Encrypted and (In)authentic: Telegram and the Resurgence of Iran-Backed Paramilitaries in Iraq

Sleibi, Abdullatif

Encrypted and (In)authentic: Telegram and the Resurgence of Iran-Backed Paramilitaries in Iraq

Submitted by: Abdullatif (Latif) Sleibi
MSc of International Relations & Diplomacy (MIRD)

Submitted on: 19/05/2022
Word count: 23,886

Supervisor: Dr. Graig Klein
Second reader: Dr. Jaroslaw Kantorowicz

Abstract:

Rationalist perspectives have dictated that violent non-state actors will take credit for their attack(s). Despite this, there has been an increasing tendency for such groups to make the strategic choice to blame their attacks on fictitious groups in order to avoid political and social repercussions. This can be seen in the illustrative case of Iran-backed paramilitaries in Iraq, and their hurried creation of fictitious Telegram-based surrogates. Such groups have exploited the online Telegram space, where anonymity and encryption allows for the effortless creation of 'shell' groups with no physical presence, to exert an outsized impact on public perception through a combination of extremist rhetoric and graphic violence. The existence of such *inauthentic* groups is hazardous given the uncertainty it creates for the public and the broader state apparatus. Combining research on strategic credit-claiming and militias in the online space, this study has sought to provide an exploratory analysis of *the Telegram content of Iran-backed paramilitary surrogates in Iraq, and what it can reveal about their authenticity.* Particularly, this study has sought to provide a methodology that can help classify fictitious surrogates. To do so, a mix of textual analysis, network analysis, and machine learning was employed. At the core of the analysis is the prediction that high similarity in published content between surrogates and their paramilitary sponsor is a likely indicator of *inauthenticity.* This study concludes that this prediction is accurate, and that similarity is a useful indicator of authenticity; despite various limitations, the study provides important groundwork upon which future research can be conducted.

1. Introduction

1.1: Overview of Topic

Organised armed groups, specifically violent non-state actors (VNSAs), have been historically known to claim responsibility for attacks to differentiate themselves from ordinary criminals, and either highlight the purpose of the accomplished violence, or prevent it from being obscured (LaFree & Dugan, 2004; Rapoport, 1997). Despite this, VNSAs have been also found to strategically decide when they wish to claim credit for their attacks (Kearns, 2021). A large portion of modern violent incidents, often assumed to be committed by armed or terrorist groups, usually go unclaimed (LaFree & Dugan, 2014). The reasoning behind this ostensibly counterintuitive behaviour likely hinges on public backlash and political returns (Lake, 2002). VNSAs like militias are quite cognisant of their need for public support, and will therefore customarily avoid taking credit for attacks that entail harm to civilian targets and consequent negative political returns (Abrahms & Conrad, 2017). This rationale tends to persist even in the face of benefits such as publicity, intimidation of adversaries, and sending a political message (Kearns et al., 2014).

To offset their exposure to political backlash, VNSAs have accordingly taken to using methods of indirect credit-claiming. In simple terms, VNSAs have made the conscious move to lie about attacks through the means of false claiming, false blaming, or lies of omission (Kearns et al., 2014, p. 423). A prominent example of this claiming approach is the use of proxies (often in the form of surrogates) to indirectly claim credit for a violent incident. This method of false blaming is assumed to entail a genuine armed group creating a fictitious entity as a blameworthy front. This tactic can be quite tricky, as there needs to be enough credible evidence that the fictitious entity is indeed responsible for the attack, and more significantly, minimal to no

evidence that the entity was invented by the true perpetrator (Bale, 1996; Hoffman, 1997). Distinctly, it can then be argued that at the core of this false credit-claiming approach, are questions of *authenticity*. As a concept, *authenticity* can be functionally articulated as the degree of autonomy an actor has in representing itself through relatable rhetoric, competent action, and attitudinal evaluations of their relationships (Ryan & Deci, 2000; Tracy & Robles, 2013, pp. 22-23; Van Leeuwen, 2001). For a fictitious entity to be a credible front, it must demonstrate that it is *authentic*, or mask the features that make it appear *inauthentic.* In practice, a prominent example of false credit-claiming, and specifically false blaming, can be seen in the case of Iran-backed militias in Iraq and their exploitation of proxies (precisely surrogates).

To put things in context, Iran-backed militia groups have been observed to have modernised their appeal in the last decade, and as such, have succeeded in garnering an impressive online following and audience, especially on social media services like Telegram (Knights et al., 2021a). In particular, major Iran-backed paramilitaries have built a deliberate network of proxies in the form of surrogates[1], and exploited said groups and their multitude of Telegram channels to disseminate propaganda to a large public audience; in doing so, they have exerted an outsized impact on public perception by employing a combination of extremist rhetoric and graphic violence (Berger & Morgan, 2015; Ward, 2020). Markedly, the conduct of Iran-backed paramilitaries on Telegram mirrors the Islamic State's (IS), Al-Qaeda's or Boko Haram's use of social media and online platforms [Twitter, Telegram, Youtube] (Chatfield et al., 2015; Pieslak et al., 2021; Weimann, 2010; Zenn, 2020). Altogether, the ability of such groups to influence audiences through online platforms and social media, and their widespread use by

[1] When referring to surrogates, this project is pointing to the new wave of Iran-backed paramilitary surrogate groups. Although some of these groups may be semi-autonomous, all have been linked to major Iran-backed paramilitaries, namely the muqawama, that have operated in Iraq over the last two decades. This study will use the following terms to refer to such groups: surrogate, proxy, proxy surrogate.

VNSAs, has therefore motivated a growing body of research on these groups' online strategies and their effect on conflict settings (Conway, 2017; Conway et al., 2019; Mroszczyk & Abrahms, 2021; Pieslak et al., 2021; Prucha, 2016; Weimann, 2015).

Beyond the broader academic dimension, Iran-backed paramilitaries continue to be a serious player in the current Iraqi political-security landscape (Knights et al., 2021a). Accordingly, front-line researchers and policy-makers have again and again scrutinised Telegram channels belonging to formal Iran-backed militias (hereafter the muqawama[2]), such as Kata'ib Hezbollah, Kata'ib Sayyid al-Shuhada, or Asa'ib Ahl al-Haq, who have a conspicuous physical presence[3], and a clear record of military operations (Mansour, 2021; Nada & Rowan, 2021). However, this scrutiny has also extended to a set of seemingly contemporary paramilitary actors. Over the last two years, there has been a clear resurgence in violence, digital communication, and online displays of brutality by a new wave of paramilitary surrogate groups with strong links to enduring Iran-backed militia networks. Consequently, such a resurgence has garnered renewed interest in the more fundamental behaviour, formation, and origin of surrogate Iran-backed paramilitaries (Badawi, 2021), particularly as a response to the rising frequency of unclaimed or falsely claimed attacks and violent incidents (Knights & Smith, 2022).

