Some aspects of the method could be extrapolated from the other sections of the papers, but they did not provide sufficient information to replicate the study or otherwise assess the reliability of the findings. On a similar note, a paper that analysed ISIS propaganda in Indonesia [4] did have a methodology section but it did not have an explicit sample size nor a replicable explanation of how the sample was obtained.

One paper was a systematic literature review on how exposure to extremist content online related to violent radicalization [1] with only 11 eligible papers. While the number of papers a review needs is dependent on multiple factors, systematic literature reviews tend to need to survey all available empirical data, and 11 papers are unlikely to cover a topic that wide. They surveyed thousands of papers so perhaps a less strict screening phase could have given them a better overview of the literature.

Another paper [16] aimed to identify the primary mode of radicalisation in German foreign fighters. It tried to find which theory was the main cause of radicalisation and as such constructed a method that made them mutually exclusive. This is a viable method to determine which theory is the main radicaliser, but it is unusual for the social sciences. It made the results not as informative as they could have been since social media not being the only radicaliser does not preclude it from being an important one.

The paper [25] with the biggest limitation in credibility tried to determine how personality affected reactions and radicalisation potential due to exposure to extremist content online. To do this they first assessed the level of radicalisation of participants based on a social cause of their own choosing. Then, they gave them propaganda of a specific extremist group that did not necessarily coincide with the one chosen before then. Finally, they assessed their level of radicalisation again using the same test but with the extremist group they had just been exposed to as the social cause. So, a person could have answered the initial radicalisation test with Black Lives Matters in mind, for example, and then answered the second about IS and Jihadist radicalisation. They acknowledged this limitation multiple times, but upon seeing the decrease in radicalism scores between tests they also claimed that they “found that exposure to online violent extremist content did not increase tendencies for political mobilization.” This is misleading, especially since they had a control group that was not exposed to radical content and their radicalisation test scores were lower, and it poses the credibility of the paper into question.

5.2 Limitations of The Review

The main contribution of literature reviews is the interpretation and synthesis of the papers they review. However, interpretations are by nature subjective, making the conclusions reached subject to the reviewer's bias. Many categorisations in this review were done based on my understanding of the methods and results and discipline of the paper, which were often not stated explicitly or otherwise ambiguous. The choices I made may differ from those another person reading the papers would make, especially considering I do not have any formal education in most of the disciplines the papers were drawn from. This bias in interpretation is exacerbated because this review only had one reviewer.

Literature reviews are also affected by publication bias, which refers to the tendency to only publish studies with strong findings [91]. This is especially significant in this review because one of its main aims is to determine how influential social media is in modern radicalisation. Papers that did not find any strong correlation may have not been published, skewing the results, and overinflating the importance of social media.

Semi-systematic literature reviews are less comprehensive than their systematic counterpart. In this case, I only considered papers from two databases, so I may have missed relevant studies that could have been included and enhanced the results. Similarly, because I did not look at all the literature I cannot provide a comprehensive overview of all academic conversations on the topic, which may make my findings less generalisable.

I did not filter for quality when selecting my sample. This means that the papers I reviewed are not necessarily academically rigorous or otherwise unbiased and may influence my results. This is especially relevant due to the relatively small sample size of this review (30 papers).

5.3 Possible Future Expansions

A few choices in this review were made specifically because of the scope of this dissertation. For one, if I had more time more databases would be added, or ideally the majority of the literature on the topic would be considered at the beginning of the screening process, to address the third limitation mentioned above. This would probably allow for more perpetrators of radicalisation and more analysed countries, thus offering a more generalisable understanding of social media's impact on all forms of radicalisation.

Beyond that, if the quantitative aspect of the first part of the review were to be expanded, another reviewer would be ideal, to help categorise the papers and thus have a more objective grouping.

If this project was further extended, converting it into an integrative review could be extremely useful. Of the papers analysed, a few proposed theories and frameworks of radicalisation based on their sample data or case studies, but a literature review could create a theoretical model that extended beyond specific contexts. As mentioned in the *Methodology* section, this was not feasible for this dissertation because integrative reviews require a comprehensive understanding of all areas the topic discusses, which in this case would be part of at least 22 different disciplines, but given the time to achieve this, it would be even more useful than a semi-systematic review in the creation of de-radicalisation programs.

Chapter 6: Conclusion

This paper explores the academic conversation around social media's impact on political radicalisation, through a semi-systematic literature review. I addressed and attempted to answer all the aims expressed in my introduction.

When trying to determine who is most likely to be affected by radicalisation through social media, I found no general consensus in terms of demographic other than younger people. This is in itself a useful finding, as it shows that monitoring accounts or pushing de-radicalisation content based on demographic profiles is unlikely to cover enough of the group that is actually at risk. Some non-demographical aspects were found. These include a higher propensity for anger, impulsiveness, belief in conspiracy theories, and willingness to actively seek out and engage with radical content. Some are impossible to account for, but engagement and search terms could be considered when creating algorithms to detect those in the ideological radicalisation phases.