To be precise, up to 30 new surrogate groups have risen to prominence in Iraq, with many such groups being discursively credited with anti-United States (US) coalition operations and attacks targeting domestic rivals after posting evidence on social media (Seligman, 2021). Yet, their *authenticity*, namely if such groups are actually fake profiles employed by larger groups,

[2] The muqawama is a self-styled Tehran-backed resistance faction operating out of the broader Popular Mobilisation Forces (PMF) umbrella organisation (Knights, 2019). This faction has taken a leading role in coordinating attacks against US-led coalition forces. The following groups belong to the muqawama: Kata'ib Hezbollah, Kata'ib Sayyid al-Shuhada, Asa'ib Ahl al-Haq, elements of the Badr Organisation, Kata'ib al Imam Ali, Harakat Hezbollah al-Nujba (Nada & Rowan, 2021).

[3] Well known groups belonging to factions like the muqawama have an obvious physical presence and voice, particularly seen through representatives and spokespeople (Mansour, 2021; Knights et al., 2020). The same level of on-the-ground social/political visibility has not been seen in regards to newer surrogates and groups (Badawi, 2021).

has been challenged by political analysts working in the region (Badawi, 2021). The appearance of a new wave of politically motivated surrogates in the months following a major US airstrike that killed two prominent Iraqi paramilitary leaders[4] has been argued to be a rather convenient response that is unlikely to be organic (Elias, 2020; Knights et al., 2021a).

Additionally, taking into account that paramilitary operations in Iraq are semi-legal and commonly exist within/are linked to a state-sponsored hierarchy called the Popular Mobilisation Forces[5] (al-hashd al-sha'abi or PMF), essentially a *de facto* national guard (Mansour, 2018), evidence of a new group's formation and activity is often difficult to conceal (Mansour, 2021; Knights et al., 2020). Keeping this disparity of evidence and activity in mind, analysts and policy-makers in the region have taken to subsequently questioning the *authenticity* of these new groups, as there appears to be a lack of concrete evidence indicating the physical presence, or position, occupied by a majority of these new-wave surrogates in the Iraqi paramilitary (PMF) hierarchy (Badawi, 2021; Elias, 2020; Knights et al., 2021a).

Taking this physical ambiguity into account, and previous incidents where social media outlets like Telegram or Twitter were used by major Iraqi paramilitary groups to create 'shell' profiles for the purpose of publicising their attacks (Badawi, 2021; Knights et al., 2020), the *authenticity* of these new paramilitaries can be reasonably scrutinised. To this end, this study will aim to explore the following research question (RQ): what can published Telegram content reveal about the *authenticity* of the new wave of Iran-backed paramilitary groups in Iraq? This study will employ textual analysis, network analysis, and machine learning methodologies to

[4] This refers to the killing of the two iconic paramilitary leaders, Qassem Soleimani and Abu Mahdi al-Muhandis. The latter served as vice chairman and operational commander of the PMF, and filled the role of the senior Iraqi representative of Iran's Islamic Revolutionary Guard Corps-Qods Force (IRGC-QF) (Knights, 2022), while the former served as the commander of the aforementioned Qods Force (Knights et al., 2020).

[5] The PMF is an umbrella organisation that coordinates militias operating in Iraq and serves as part of the security apparatus of the state. The organisation was initially created in order to coordinate militias/paramilitary groups and their efforts to combat Islamic State (IS) forces in Iraq (Nada & Rowan, 2021). The PMF is composed of upwards of 140,000 fighters divided over around 70 groups (CEDOCA, 2019, pp. 17-18).

study the online discourse of the muqawama and their surrogates, and establish whether similarities in discourse can indicate the *authenticity* of a surrogate group.

1.2: Relevance and Background

Taking a step back, it is worth concisely examining the broader structure of the Iraqi paramilitary network, and the sponsorship system underlying it. At the start of the chain, sits the Islamic Republic of Iran, the principal sponsor backing muqawama paramilitaries that officially operate in Iraq as part of the PMF. Overall, this Islamic regime has been well-known to employ a hybrid approach in the pursuit of its strategic objectives in the Middle East and North Africa (MENA) region, with it leveraging a range of (un)conventional capabilities, including proxy (militia) forces (Dalton, 2017, p. 2; McInnis, 2016). Nevertheless, recent behaviour has shown that the muqawama has evolved past the yoke of absolute Iranian control (Badawi, 2021; Malik, 2021), and lengthened the chain to become a sponsor in its own right. Essentially, going beyond Iran-coordinated military manoeuvres, such militias have taken to using social media (encrypted messaging) to publicise independent political-strategic objectives through widely accessible online discourse (Knights, et al., 2021b), and more importantly, propagate evidence of their pursuits through their own network of proxies (Badawi, 2021; Knights et al., 2020). In a sense, the muqawama has created a sequence of second-order paramilitary surrogates to act as proxies, some (allegedly) fictitious, and some authentic, to achieve its objectives.

Looking at this issue more broadly, the study of proxy relationships, such as those seen between Iran and its paramilitaries, has been a time-tested tradition in the field of international relations, especially in the context(s) of terrorism, radicalism, and insurgency (Byman et al., 2001; Byman, 2005). Moreover, recent conflicts, such as those in Syria, Ukraine, Yemen, and

Iraq, have further galvanised research in the field, but with a primary focus on how state actors instrumentally exploit proxy groups (Hughes, 2012; Mumford, 2013; Piazza, 2017). To some extent, the classification of these groups as something beyond a mere tool of the state has remained a contentious issue (Staniland, 2015; Moghadam & Wyss, 2020). Nonetheless, a growing body of literature has sought to change the scope of conventional analysis and specifically investigate intergroup (militia) relationships (Bakke et al., 2012; Bacon, 2018; Conrad et al., 2021). This examination of intergroup dynamics has taken various forms, with one point of interest being the behaviour of militia, terrorist, and extremist groups in the dimension of online media (Hughes & Meleagrou-Hitchens, 2017; Shehabat et al., 2017). The present study seeks to take this examination further, while also accounting for literature on strategic credit-claiming (Abrahms & Conrad, 2017; Rapoport, 1997; Kearns, 2021), to investigate the use of second-order proxies on social media for the purpose of false blaming.

In line with the above, the case study of Iran-backed paramilitaries operating in Iraq is well suited as a first step in exploring this topic. This is on account of the success such groups have had in growing an online following on platforms like Telegram (Knights et al., 2021a), and their ability to disseminate propaganda, extremist rhetoric, and audio-visual demonstrations of graphic violence to a large public audience (Berger & Morgan, 2015; Ward, 2020). A key consideration here is the rhetoric used in the online discourse of paramilitaries like those belonging to the muqawama. Statements seen on social media outlets are politically motivated, and often go beyond merely expressing violence as they seek to advance religious, nationalist, ethnic, or ideological causes (Etaywe & Zappavigna, 2021). These expressions work to demonstrate the *authenticity* of paramilitaries, as they carry original underpinnings of sociocultural meaning, and can serve as an indication of authentic identity in their repetition on

mediums like Telegram (Du Bois, 2007; Etaywe & Zappavigna, 2021, pp. 4-5). Nonetheless, identifying the *authenticity* of a militia group using its published content is not an easy task.

Most investigative attempts to date on the *authenticity* of Iraq-based paramilitaries have relied on rare examples of uncensored political discourse, overt instances of power grabs or internal strife, and infrequent media-related demonstrations of allegiance (Al-Hura, 2020; Knights, 2020; Knights, 2021; Mansour, 2021). As a result of this, an accurate taxonomy of the Iran-backed paramilitary network in Iraq has remained mostly elusive, even with predictions on which groups are likely to be fictitious (Knights et al., 2021a, pp. 5-6). This is problematic, as without a clear view of actor dynamics within this network, the threat it truly poses also remains uncertain. The societal risk inherent to such uncertainty lies in said groups' extensive popular legitimacy, evident capacity to pursue social and political agendas through violence, and participation in Iraqi state institutions like the national parliament (Malik, 2021; Thurber, 2014). Moreover, the fact that Telegram allows muqawama groups to strategically exploit proxies can further entrench these groups in the Iraqi political landscape, and create a dilemma of attribution as the muqawama uses its proxies to escape conviction and consequence.

To some degree, this can be viewed as a transformation of the notion of strategic credit-claiming (false blaming) mentioned earlier. By capitalising on the fact that the *authenticity* of new-wave surrogates is difficult to immediately establish, the muqawama is able to use the flexibility of online platforms like Telegram, and the active paramilitary environment in Iraq, to conceal their violent endeavours. Given the accessibility of encrypted messaging services, and their increasing use by VNSAs (Conway et al., 2019, pp. 1-24; Weimann, 2004), a better theoretical and empirical understanding of this transformed credit-claiming approach is essential. Conducting research on the evolution of claiming decisions can have implications for academic

scholarship, governmental policy, and public opinion. Developing a model on proxy-based credit-claiming can help with academic data collection initiatives, particularly when the objective is to establish a formal group's level of activity and distinguish it from a potentially (in)authentic surrogate. With this baseline, incidents can more effectively be traced to their true perpetrators, which can assist the efforts of relevant security forces. Additionally, a more precise tool for tracing attacks can help erase some of the public uncertainty inherent to the activities of paramilitaries, and may also dampen the effect of paramilitary propaganda on impressionable audiences.

Taking the above into account, the present study will seek to investigate the *authenticity* of the new wave of Iran-backed paramilitary groups in Iraq using a methodology involving a similarity-based textual analysis of Telegram content, and develop a machine learning model that can determine the degree of a surrogate's authenticity in comparison to an established criteria. This is a somewhat novel approach to the study of such militias, as the relationship between non-state (militia) sponsors and proxies is still a growing field (Moghadam, & Wyss, 2020). Despite this, much research has been conducted on how non-state actors utilise Telegram (Shehabat et al., 2017; Ward, 2020; Zenn, 2020), or strategically claim credit (Abrahms & Conrad, 2017; Kearns, 2021). This project will engage with this literature in the next section, and subsequently expand on traditional approaches to the study of proxy (militia) relationships, and options for strategic credit-claiming, through the case of Iran-backed Iraqi paramilitaries. Next, the essential features of the methodology and theoretical framework behind this study will be elaborated upon. Finally, this study will conclude with a discussion of the results and limitations to the proposed model, policy implications that may stem from these findings, and available avenues for future research.

2. Definitions: Militias and their Relationships

Before proceeding to the literature review, this study will briefly outline the actors under investigation and the hierarchy within which they operate. As mapped out in the introduction, although the VNSAs operating in Iraq belong to the broader categorization of militias, they are also known as paramilitaries, and more importantly, proxies of the Islamic Republic of Iran. Beyond this multitude of titles, Iran-backed paramilitaries have also taken to wearing the hat of 'sponsor' to their own set of (proxy) surrogate groups, a key feature of their false blaming strategy. For the sake of clarity, and to assist in the later exploration of literature on strategic credit-claiming, VNSA usage of social media, and representation through online discourse, as well as the analysis proper, this section will clearly define militias and their subcategories before continuing to an exploration of how VNSAs, like the muqawama, can take on functional proxy sponsorship roles. Overall, the intention here is to ascertain what is meant by an *authentic* paramilitary group, and subsequently, examine how non-state actors such as paramilitary militias can construct and exploit their own proxy networks with the aim of realising political objectives.

2.1: Defining Militias and their Subcategories

Conspicuously, despite their continued importance in multilateral conflicts, militias have regularly been insufficiently defined in the study of armed groups (Thurber, 2014, p. 900). Attempts to define militias often focus on the presence of a multitude of key characteristics: irregular armed force, operating within a failed or failing state, representing ethnic/religious/tribal/clan or other communal groups, no formal military training, skilled unconventional fighters, and operating autonomously or in service of a state (Shultz et al., 2004). Noticeably, several of these characteristics are contradictory, and this speaks to the broad

spectrum of attributes and characteristics contemporary militias possess. To some extent, the term 'militia' is a convenient placeholder until more appropriate terms can be found. With this in mind, scholarship has taken to focusing on subcategories of militias, namely: paramilitaries, foreign proxies, and warlords (Jackson, 2003; Bruce, 2004; Ahram, 2011).

Of interest to the greater study, is the potential role of Iranian militias as paramilitaries and foreign proxies.[6] For one, said actors can be viewed as paramilitaries: a subset of non-state actors that resort to violence in the service of a state (Thurber, 2014). A notable reason paramilitaries are a popular arrangement for states is their potential to allow for brutal repression and violence without the state-accountability inherent to any conducted abuses (Ahram, 2011). This is especially useful when operating on foreign territory, as is the case with Iran directing militias in Iraq, Syria, Palestine, and Lebanon. Furthermore, paramilitaries have been seen to operate under a combination of ideological, psychological, and material motivations (Bruce, 2001; Groh, 2019; Hughes, 2012, p. 11). On the one hand, paramilitaries have shown a tendency to form as a response to violence by insurgents (Bruce, 2001), and on the other, some have formed after receiving economic support (Mumford, 2013, p. 11). Not unlike mercenaries, such paramilitaries can either receive direct material benefits from state sponsors or profit indirectly as a result of the authority garnered in their areas of operation (Rozema, 2008). Historically, such economic motives have made paramilitaries hard to control, with violence relapsing even after the threat of insurgency has subsided (Bruce, 2004).