This review also identified the social media platforms most studied and the features that were used as a means to radicalisation. Facebook was the most discussed and analysed platform, though Telegram was the most associated with radicalisation, as it was the most studied when compensating for the size of the platforms themselves. The features and affordances most exploited were the user-generated nature of social media content, measurable engagement metrics, anonymity, and the accessibility of social media platforms, though there were many more. Not all of these features could or should be altered just because of how they are exploited. User-generated content is what makes social media social media for example, and making platforms less accessible merely to prevent a small minority of users from exploiting them is outrageous, but some others could be at least monitored. Instagram made likes optionally removable [92], so it isn't unheard of for social media platforms to hide the measurable engagement features, and something similar could be considered, at least for accounts that appear to be radical or otherwise politically polarising.

There was significant overlap regarding the main radicalisation perpetrators: the far-right and Jihadist extremists, the most popular social media features used for radicalisation, and the types of radical content. Within these categories, no papers contradicted each other or otherwise implied that a finding of some other paper was not corroborated when there was overlap.

I also identified three areas of academic disagreement: the most vulnerable demographic, the existence and efficacy of echo chambers, and the extent to which social media is a driving factor of radicalisation. The demographic and echo chambers conversations could benefit from studies that attempt to create one model or otherwise categorise how each subcategory is related to some aspect of radicalisation or social media platform respectively.

Lastly, I found some gaps in my sample that could be interesting for future studies to explore. The papers were mostly focused on the Western world, despite radical action occurring in other areas more frequently. Future research should expand into these areas, I propose North Africa, Afghanistan, Iraq, and Myanmar as possible locations to explore. The papers were also almost exclusively focus on either Islamic or far-right radicalisation, so future research might want to consider other perpetrators.

It is also important for research to remain up to date with the ever-changing social media landscape, so platforms that are becoming more popular for radical groups, like Telegram, should continue to be explored. I also propose that future papers could consider researching TikTok as a platform and how it aids radicalisation, due to its fast-paced increase in size and the younger demographic using the app.

Finally, while I decided to analyse radicalisation as a whole, most papers focused on violent extremism, as is evidence from 'violence' and 'terrorism' being within the top 10 mentioned keywords and 'political violence' and 'violent radicalisation' being present more than once (for a complete list of all keywords see appendix B). This is understandable—actions are easier to track than thoughts and people perpetrating violence are a more urgent issue—but the earlier stages of radicalisation and even radicalisation that stops at the cognitive aspect or engages in support of violent actions instead of propagation, are still relevant and could be useful to determine how to stop individuals from becoming violent. This would be difficult to study since people are unlikely to be honest in surveys asking about their most radical thoughts, but it is a notable gap.

Overall, this study provides important insights into the impact of social media on political radicalisation and the scholarship surrounding the issue. It identified platform features, content types and other patterns that those working at or owning social media companies should know and consider when creating prevention and de-radicalisation programs and that could hopefully be useful in identifying radical behaviour better.

References

[1] G. Hassan, et al., "Exposure to Extremist Online Content Could Lead to Violent Radicalization: A Systematic Review of Empirical Evidence," *International Journal of Developmental Sciences*, vol. 12, no. 1-2, 2018, pp. 71-88; DOI 10.3233/DEV-170233.

[2] P. Mantello, "Fatal portraits: The selfie as agent of radicalization," *Sign Systems Studies*, vol. 49, no. 3-4, 2021, pp. 566-589; DOI 10.12697/SSS.2021.49.3-4.16.

[3] T. Riebe, et al., "From Conspiracies to Insults: A Case Study of Radicalisation in Social Media Discourse," *Proc. Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI)*, 2018, pp. 595-603.

[4] S. Djelantik, "Islamic State and the social media in Indonesia," *Journal of Content, Community and Communication*, vol. 9, no. 2019, 2019, pp. 146-155; DOI 10.31620/JCCC.06.19/20.

[5] Maly, "Metapolitical New Right Influencers: The Case of Brittany Pettibone," *Social Sciences-Basel*, vol. 9, no. 7, 2020; DOI 10.3390/socsci9070113.

[6] M. Bloom, et al., "Navigating ISIS's Preferred Platform: Telegram (1)," *Terrorism and Political Violence*, vol. 31, no. 6, 2019, pp. 1242-1254; DOI 10.1080/09546553.2017.1339695.

[7] J.D. Gallacher, et al., "Online Engagement Between Opposing Political Protest Groups via Social Media is Linked to Physical Violence of Offline Encounters," *Social Media and Society*, vol. 7, no. 1, 2021; DOI 10.1177/2056305120984445.

[8] A. Lipiński and A. Stępińska, "Polish Right-Wing Populism in the Era of Social Media: The Unexpected Careers of Paweł Kukiz and Janusz Korwin-Mikke," *Problems of Post-Communism*, vol. 66, no. 1, 2019, pp. 71-82; DOI 10.1080/10758216.2018.1484667.

[9] G. Tang, et al., "Politically Motivated Internet Addiction: Relationships among Online Information Exposure, Internet Addiction, FOMO, Psychological Well-being, and Radicalism in Massive Political Turbulence," *International Journal of Environmental Research and Public Health*, vol. 17, no. 2, 2020; DOI 10.3390/ijerph17020633.