To a large extent, this phenomenon can be seen in the case of Iran-backed paramilitaries and their additional role as foreign proxies. A proxy can be defined as an actor that serves as the strategic agent of a sponsor (here, state) for tangible benefits (Salehyan, 2010, p. 503; Hughes,

[6] Although the topic of warlordism has been heavily explored in militia literature (Jackson, 2003; MacKinlay, 2000), it does not suitably explain the arrangement of Iran-backed militias.

2012, p. 11). However, such actors are not limited to operating under local regimes as they can also serve foreign governments (Thurber, 2014, p. 904). In this formulation, paramilitaries may operate as an ally of the regime and target a common adversary of both the local and foreign states. Alternatively, paramilitaries can take a more ambiguous position, and balance between the offer of support and threat of insurgency, with the determinant being whether the local regime is acting in accordance with the desires of the foreign state.

Iran has formed such proxy linkages with armed groups across the MENA region, with notable actors including Hamas, Hezbollah, and the numerous Iraqi Shia groups discussed in this paper (Forrest, 2009; Thurber, 2014). Nevertheless, it is necessary to keep in mind that the muqawama groups investigated in this study not only exert pressure on the Iraqi government in their role as Iranian proxies, but also pursue their own strategic interests that go beyond their obligations to Iran or the Prime Minister of Iraq (Knights et al., 2021b; Nada & Rowan, 2021, pp. 1-2). As was conceptualised earlier, this study perceives *authentic* paramilitaries to be able to autonomously represent themselves through relatable rhetoric, competent action, and attitudinal evaluations of their relationships (Ryan & Deci, 2000; Tracy & Robles, 2013, pp. 22-23; Van Leeuwen, 2001). Accordingly, having the capacity to pursue an independent agenda detached from the interests of a sponsor, is a distinct feature of the Iraqi paramilitary identity, and a factor behind their *authenticity.*

Remarkably, considerable literature has been attributed to studying the relationship between Iran and its proxies (Knights et al., 2021), however little has been done in relation to the role of state proxies, in the Iranian case and more generally, as sponsors with their own strategic objectives (Byman et al., 2001; Mumford, 2013). To be clear, ‘sponsors’ have been consistently conceptualised along state-centric definitions, with a sponsor-proxy relationship being analogous

to a state-VNSA relationship, despite examples showing a contrary reality of VNSA sponsorship arrangements (Moghadam & Wyss, 2020). The following subsection will address this topic in more detail, and demonstrate its applicability to the case of Iran-backed paramilitaries.

2.2: Non-state proxy relationships

Positioning proxies as VNSAs (i.e. paramilitaries) that operate under the sponsorship of the state has been a conventional practice dating back to the Cold War (Bar-Siman-Tov, 1984; Dunér, 1981, pp. 350-360; Mumford, 2013). Although this state-centric perception may have been apt at the time, it has obscured the reality of present-day proxy relationships. There has been a growing trend of ideologically and geographically diverse non-state actors adopting sponsorship roles that are analogous to those customarily held by states. Notable examples include militias and groups such as Hezbollah, the National Patriotic Front of Liberia, and the People's Protection Units (Moghadam & Wyss, 2020, p. 120). Such groups have employed their sponsorship role, in combination with other tactics, to embed themselves in their respective areas of operation in a manner similar to Iran-backed paramilitaries. This subsection will expand on the rise of VNSAs as sponsors and indicate how such actors may differ from states in their use of proxies. The aim here is not necessarily to illustrate how state sponsors and VNSA sponsors differ in absolute terms, but to establish why VNSA sponsorship arrangements are distinct. This inference can be then used to lay the groundwork for a discussion in the next subsection on the transformed proxy-based claiming strategy employed by the muqawama.

Overall, a shared feature of VNSA and state sponsorship arrangements is the severe and violent repercussions such proxy relationships tend to generate (Cunningham et al., 2013, pp. 526-528). The provision of external support to belligerents, either by a state or by a VNSA, can

escalate civil wars, insurgencies, and other forms of political violence to higher levels of lethality while also risking their internationalisation (Schultz, 2010). At the core of this escalation, is the increased number of violent belligerents that sponsorship arrangements add to a conflict theatre, and the challenges brought about by the indirect warfare tactics inherent to proxy-use (Clarke, 2019; Hughes, 2016). A clear example of this is the attribution of violent activities to fictitious proxies by the muqawama. In attempting to interpret such a proxy-relationship, conventional (state-centric) insights would dictate that sponsors employ proxies in a cost-effective manner, and that proxies are an indirect and predominantly military tool, even when used to achieve political objectives (Moghadam, & Wyss, 2020, p. 121). Taking this approach in relation to Iraqi paramilitaries can be problematic as its view of proxies as a tool primarily aimed at achieving military-centric outcomes, rather than as a political instrument, may ignore the true value of proxies to Iraqi paramilitaries.

Unlike states, VNSA sponsors use proxies as ‘political assets’(Moghadam, & Wyss, 2020, pp. 128-132). In such a formulation, proxies are used to advance relatively specific aims of non-state sponsors, which are often based around consolidating political power and legitimacy (Podder, 2017; Suchman, 1995). In practice, this means that proxies in the hands of VNSAs are more than just a blunt weapon to be pointed at enemies, with many proxies helping/being used to win over local communities and achieve social and political legitimacy (popularity and support) (Grynkewich, 2008). The military benefits provided by proxies are certainly appreciated, but VNSA sponsors base part of their legitimacy on their ability as fighters, and thus are hesitant to delegate their offensive combat operations to proxies (Malet, 2013; Moghadam, 2008). Although most non-state sponsors face additional constraints in terms material capabilities, and governance deficits (Weinstein, 2006), the same does not apply to the muqawama.

The muqawama, as a formal Iraqi paramilitary faction, is a high-capacity sponsor. Given its integration into the Iraqi security apparatus through its membership in the PMF (Mansour, 2018; Mansour, 2021), this faction has substantial financial and military means at its disposal. Nevertheless, the use of proxies as a political asset is a key feature of how it operates in Iraq. Beyond its role under the PMF organisational umbrella, ambitious paramilitaries within the muqawama have also sought to enter the Iraqi electoral system (Nada & Rowan, 2021). This is a tricky manoeuvre when considering their violent rejection of US and Turkish presence in Iraq (Ezzeddine & van Veen, 2021). Most non-state sponsors welcome the chance to provoke what they have termed as an 'occupier' to overreaction (Crenshaw, 1981), and often attempt to goad international powers into conflicts that may enhance their legitimacy and appeal (Kydd & Walter, 2006; Neumann & Smith, 2007). However, despite this urge, the Iraqi government's functional relationship with the US forces a degree of subterfuge to muqawama operations. In essence, to remain legitimate participants in the Iraqi political system, muqawama groups employ proxies to conceal their operations against entities such as the US (Knights et al., 2021a). More significantly, fictitious surrogates are used as a way to avoid public repercussions and criminal conviction while maintaining political legitimacy.

3. Literature Review

Taking the above insights on militia identification and the structural logic behind VNSA-proxy sponsorship arrangements, core concepts and ideas underlying this study can now be explored in a review of relevant literature and scholarship. Overall, in attempting to answer the proposed RQ of this study, it is essential to examine scholarly work on how VNSAs strategically approach credit-claiming, the significance of their discursive practices, their evolved

online presence, and what this all means in terms of *authenticity*. To this end, this section will first start by describing the particularities of strategic credit-claiming, and the various ways paramilitary actors can falsely attribute credit for an attack, or more generally, lie about its perpetrator. Subsequently, this study will explore the significance of online discourse, particularly on social media, as a mechanism for large-scale coordination, securing public legitimacy, and *authentic* representation. Finally, this section will end with an inspection of literature on the role played by social media platforms in the operations of VNSAs, with a specific focus on the Telegram application and its encryption-based popularity.

3.1: Strategic credit-claiming

Credit-claiming, as a practice, has often been studied in the dimension of terrorism (Crenshaw, 1981; Fromkin, 1975). Nevertheless, by focusing on the performative nature of such acts, and their objective of gaining a response from an audience and progressing one's goals (Jenkins, 1974), their rationale can also be applied to militia behaviour. This is especially relevant in the case of Iraqi paramilitaries who have been designated as terrorists on multiple occasions.[7] Customarily, rationalist perspectives suggest that terrorism is a strategic act that seeks to communicate a message or express a grievance (Jenkins, 1974). Groups that commit (violent) terrorist acts are then assumed to be rational actors who utilise costly signaling to achieve their objectives (Kydd & Walter, 2006; Lake, 2002). Furthermore, groups are also expected to claim credit for their attack(s), as violence without explanation is considered to be a poor form of communication (Kearns, 2021), especially since the target will be unable to know, or redress, the group's grievances (Abrahms & Conrad, 2017).

[7] Iraqi paramilitary groups have targeted US forces on numerous occasions. This has motivated the US to designate several groups as terror organisations while also labeling key leaders/figures behind these groups as terrorists (Nada & Rowan, 2021).

Although rationalist perspectives on credit-claiming have dominated the field for a significant period of time (Lake, 2002; Pluchinsky, 1997), more recent scholarship has established that the majority of (terrorist) attacks go unclaimed (LaFree & Dugan, 2014). While there are indeed many reasons to claim an attack, including publicity, intimidation of adversaries, and sending a political message (Kearns et al., 2014), it can also be rather detrimental. Groups that use violence akin to terrorism hold survival as a primary objective (Crenshaw, 2001). Consequently, such groups must balance the need to gain supporters through performative feats of violence with their concerns about possible backlash from the general populace (Lake, 2002; Pluchinsky, 1997). Notably, a similar concern is expressed by VNSA sponsors, and has been demonstrated as a motivator behind the use of proxies. In essence, VNSAs are only likely to claim credit when the expected political return from the exercise is positive (Abrahms & Conrad, 2017, p. 281). If the anticipated return from claiming credit is negative, which often occurs when civilian targets are harmed (Findley & Young, 2007), an attack is likely to go unclaimed lest it undermine the political ambitions of the VNSA.

Furthermore, beyond not taking credit for committed violence, VNSAs have also been known to issue concrete threats of violence that they do not end up carrying out (Brown, 2020; Mroszczyk & Abrahms, 2021). This specific mode of behaviour is beyond the scope of this study of Iran-backed paramilitaries, but it does match the rationale of strategic credit-claiming. Violent acts (e.g. terrorism) are to a large degree a political communication strategy expressed by their perpetrator (Pape, 2008), and as such, these acts must balance between generating fear and attracting sympathy (Crenshaw, 1981). Threats can limit the detriments associated with an actual, claimed, attack by manipulating audience perceptions, inducing fear, and generating desired concessions (Brown, 2020), but without the negative political return resulting from casualties.

Moreover, a threat that is a precursor to a real incident, or a threat that has generated violence after noncompliance, has the added benefit of legitimising a VNSAs ability and local authority Kydd & Walter, 2006). In essence, VNSAs can derive benefits and concessions, intimidate civilians, and sharpen the impact of future violence by substituting a genuine act for a threat of its incidence. Altogether, this strategic theory on the value of threatened violence is quite applicable to groups that place violence, fear, and the ability to cause harm, as key objectives (Abrahms, 2013; O'Shaughnessy & Baines, 2009), yet it stumbles when applied to restrained VNSAs looking to display a more moderate face.

Taking into account the urge of some politically motivated and socially active VNSAs to show moderation, credit-claiming can be considered from an alternate perspective, one oriented around false attribution. Holding the rationalist assumption that actors who claim an attack are the ones who committed it (Lake, 2002), the discernible exercise of credit-claiming can be expanded to pose a different question: why do some groups falsely claim credit for attacks/terrorism? Or, why do groups lie about terrorism? The frequency of false claims is certainly difficult to establish, but the practice is not uncommon (Kearns et al., 2014, pp. 423-424). When a group falsely claims credit, it is normally an attempt to project strength and attract attention that it would normally be unable to do as a result of capacity constraints (Kearns, 2021).

Alternatively, VNSAs can falsely attribute credit through the means, and with the intent, of false claiming, false blaming, or lies of omission (Kearns et al., 2014, p. 423). For one, a group can falsely attribute credit for (or blame) an attack it committed to a rival third-party or fictitious group of its invention, an act that has been termed “false flag terrorism” (Kearns et al., 2014, pp. 425). Other options for false credit-claiming also include the “hot potato problem”

which involves a group blaming an attack that it did not commit on an adversary, or the more ambiguous choice to lie through omission and commit an attack without taking credit or blaming it on a third-party (Kearns et al., 2014, pp. 424-426). Reasons that may be given for any of the above (false) credit-claiming strategies will inevitably rely on context, but referring back to the ongoing legitimacy deficit faced by VNSAs (Podder, 2017; Suchman, 1995; Weinstein, 2006), and the persistent need to balance fear and sympathy (Crenshaw, 1981), broader explanations can be presupposed.

Overall, (extremist) violence often backfires (Mroszczyk & Abrahms, 2021, p. 426). Overt violence, particularly against civilians, can strengthen the resolve of local authorities (Berrebi & Klor, 2008), lower the odds of governmental concessions (Abrahms & Gottfried, 2016; Gaibulloev & Sandler, 2009), erode popular support (Crenshaw, 2007), and potentially expedite organisational demise (Cronin, 2009). With this in mind, VNSAs, particularly those with long-term political ambitions, seek to present a more moderate front and will hence use the above strategies to deny (their) organisational involvement (Abrahms & Conrad, 2017). Moreover, as a partly learned behaviour, VNSAs have been shown to claim less credit over time, with results strengthening for groups that had survived more than a decade (Abrahms et al., 2018).