[10] M. Costello, et al., "Predictors of Viewing Online Extremism Among America's Youth," *Youth & Society*, vol. 52, no. 5, 2020, pp. 710-727; DOI 10.1177/0044118x18768115.

[11] G.F. Hollewell and N. Longpre, "Radicalization in the Social Media Era: Understanding the Relationship between Self-Radicalization and the Internet," *International Journal of Offender Therapy and Comparative Criminology*, vol. 66, no. 8, 2022, pp. 896-913; DOI 10.1177/0306624x211028771.

[12] D. Reinisch, "Teenagers and young adults in dissident Irish republicanism: a case study of Na Fianna eireann in Dublin," *Critical Studies on Terrorism*, vol. 13, no. 4, 2020, pp. 702-723; DOI 10.1080/17539153.2020.1820207.

[13] R. Evangelista and F. Bruno, "Whatsapp and political instability in Brazil: Targeted messages and political radicalisation," *Internet Policy Review*, vol. 8, no. 4, 2019; DOI 10.14763/2019.4.1434.

[14] L. Frischlich, ""Resistance!": Collective Action Cues in Conspiracy Theory-Endorsing Facebook Groups," *Media and Communication*, vol. 10, no. 2, 2022, pp. 130-143; DOI 10.17645/mac.v10i2.5182.

[15] M.H. Su, et al., "Social Media Expression, Political Extremity, and Reduced Network Interaction: An Imagined Audience Approach," *Social Media + Society*, vol. 8, no. 1, 2022; DOI 10.1177/20563051211069056.

[16] S.C. Reynolds and M.M. Hafez, "Social Network Analysis of German Foreign Fighters in Syria and Iraq," *Terrorism and Political Violence*, vol. 31, no. 4, 2019, pp. 661-686; DOI 10.1080/09546553.2016.1272456.

[17] J. Githigaro and A. Kabia, "An evaluation of factors pushing youth from Majengo, Mombasa Kenya into al-Shabaab: a methodological and theoretical analysis," *Critical Studies on Terrorism*, vol. 15, no. 2, 2022, pp. 381-398; DOI 10.1080/17539153.2022.2048989.

[18] J. Droogan, et al., "Innovation and terror: an analysis of the use of social media by terror-related groups in the Asia Pacific," *Journal of Policing, Intelligence and Counter Terrorism*, vol. 13, no. 2, 2018, pp. 170-184; DOI 10.1080/18335330.2018.1476773.

[19] J. Johnson, "The Self-Radicalization of White Men: "Fake News" and the Affective Networking of Paranoia," *Communication Culture & Critique*, vol. 11, no. 1, 2018, pp. 100-115; DOI 10.1093/ccc/tcx014.

[20] L. Pauwels and N. Schils, "Differential Online Exposure to Extremist Content and Political Violence: Testing the Relative Strength of Social Learning and Competing Perspectives," *Terrorism and Political Violence*, vol. 28, no. 1, 2016, pp. 1-29; DOI 10.1080/09546553.2013.876414.

[21] W. Pedersen, et al., "At risk for radicalization and jihadism? A population-based study of Norwegian adolescents," *Cooperation and Conflict*, vol. 53, no. 1, 2018, pp. 61-83; DOI 10.1177/0010836717716721.

[22] H. Marcks and J. Pawelz, "From Myths of Victimhood to Fantasies of Violence: How Far-Right Narratives of Imperilment Work," *Terrorism and Political Violence*, vol. 34, no. 7, 2022, pp. 1415-1432; DOI 10.1080/09546553.2020.1788544.

[23] A. Cherney, et al., "Understanding youth radicalisation: an analysis of Australian data," *Behavioral Sciences of Terrorism and Political Aggression*, vol. 14, no. 2, 2022, pp. 97-119; DOI 10.1080/19434472.2020.1819372.

[24] L. Windsor, "The Language of Radicalization: Female Internet Recruitment to Participation in ISIS Activities," *Terrorism and Political Violence*, vol. 32, no. 3, 2020, pp. 506-538; DOI 10.1080/09546553.2017.1385457.

[25] N. Shortland and P. McGarry, "The Personality and Propaganda Puzzle: Exploring the Effect of Personality on Exposure to Extremist Content Online," *Psychology of Violence*, vol. 12, no. 1, 2022, pp. 1-10; DOI 10.1037/vio0000396.

[26] C. Mazza, et al., "Mapping the impact of ISIS propaganda narratives: Evidences from a qualitative analysis in Italy and France," *Comunicazioni Sociali*, vol. 2018, no. 3, 2018, pp. 432-447; DOI 10.26350/001200_000033.

[27] C. Rousseau, et al., "Radicalization to Violence: A View from Cultural Psychiatry," *Transcultural Psychiatry*, vol. 58, no. 5, 2021, pp. 603-615; DOI 10.1177/13634615211048010.

[28] K.V. Bhatia, "'IN TIMES OF CRISIS, FOLLOWERS OF ONE TRUE GOD UNITE': SOCIAL MEDIA AND THE FORMATION OF ONLINE RELIGIOUS SILOS," *Medijske Studije-Media Studies*, vol. 10, no. 19, 2019, pp. 119-139; DOI 10.20901/ms.10.19.7.