This moderate form of branding can also extend to denying principal intent, where the leadership of a VNSA will acknowledge the committed violence, specifically in the case of civilian casualties, but will deny that this action reflects its true intentions (Abrahms, 2018). Public image is a near constant consideration for VNSAs, and as a result, making moves to distance the core organisational image away from violent offences is an expected tendency. In practice, this can manifest in the form of image restoration techniques like apologies or

scapegoating to distance a VNSA from the offence (Matesan & Berger, 2017; Mroszczyk & Abrahms, 2021, p. 427). Predictably, muqawama groups have experienced internal differences which have necessitated denial of principal intent (Knights et al., 2021a), however due to scope limitations and an alternate research focus, this study will primarily focus on false attribution as part of strategic credit-claiming.

All in all, with account to the value of threats of violence, general denial of involvement, and denial of principal intent, this study finds "false flag terrorism" to be a rather apt view of how the muqawama operates. In simple terms, fictitious proxies are created for the purpose of taking credit for an attack or violence committed by muqawama groups seeking to protect their public image. To further investigate how such false flag terrorism can be tackled/measured, especially in the methodology of this study, the following sub-section will compile relevant theoretical insights on how online content, specifically on social media platforms, can assist in determining the *authenticity* of new-wave paramilitary surrogates in Iraq.

2.4: Online Discourse and Authenticity

Online information environments including forums, audio-visual distribution platforms, and especially social media networks have grown exceedingly popular in the eyes of VNSAs (Conway et al., 2019; Mroszczyk & Abrahms, 2021; Prucha, 2016). A distinct draw of social media networks is their capacity to concretely broadcast a VNSAs worldview through easily accessible online content (Prucha, 2016). A prominent example of this can be seen in the drive of Jihadist's to spread theological writings and statements on online mediums under the premise of identity-building, with notable cases including Al-Qaeda or IS (Bloom et al., 2019; Lohlker, 2016). Modern audio-visual communication on social media has satisfied this goal by illustrating

who a group is, what it is fighting for, and whom it is fighting against (Prucha, 2016, p. 48). To a large degree, this online discourse is a manner by which VNSAs express *authenticity*. It should be stressed that although similarities can be seen in the discourses of analogous VNSAs, which tends to lead to lump categorisations of 'Islamist' or 'Jihadist' (Lesser et al., 1999; Ranstorp, 1996), no single narrative exists amongst VNSAs (Prucha, 2016, pp. 48-50). Each group enforces a coherent worldview which advances religious, nationalist, ethnic, or ideological causes (Etaywe & Zappavigna, 2021). This coherent expression produces an *authentic* representation of a VNSAs identity, as its repetition and demonstration on social media platforms (like Telegram) reinforces and embeds it in the minds of the audience (Du Bois, 2007).

When applying this principle to genuine and fictitious proxies of the muqawama, a distinction can be made in terms of their online discursive expressions. Although surrogates of VNSA sponsors operate as political ancillaries (Moghadam, & Wyss, 2020, pp. 128-132), these groups are known to carry their own ambitious, local identities, and attitudinal evaluations (Schlichte & Schneckener, 2015). Such features provide surrogates with local legitimacy, and are often the reason VNSA sponsors take such groups under their wing (Grynkewich, 2008; Podder, 2017; Suchman, 1995). With this in mind, it can be inferred that genuine proxy surrogates, like their VNSA sponsors, will express an *authentic* worldview in their online communications. A fictitious proxy on the other hand, will only offer a duplication of the worldview conveyed by its 'sponsor', essentially plagiarising its content.

Without a careful and large-scale analysis, this distinction between fictitious and genuine discourse can be difficult to trace, but the accessibility of Iraqi paramilitary content on Telegram can make such an investigation into surrogate authenticity possible. Looking at the Iraqi paramilitary Telegram environment, false flag terrorism is not easy to perceive, as an external

overview shows a conventional VNSA sponsor-proxy arrangement. Militias are observed to engage in a form of 'outbidding' behaviour,[8] seen through their demonstrative violence against 'enemies' such as US-coalition military forces (Knights, Malik, & Smith, 2021a; Seligman, 2021). The goal of this violence seems to be a classic attempt at portraying competence in the eyes of the public, an ambition which is in line with VNSA attempts at gaining local legitimacy (Bloom, 2005; Kydd & Walter, 2006). As groups attempt to captivate audiences with the spectacle of their activities, more severe forms of violence are performed and posted on Telegram (Conrad and Greene 2015; Ward, 2020). To some degree, this appears counterproductive to militia longevity (Staniland, 2012), but Iran-backed paramilitary groups (particularly the muqawama) have innovated and employed this performative violence on their channels to motivate members and deceive adversaries (Gaibulloev, Hou, & Sandler 2020; Kenny, 2007). Essentially, by conducting false flag terrorism in an environment with genuine surrogates, muqawama groups conceal their inauthentic proxies behind a curtain of real surrogates. The next subsection will elaborate on how Telegram provides the toolkit necessary towards virtual VNSA operations, and specifically the mode of credit-claiming promised by false flag terrorism.

2.5: Militias on Telegram

Telegram, like other encrypted messaging services, has become a tool for militias, extremist, and insurgent groups to share information, reinforce messages, and plan activities (Walther & McCoy, 2021). Largely, previous notions on how such groups use social media for the single-minded purpose of planning attacks has transitioned into a perception of VNSAs as

[8] Although the theory of outbidding, which posits that groups would try to 'one up' each other in terms of their (violent) attack(s) does provide a logical explanation for the demonstrative violence, it does not predict a deceptive performance.

virtual entrepreneurs who promote objectives beyond violent activities (Hughes & Meleagrou-Hitchens, 2017). Principally, the Telegram militia landscape can be framed as an abstract socio-sphere that exceeds traditional boundaries (Habermas, & Burger, 1989; Shehabat et al., 2017, pp. 28-32). Through this sphere, VNSAs are capable of acting stigmerigically, meaning in a coordinated manner but without constant communication or control. Telegram is key to this mode of functioning as it offers a unique blend of privacy and accessibility, where users can interact in secure [private] one-to-one or [public] one-to-many communication channels (Prucha, 2016; Shehabat et al., 2017).