[29] D. Lucić, et al., "Sentiment analysis of the Syrian conflict on Twitter," *Medijske Studije*, vol. 11, no. 22, 2020, pp. 46-61; DOI 10.20901/ms.11.22.3.

[30] A. Seto, "Islamist buzzers: Message flooding, offline outreach, and astroturfing," *Austrian Journal of South-East Asian Studies*, vol. 12, no. 2, 2019, pp. 187-208; DOI 10.14764/10.ASEAS-0021

[31] University of Maine, "Social Media Statistics Details - Undiscovered Maine - University of Maine," *Undiscovered Maine*, Sep. 02, 2021. https://umaine.edu/undiscoveredmaine/small-business/resources/marketing-for-small-business/social-media-tools/social-media-statistics-details/

[32] J. Wakefield, "Christchurch shootings: Social media races to stop attack footage," *BBC News*, Mar. 16, 2019. Accessed: May 03, 2023. [Online]. Available: https://www.bbc.co.uk/news/technology-47583393

[33] S. Sandberg, "Facebook Chief Operating Officer Sheryl Sandberg's letter to New Zealand," *NZ Herald*, Mar. 30, 2019. Accessed: May 03, 2023. [Online]. Available: https://www.nzherald.co.nz/business/facebook-chief-operating-officer-sheryl-sandbergs-letter-to-new-zealand/UAPCQMTI645ICB734DKQW25FQQ/?c_id=3&objectid=12217454

[34] N. Nix, "Facebook bans hate speech but still makes money from white supremacists," *Washington Post*, Aug. 10, 2022. Accessed: May 03, 2023. [Online]. Available: https://www.washingtonpost.com/technology/2022/08/10/facebook-white-supremacy-ads/

[35] P. Gill, E. Corner, M. Conway, A. Thornton, M. Bloom, and J. Horgan, "Terrorist Use of the Internet by the Numbers," *Criminology & Public Policy*, vol. 16, no. 1, pp. 99–117, Feb. 2017, doi: https://doi.org/10.1111/1745-9133.12249.

[36] J. Whittaker, "The online behaviors of Islamic state terrorists in the United States," *Criminology & Public Policy*, vol. 20, no. 1, Jan. 2021, doi: https://doi.org/10.1111/1745-9133.12537.

[37] M. Jensen, "The Use of Social Media by United States Extremists," Jul. 2018. Available: https://www.start.umd.edu/pubs/START_PIRUS_UseOfSocialMediaByUSExtremists_ResearchBrief_July2018.pdf

[38] M. Walker and K. E. Matsa, "News Consumption Across Social Media in 2021," *Pew Research Center's Journalism Project*, Sep. 20, 2021. Accessed: May 03, 2023. [Online]. Available: https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/

[39] C. T. Carr and R. A. Hayes, "Social Media: Defining, Developing, and Divining," *Atlantic Journal of Communication*, vol. 23, no. 1, pp. 46–65, Feb. 2015, doi: https://doi.org/10.1080/15456870.2015.972282.

[40] T. Aichner, M. Grünfelder, O. Maurer, and D. Jegeni, "Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019," *Cyberpsychology, Behavior, and Social Networking*, vol. 24, no. 4, Apr. 2021, doi: https://doi.org/10.1089/cyber.2020.0134.

[41] B. Wellman, J. Salaff, D. Dimitrova, L. Garton, M. Gulia, and C. Haythornthwaite, "Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community," *Annual Review of Sociology*, vol. 22, no. 1, pp. 213–238, Aug. 1996, doi: https://doi.org/10.1146/annurev.soc.22.1.213.

[42] K. Sajithra, "Social Media – History and Components," *IOSR Journal of Business and Management*, vol. 7, no. 1, pp. 69–74, 2013, doi: https://doi.org/10.9790/487x-0716974.

[43] D. M. Boyd and N. B. Ellison, "Social Network Sites: Definition, History, and Scholarship," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210–230, Dec. 2007, doi: https://doi.org/10.1111/j.1083-6101.2007.00393.x.

[44] A. M. Kaplan and M. Haenlein, "Users of the World, Unite! The Challenges and Opportunities of Social Media," *Business Horizons*, vol. 53, no. 1, pp. 59–68, Jan. 2010.

[45] P. N. Howard and M. R. Parks, "Social Media and Political Change: Capacity, Constraint, and Consequence," *Journal of Communication*, vol. 62, no. 2, pp. 359–362, Feb. 2012, doi: https://doi.org/10.1111/j.1460-2466.2012.01626.x.

[46] E. Ortiz-Ospina and M. Roser, "The rise of social media," *Our World in Data*, Sep. 2019, Accessed: May 03, 2023. [Online]. Available: https://ourworldindata.org/rise-of-social-media?ref=tms

[47] A. Perrin, "Social Media Usage: 2005-2015," Pew Research Center, Oct. 2015. Accessed: May 03, 2023. [Online]. Available: https://assets.pewresearch.org/wp-content/uploads/sites/14/2015/10/PI_2015-10-08_Social-Networking-Usage-2005-2015_FINAL.pdf

[48] D. M. D. Silva, "'Radicalisation: the journey of a concept,' revisited," *Race & Class*, vol. 59, no. 4, pp. 34–53, Jan. 2018, doi: https://doi.org/10.1177/0306396817750778.

[49] P. R. Neumann *et al.*, "Papers from the First International Conference on Radicalisation and Political Violence Perspectives on Radicalisation and Political Violence," 2008. Accessed: May 03, 2023. [Online]. Available: https://www.nonviolent-conflict.org/wp-content/uploads/2016/11/Perspectives-on-Radicalisation-Political-Violence.pdf

[50] M. Sedgwick, "The Concept of Radicalization as a Source of Confusion," *Terrorism and Political Violence*, vol. 22, no. 4, pp. 479–494, Sep. 2010, doi: https://doi.org/10.1080/09546553.2010.491009.

[51] "radicalization," *OED Online*. Oxford University Press, Mar. 2022. Accessed: May 03, 2023. [Online]. Available: https://www.oed.com/view/Entry/272267?redirectedFrom=radicalisation&

[52] "radical," *OED Online*. Oxford University Press, Dec. 2022. Accessed: May 03, 2023. [Online]. Available: https://www.oed.com/view/Entry/157251

[53] J. Harman *et al.*, *AN ACT To prevent homegrown terrorism, and for other purposes.* 2007. Available: https://www.govtrack.us/congress/bills/110/hr1955/text

[54] Home Office, "Get help if you're worried about someone being radicalised," *GOV.UK*, Jul. 2022. https://www.gov.uk/guidance/get-help-if-youre-worried-about-someone-being-radicalised (accessed May 03, 2023)

[55] AIVD, "From dawa to jihad The various threats from radical Islam to the democratic legal order," Mar. 2005. Accessed: May 03, 2023. [Online]. Available: https://irp.fas.org/world/netherlands/dawa.pdf

[56] P. R. NEUMANN, "The trouble with radicalization," *International Affairs*, vol. 89, no. 4, pp. 873–893, Jul. 2013, doi: https://doi.org/10.1111/1468-2346.12049.

[57] A. Schmid, "Radicalisation, De-Radicalisation, Counter-Radicalisation: A Conceptual Discussion and Literature Review," *Terrorism and Counter-Terrorism Studies*, vol. 4, no. 2, 2013, doi: https://doi.org/10.19165/2013.1.02.

[58] A. Hoskins and B. O'Loughlin, "Media and the myth of radicalization," *Media, War & Conflict*, vol. 2, no. 2, pp. 107–110, Jun. 2009, doi: https://doi.org/10.1177/1750635209105608.

[59] A. Kundnani, "Radicalisation: the journey of a concept," *Race & Class*, vol. 54, no. 2, pp. 3–25, Sep. 2012, doi: https://doi.org/10.1177/0306396812454984.

[60] A. Orsini, "What Everybody Should Know about Radicalization and the DRIA Model," *Studies in Conflict & Terrorism*, vol. 46, no. 1, pp. 1–33, Mar. 2020, doi: https://doi.org/10.1080/1057610x.2020.1738669.

[61] B. Doosje, F. M. Moghaddam, A. W. Kruglanski, A. de Wolf, L. Mann, and A. R. Feddes, "Terrorism, radicalization and de-radicalization," *Current Opinion in Psychology*, vol. 11, no. 11, pp. 79–84, Oct. 2016, doi: https://doi.org/10.1016/j.copsyc.2016.06.008.

[62] C. R. Mccauley and S. Moskalenko, *Friction how conflict radicalizes them and us*. Oxford New York Oxford University Press, 2011.

[63] D. C. Hofmann, "Clark McCauley and Sophia Moskalenko.Friction: How Radicalization Happens to Them and Us," *Terrorism and Political Violence*, vol. 24, no. 5, pp. 865–866, Nov. 2012, doi: https://doi.org/10.1080/09546553.2012.728938.

[64] A. W. Kruglanski, BélangerJ. J., and Rohan Gunaratna, *The three pillars of radicalization: needs, narratives, and networks*. New York, Ny: Oxford University Press, 2019.

[65] H. Snyder, "Literature Review as a Research methodology: an Overview and Guidelines," *Journal of Business Research*, vol. 104, no. 1, pp. 333–339, Nov. 2019, doi: https://doi.org/10.1016/j.jbusres.2019.07.039.

[66] D. Evans and I. Kowanko, "Literature reviews: evolution of a research methodology," *Australian Journal of Advanced Nursing*, vol. 18, no. 2, Mar. 2000.

[67] J. Davis, K. Mengersen, S. Bennett, and L. Mazerolle, "Viewing systematic reviews and meta-analysis in social research through different lenses," *SpringerPlus*, vol. 3, no. 1, p. 511, 2014, doi: https://doi.org/10.1186/2193-1801-3-511.

[68] D. J. MacInnis, "A Framework for Conceptual Contributions in Marketing," *Journal of Marketing*, vol. 75, no. 4, pp. 136–154, Jul. 2011, doi: https://doi.org/10.1509/jmkg.75.4.136.