Notably, one-to-one communication on Telegram is encrypted, essentially allowing for secret chats between users. This ensures that only parties to a conversation can access the exchange. Such encryption is a relatively common feature of cloud-messaging services, but Telegram takes security a step further by allowing single users in an exchange the chance to permanently delete all shared content for all participants (Hamburger, 2014). This is a rather convenient feature, as a sense of practical anonymity is constructed within VNSA virtual networks. Conversations can be easily started and ended without lasting evidence. Additionally, given the multiple chains of communication, a sort of segmentation can be created between one-to-one interactions and one-to-many announcements released by central information hubs. These are often seen in the communication(s) of major groups like IS, Hamas, Al-Qaeda, Hezbollah, or the muqawama[9] (Bloom et al., 2019; Knights et al., 2021a; Prucha, 2016, p. 51). In essence, Telegram allows VNSAs to operate separately while also existing on a group-community level (Palasinski & Bowman-Grieve, 2017).

[9] These information hubs tend to be the Telegram channels of muqawama groups, or propaganda-news channels owned by muqawama paramilitaries. A prominent example of the latter is the extremely popular channel going by the name of “Sabereen News” (Knights et al., 2021d).

Adding to this internal anonymity which works to limit options for effective infiltration (Rogers, 2003), Telegram is particularly attractive to VNSAs as a result of its low barriers to entry. The messaging application allows for free cross-platform communication, and the exchange of data (video, audio, text) in a secure environment (Bloom et al., 2019). Moreover, users do not have to rely on external links or platforms as all data can be securely downloaded to, and accessed on, the app itself (Prucha, 2016). Furthermore, accounts can be created and terminated instantly, which can complicate attempts to trace the full size or activity of a VNSA network. This makes strategic credit-claiming, specifically false flag terrorism, rather easy on Telegram, as ease of account creation, instantaneous sharing of audio-visual evidence of violence, user anonymity, and encryption of private content in addition to the freedom to delete all evidence when necessary, are conducive features to the creation of fictitious and temporary online identities.

In essence, Telegram provides VNSAs with the essential tools necessary towards creating a multipurpose digital environment that latently connects surrogates to widely-accessible central hubs (Shehabat & Mitew, 2018, p. 84), while also concealing the *authenticity* of surrogates, thus ensuring a robust, secretive, and stigmergic virtual operation. Significantly, hubs are key to the operation, as they work to share coherent and centralised propaganda messages while also (indirectly) coordinating activities and establishing one-to-one connections between (in)authentic surrogates. The following section will compile relevant theoretical insights on how proxies can be exploited on Telegram, and provide a compact theoretical explanation detailing how online content on the platform can assist in determining the *authenticity* of new-wave paramilitary surrogates in Iraq.

4. Theoretical Framework

To empirically examine the identity of new-wave surrogates operating in Iraq, *authenticity,* the key concept investigated by the study*,* must be manifestly situated within a broader framework of theories on representation, proxy sponsorship, credit-claiming, identity-building, and online discourse. This section will echo the main concepts and ideas found in the literature review that allow for a theoretical lens through which to view how Iran-backed paramilitaries in Iraq organise their sponsorship arrangements, and use surrogates as proxies with the purpose of indirect credit-claiming.

This study has defined authentic paramilitaries as those are able to autonomously represent themselves through relatable rhetoric, competent action, and attitudinal evaluations of their relationships (Ryan & Deci, 2000; Tracy & Robles, 2013, pp. 22-23; Van Leeuwen, 2001). In practice, this means that to be *authentic* in the Iraqi paramilitary landscape, a surrogate must have the capacity to pursue an independent agenda detached from the interests of a sponsor. This fits prevailing theories on how non-state sponsors employ proxies, as the local legitimacy and operational capacity such surrogate actors carry is a significant draw to non-state sponsors (Grynkewich, 2008). VNSA sponsors face a constant legitimacy deficit that pushes them to use surrogates to consolidate political power, authority, and political-social legitimacy in their areas of operation (Podder, 2017; Suchman, 1995), essentially making proxies a valuable political asset (Moghadam, & Wyss, 2020, pp. 128-132). Keeping this in mind, VNSAs must then utilise proxies to balance their need to gain supporters through performative feats of violence with the possible backlash such violence may incur from the general populace (Lake, 2002; Pluchinsky, 1997).

Principally, the balancing act VNSAs face when it comes to conducting violent operations in the public eye presents a credit-claiming dilemma. On the one hand, taking credit for an act of violence allows groups to gain publicity, send a political message, or intimidate rivals (Kearns et al., 2014), on the other, it poses a risk to their primary objective of survival and influence (Crenshaw, 2001; Mroszczyk & Abrahms, 2021). This means that VNSAs are only likely to claim credit when the expected political return from the exercise is positive (Abrahms & Conrad, 2017, p. 281). If the anticipated return from claiming credit is negative, which often occurs when civilian targets are harmed (Findley & Young, 2007), an attack is likely to go unclaimed lest it undermine the political ambitions and long-term survival of the VNSA. This often leads to a strategic attitude to credit-claiming (Kearns, 2021). As observed in the behaviour of muqawama groups and their surrogates (Elias, 2020), this can manifest as the false attribution of attacks to a rival third-party or fictitious (*inauthentic*) group, namely being the process 'false flag terrorism' (Kearns et al., 2014, p. 425).

Altogether, false flag terrorism is a rather apt view of how the muqawama operates in Iraq. Simply put, fictitious proxies are created for the purpose of taking credit for an attack or violence committed by a muqawama group. This false attribution masks responsible paramilitary groups who are seeking to protect their public image, relations with Iraqi state institutions, and political ambitions. Telegram is pivotal to this process as ease of account creation, instantaneous sharing of audio-visual evidence of violence (Prucha, 2016), user anonymity, and encryption of private content (Bloom et al., 2019), in addition to the freedom to delete all evidence when necessary (Hamburger, 2014), are all conducive features to the creation of fictitious and temporary online identities. Moreover, the fact that Telegram provides paramilitaries with the crucial tools necessary towards the creation of a multipurpose digital environment, which

latently connects surrogates to widely-accessible information central hubs (Shehabat & Mitew, 2018, p. 84), is a major boon to false flag terrorism. Using such hubs, muqawama sponsors are able to easily conceal the *authenticity* of surrogates using Telegram's built-in features, thus ensuring a robust, secretive and stigmergic virtual operation.

In addition to their virtual boons, social media hubs are a vital communication asset for VNSAs. Information hubs are seen to share coherent and centralised propaganda messages while also (indirectly) coordinating activities and establishing one-to-one connections between *inauthentic* surrogates (Shehabat & Mitew, 2018, p. 84). More significantly, these media channels gave paramilitaries the power to concretely broadcast their worldview through easily accessible online content (Prucha, 2016). To a large degree, it can be argued that this online discourse is a manner by which paramilitaries express *authenticity*. Each group enforces a consistent worldview which carries original underpinnings of sociocultural meaning and advances religious, nationalist, ethnic, or ideological causes (Etaywe & Zappavigna, 2021). This coherent expression then serves to produce an *authentic* representation of a paramilitary's identity, as its repetition and demonstration on platforms like Telegram reinforces and embeds it in the minds of the audience. With this in mind, it can be inferred that proxy surrogates, like their VNSA sponsors, will express an *authentic* worldview in their online communications. A fictitious proxy on the other hand, will only offer a duplication of the worldview conveyed by its 'sponsor'.