[69] A. Eykelbosh and D. Fong, "Conducting a Literature Search & Semi-Systematic Review: the NCCEH Approach," Feb. 2017. Accessed: May 03, 2023. [Online]. Available: https://ncceh.ca/sites/default/files/Conducting_Semi_Systematic_Review_NCCEH_Approach-Eykelbosh-Fong.pdf

[70] P. S. Gautam, "Sufi-Barelvis, Blasphemy and Radicalization: A Critical Analysis," *Strategic Analysis*, vol. 46, no. 5, pp. 459–472, Sep. 2022, doi: https://doi.org/10.1080/09700161.2022.2115237.

[71] J. Hinds, E. J. Williams, and A. N. Joinson, "'It wouldn't happen to me': Privacy concerns and perspectives following the Cambridge Analytica scandal," *International Journal of Human-Computer Studies*, vol. 143, no. 2020, p. 102498, Nov. 2020, doi: https://doi.org/10.1016/j.ijhcs.2020.102498.

[72] A. Mondon and A. Vaughan, "The Trump Presidency and Mainstreaming of Far-Right Politics," *www.gale.com*, 2021. https://www.gale.com/intl/essays/aurelien-mondon-antonia-vaughan-trump-presidency-mainstreaming-far-right-politics (accessed May 03, 2023).

[73] B. E. Sawe, "English Speakers By Country," *WorldAtlas*, May 14, 2018. https://www.worldatlas.com/articles/english-speakers-by-country.html (accessed May 03, 2023).

[74] Vision of Humanity, "Global Terrorism Index," *Vision of Humanity*, 2023. https://www.visionofhumanity.org/maps/global-terrorism-index/#/ (accessed May 03, 2023).

[75] "Social media penetration in Africa by region 2021," *Statista*, Feb. 2022. https://www.statista.com/statistics/1190628/social-media-penetration-in-africa-by-region/ (accessed May 03, 2023).

[76] S. Kemp, "Digital 2022: Afghanistan," *DataReportal – Global Digital Insights*, Feb. 15, 2022. https://datareportal.com/reports/digital-2022-afghanistan (accessed May 03, 2023).

[77] S. Kemp, "Digital 2023: Iraq," *DataReportal – Global Digital Insights*, Feb. 13, 2022. https://datareportal.com/reports/digital-2023-iraq#:~:text=More%20broadly%2C%2075.7%20percent%20of (accessed May 03, 2023).

[78] We Are Social, "Myanmar: social media penetration 2016-2022," *Statista*, Feb. 2022. https://www.statista.com/statistics/883751/myanmar-social-media-penetration/#:~:text=Social%20media%20users%20as%20share%20of%20total%20population%20Myanmar%202016%2D2022&text=In%202022%2C%20about%2037.7%20percent (accessed May 03, 2023).

[79] Institute for Economics & Peace, "GLOBAL TERROR INDEX 2023," Mar. 2023. https://www.visionofhumanity.org/wp-content/uploads/2023/03/GTI-2023-web-170423.pdf (accessed May 03, 2023).

[80] E. M. Kearns, A. E. Betus, and A. F. Lemieux, "Why Do Some Terrorist Attacks Receive More Media Attention Than Others?," *Justice Quarterly*, vol. 36, no. 6, pp. 985–1022, Jan. 2019, doi: https://doi.org/10.1080/07418825.2018.1524507.

[81] S. Jones, C. Doxsee, and N. Harrington, "The Escalating Terrorism Problem in the United States," *Center for Strategic and International Studies*, Jun. 17, 2020. https://www.csis.org/analysis/escalating-terrorism-problem-united-states (accessed May 03, 2023).

[82] L. Andersson, "What's left of the radical left online? Absence of communication, political vision, and community in autonomist web milieus in Sweden," *New Media & Society*, vol. 20, no. 1, pp. 384–398, Nov. 2016, doi: https://doi.org/10.1177/1461444816677533.

[83] R. Tan, "Terrorists' love for Telegram, explained," *Vox*, Jun. 30, 2017. https://www.vox.com/world/2017/6/30/15886506/terrorism-isis-telegram-social-media-russia-pavel-durov-twitter (accessed May 03, 2023).

[84] C. Bucknell Bossen and R. Kottasz, "Uses and gratifications sought by pre-adolescent and adolescent TikTok consumers," *Young Consumers*, vol. 21, no. 4, Nov. 2020, doi: https://doi.org/10.1108/yc-07-2020-1186.

[85] Insider Intelligence, "TikTok users worldwide (2020-2025)," *Insider Intelligence*, Jun. 01, 2022. https://www.insiderintelligence.com/charts/global-tiktok-user-stats/ (accessed May 03, 2023).

[86] C. H. Basch, Z. Meleo-Erwin, J. Fera, C. Jaime, and C. E. Basch, "A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok," *Human Vaccines & Immunotherapeutics*, vol. 17, no. 8, pp. 1–5, Mar. 2021, doi: https://doi.org/10.1080/21645515.2021.1894896.

[87] House of Commons Library, "Terrorism in Great Britain: the statistics," Jul. 2022. Accessed: May 03, 2023. [Online]. Available: https://researchbriefings.files.parliament.uk/documents/CBP-7613/CBP-7613.pdf

[88] Australian Bureau Of Statistics, "Mental health, 2017-18 financial year | Australian Bureau of Statistics," *www.abs.gov.au*, 2018. https://www.abs.gov.au/statistics/health/mental-health/mental-health/2017-18

[89] M. Cinelli, G. D. F. Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, "The echo chamber effect on social media," *Proceedings of the National Academy of Sciences*, vol. 118, no. 9, Mar. 2021, doi: https://doi.org/10.1073/pnas.2023301118.

[90] A. Guess, B. Lyons, B. Nyhan, and J. Reifler, "Avoiding the echo chamber about echo chambers: Why selective exposure to like-minded political news is less prevalent than you think," Jan. 2018. Accessed: May 03, 2023. [Online]. Available: https://www.researchgate.net/publication/330144926_Avoiding_the_echo_chamber_about_echo_chambers_Why_selective_exposure_to_like-minded_political_news_is_less_prevalent_than_you_think

[91] A. S. Nair, "Publication bias - Importance of studies with negative results!," *Indian Journal of Anaesthesia*, vol. 63, no. 6, p. 505, Jun. 2019, doi: https://doi.org/10.4103/ija.ija_142_19.

[92] C. Criddle, "Instagram lets users hide likes to reduce social media pressure," *BBC News*, May 26, 2021. Available: https://www.bbc.co.uk/news/technology-57254488

Appendix A

This appendix is a table showing the complete list of papers in my review sample.

	Title	Author(s)	Journal	Year
1	Exposure to Extremist Online Content Could Lead to Violent Radicalization: A Systematic Review of Empirical Evidence	Ghayda Hassan; Sébastien Brouillette-Alarie; Séraphin Alava; Divina Frau-Meigs; Lysiane Lavoie; Arber Fetiu; Wynnpaul Varela; Evgueni Borokhovski; Vivek Venkatesh; Cécile Rousseau; Stijn Sieckelinck	International Journal of Developmental Sciences	2018
2	Fatal portraits: The selfie as agent of radicalization	Peter Mantello	Sign System Studies	2021
3	From Conspiracies to Insults: A Case Study of Radicalisation in Social Media Discourse	Thea Riebe, Katja Pätsch, Marc-André Kaufhold, Christian Reuter	MuC Workshopband	2018
4	Islamic State and the Social Media in Indonesia	Sukawarsini Djelantik	Journal of Content, Community and Communication	2019
5	Metapolitical New Right Influencers: The Case of Brittany Pettibone	Ico Maly	The Social Sciences	2020
6	Navigating ISIS's Preferred Platform: Telegram	Mia Bloom, Hicham Tiflati, John Horgan	Terrorism and Political Violence	2017
7	Online Engagement Between Opposing Political Protest Groups via Social Media is Linked to Physical Violence of Offline Encounters	John D. Gallacher, Marc W. Heerdink, Miles Hewstone	Social Media + Society	2021
8	Polish Right-Wing Populism in the Era of Social Media	Artur Lipiński, Agnieszka Stępińska	Problems of Post-Communism	2018
9	Politically Motivated Internet Addiction: Relationships among Online Information Exposure, Internet Addiction, FOMO, Psychological Well-being, and Radicalism in Massive Political Turbulence	Gary Tang, Eva P. W. Hung, Ho-Kong Christopher Au-Yeung, Samson Yuen	International Journal of Environmental Research and Public Health (IJERPH)	2020
10	Predictors of Viewing Online Extremism Among America's Youth	Matthew Costello, Rebecca Barrett-Fox, Colin Bernatzky, James Hawdon, Kelly Mendes	Youth & Society	2018
11	Radicalization in the Social Media Era: Understanding the Relationship between Self-Radicalization and the Internet	Georgia F. Hollewell, Nicholas Longpré	International Journal of Offender Therapy and Comparative Criminology	2021
12	Teenagers and young adults in dissident Irish republicanism: a case study of Na Fianna Éireann in Dublin	Dieter Reinisch	Critical Studies on Terrorism	2020
13	WhatsApp and political instability in Brazil: targeted messages and political radicalisation	Rafael Evangelista, Fernanda Bruno	Internet Policy Review	2019
14	"Resistance!": Collective Action Cues in Conspiracy Theory-Endorsing Facebook Groups	Lena Frischlich	Media and Communication	2022
15	Social Media Expression, Political Extremity, and Reduced Network Interaction: An Imagined Audience Approach	Min-Hsin Su, Jiyoun Suk, and Hernando Rojas	Social Media + Society	2022
16	Social Network Analysis of German Foreign Fighters in Syria and Iraq	Sean C. Reynoldsaand Mohammed M. Hafez	Terrorism and Political Violence	2017
17	An evaluation of factors pushing youth from Majengo, Mombasa Kenya into al-Shabaab: a methodological and theoretical analysis	John Githigaro and Alex Kabia	Critical Studies on Terrorism	2022

18	Innovation and terror: an analysis of the use of social media by terror-related groups in the Asia Pacific	Julian Droogan, Lise Waldek & Ryan Blackhall	JOURNAL OF POLICING, INTELLIGENCE AND COUNTER TERRORISM	2018
19	The Self-Radicalization of White Men: "Fake News" and the Affective Networking of Paranoia	Jessica Johnson	Communication Culture & Critique	2018
20	Differential Online Exposure to Extremist Content and Political Violence: Testing the Relative Strength of Social Learning and Competing Perspectives	Lieven Pauwels, Nele Schils	Terrorism and Political Violence	2016
21	At risk for radicalization and jihadism? A population-based study of Norwegian adolescents:	Willy Pedersen, Viggo Vestel, Anders Bakken	Cooperation and Conflict	2018
22	From Myths of Victimhood to Fantasies of Violence: How Far-Right Narratives of Imperilment Work	Holger Marcks, Janina Pawelz	Terrorism and Political Violence	2020
23	Understanding youth radicalisation: an analysis of Australian data	Adrian Cherney, Emma Belton, Siti Amirah Binte Norham, Jack Milts	Behavioral Sciences of Terrorism and Political Aggression	2020
24	The Language of Radicalization: Female Internet Recruitment to Participation in ISIS Activities	Leah Windsor	Terrorism and Political Violence	2018
25	The personality and propaganda puzzle: Exploring the effect of personality on exposure to extremist content online.	Neil Shortland, Presley McGarry	Psychology of Violence	2021
26	Mapping the impact of ISIS propaganda narratives: Evidence from a qualitative analysis in Italy and France	Caterina Mazza; Sara Monaci; Gabriella Taddeo	COMUNICAZIONI SOCIALI ON-LINE	2018
27	Radicalization to Violence: A View from Cultural Psychiatry	Cécile Rousseau, Neil K. Aggarwal, and Laurence J. Kirmayer	Transcultural Psychiatry	2021
28	'IN TIMES OF CRISIS, FOLLOWERS OF ONE TRUE GOD UNITE': SOCIAL MEDIA AND THE FORMATION OF ONLINE RELIGIOUS SILOS	Kiran Vinod Bhatia	MEDIJSKE STUDIJE-MEDIA STUDIES	2019
29	Sentiment Analysis of the Syrian Conflict on Twitter	Danijela Lucić, Josip Katalinić, Tomislav Dokman	Medijske studije	2021
30	Islamist buzzers: Message flooding, offline outreach, and astroturfing	Ario Seto	ASEAS	2019

Appendix B

This is a comprehensive list of all keywords used in the sample papers sorted by frequency.

Key Words	
radicalization	12
social media	10
islam	6
internet	4
radicalisation	4
violence	3
terrorism	3
influencer	2
social movement	2
political polarization	2
social movement theory	2
twitter	2
political violence	2
narratives	2
hate speech	2
online radicalization	2
violent radicalization	2
online violent radicalization	1
extremist online content	1
systematic review	1
micro-influencer	1
terror	1
selfie	1
self-portrait	1
jihad	1
branding	1
networked communities	1
social photo	1
necro-celebritism	1
social media analytics	1
populsim	1
afd	1
islamic state	1
anonymous sharing portals	1
communication	1
recruitment	1
brittany pettibone	1
new right	1
metapolitics	1
metapolitical influencer	1
isis	1
telegramchannels	1
jihadi chat rooms	1
group processes	1
social networks	1
intergroup contact	1
political extremism	1
internet addiction	1
depression	1
hong kong	1
delinquency	1
media	1
electronic	1
peers	1
technology	1
the internet	1
mturk	1
dissident irish republicans	1
na fianna éireann	1
interviews	1
ireland	1
political micro-targeting	1
surveillance	1
visibility	1
brazil	1
collective action	1
conspiracy theories	1
facebook	1
facebook groups	1
non-normative collective action	1
popularity cues	1
virtual groups	1
expression effect	1
imagined audience	1
political expression	1
political extremity	1
context collapse	1
self-presentation	1
european muslims	1
foreignfighters	1
socialnetwork analysis	1
al-shabaab	1
kenya	1
majengo	1
islamic radicalisation	1
jihadi salafism ideology	1
violent extremism	1
abu sayyaf	1
jamaat-e-islami	1
alt-right	1
uyghurs	1
white nationalism	1
masculinity	1
conspiracy theory	1
alex jones	1
donald trump	1
belgium	1
differential association	1
exposure	1
new social media	1
politicalviolence	1
social learning	1
youth delinquency	1
adolescents	1
jihadism	1
survey	1
syria	1
far right	1
dangerous speech	1
violentextremism	1
youth	1
australia	1
open-source data	1
genderdiscourse	1
islamicstate	1
socialmedia	1
reinforcement sensitivity theory	1
bas traits	1
personality	1
jihadist narratives	1
social media narratives	1
foreign fighters	1
social polarization	1
extremism	1
mental health	1
culture	1
social media ethnography	1
online religious silos	1
radical religiosity	1
adolescent studies	1
india	1
syria conflict	1
arab spring	1
syria miss ile strikes	1
sentiment analysis	1
astroturfing	1
right-wing politics	1
social media buzzer	1