Taking the above into account, it can be surmised that content published by a group on Telegram can be an indication of its *authenticity.* Moreover, when comparing discursive Telegram content published by muqawama sponsors and their surrogates, *authenticity* can be used to establish whether a new-wave surrogate is a fictitious profile constructed for the purpose

of false flag terrorism, or potentially a genuine surrogate with its own political interests and agenda that deviates away from primary muqawama interests. In a practical sense, *authenticity* is being applied with the presumption that duplication of discourse is likely to indicate that a surrogate group is *inauthentic* or fictitious. This study does not assume a binary classification, but predicts a classification that places proxy surrogates on a spectrum of *authenticity,* with placement corresponding to the degree of Telegram discourse similarity between a surrogate and a muqawama sponsor. Notably, this model does not assume to predict *authenticity* in absolute terms. The aim here is to construct a simple model that can be a first step in the classification of current new-wave proxies, and any that might appear in the future in Iraq, or other regions experiencing a similar sponsorship and credit-claiming phenomenon. In summary, this study predicts that increased Telegram-content dissimilarity from a muqawama group would indicate higher levels of *authenticity*, while similarity may allude to the fact that suspect groups are fictitious, *inauthentic*, surrogates. The next section will elaborate on the methodology this study will use to determine which surrogate groups may be *authentic*.

Table 1: List of investigated Iran-backed paramilitary groups

Militia Category	**Official Militia Name (translated)**		
Muqawama group	Kata'ib Hezbollah		
New-wave surrogates (known[10] to be fictitious)[11]	Ashab al-Kahf	Usbat al-Tha'ireen	Fariq Fatemiyoun al-Maidani
	Liwa Tha'ar al-Muhandis	Rab' Allah	Al-Wehda 10,000
	Qasim al-Jabarin	Quwat Thu al-Faqar	
Out of sample groups: New-wave surrogates: (unknown status)	Abu Jadaha	Al-Majame'e al-Khasa	Saraya al-Thawra Al-Eshreen al-Tahniya
	Saraya Tha'ar al-Shuhada	Liwa al-Shahid Ahmed Dar'am	Kata'ib Abul-Fadl al-Abbas

***All data found in this table has been collected by the author after conducting research on the political, social, and security environment in Iraq and the role paramilitaries play within it. Data was collected from political commentators and key figures on Twitter and Telegram.**

4: Methodology, Case Selection, and Data:

As described in earlier sections, the main concept under investigation in this study is the measure of paramilitary *authenticity*. The basis of this investigation is the RQ: what can published Telegram content reveal about the authenticity of the new wave of Iran-backed paramilitary groups in Iraq? Markedly, when referring to the above groups as paramilitaries, this study will reiterate its implementation of the conventional definition of the term, which is a subcategory of the broader categorisation of militias: a subset of non-state actors that resort to violence in the service of a state (Thurber, 2014). This definition has been further expanded to

[10] Certain paramilitary surrogates have been established to be fictitious by analysts and commentators working in the region. For a summarised breakdown of group affiliation(s) please see: Knights, Malik & Smith (2021a) or the series titled "Militia Spotlight: Profiles," published by the Washington Institute of Near East Policy (Knights, 2021).
[11] Early studies on the behaviour of new-wave surrogates assumed that these groups were semi-autonomous extremist cells that enjoyed indirect support from formal paramilitaries like the muqawama (Elias, 2020). Nevertheless, recent violent incidents and manoeuvres by the aforementioned surrogates and their sponsors has shown that they are more likely to be 'fake groups' that function as media façades and obscure the group truly responsible for an attack (Badawi, 2021; Knights, 2020; Knights et al., 2021a).

include proxies, who are actors that serve as the strategic (political) agent of a sponsor for tangible benefits (Salehyan, 2010; Hughes, 2012). Altogether, this project contends that paramilitary groups in Iraq are known Iranian proxies, but several well-established (muqawama) groups have taken to creating their own second-order proxies as a strategic method of credit-claiming (Kearns, 2021; Moghadam, & Wyss, 2020), and more specifically, false flag terrorism (Kearns et al., 2014). Some proxies are presumed to be genuine surrogates while others are likely to be fictitious entities that only exist on Telegram (Badawi, 2021; Elias, 2020; Malik, 2021).

All in all, taking above concepts and definitions into account, this project will seek to provide a prediction as to where the new wave of Iran-backed surrogates fit along a simplified spectrum of *authenticity.* This spectrum will be determined using an initial sample of paramilitary surrogates that are known to be fictitious[12] (see footnote for an explanation on this determination), and then used to train a machine learning algorithm that can classify out-of-sample groups. The fictitiousness of the initial sample was ascertained using data collected from external sources, and will act as a critical indicator of the validity of discourse similarity as a determinant of paramilitary *authenticity.* Significantly, as stated earlier, this study will not seek to verify in absolute terms whether a surrogate is fictitious or genuine, but alternatively aim to test the value of similarity as a determinant of *authenticity* and use it as a basis for a machine learning model. Notably, this does not allow for a final determination of whether a group is genuine or fictitious, but rather shows how distant an out-of-sample surrogate is from the sponsor, and where it stands in relation to other surrogates. This can be pointed to as a distinct issue behind the proposed methodology. Nevertheless, further steps can be taken to correct this issue, primarily in the form of an expanded model that accounts for the methodology

[12] See note(s) 10 and 11.

used in this study and builds upon it. This will be addressed in the next section as part of a discussion on research implications and study limitations. The next three subsections will elaborate on the sampling choices made by the study, and provide a complete breakdown of the tools and software used in this study.

4.1: Variables and Cases

To answer the RQ this project will operationalise *authenticity* as a continuum/spectrum along which the Telegram text-content of surrogates can be positioned. To be clear, placement on this spectrum corresponds to how distant or unalike a surrogate group's content is from the selected muqawama sponsor, with higher or lower percentages in text divergence corresponding to a surrogate's *authenticity*. To assess how a group's Telegram discourse can be positioned on this spectrum, the project will measure textual similarity (in terms of words) between new-wave channels who are known to be fictitious and a channel belonging to a muqawama group (KH). The decision to select KH as the muqawama reference point was based on the fact that it is the most prolific Iran-backed militia in Iraq (Knights et al., 2021a), and the operational reality in Iraq indicating that KH either directly controls the fictitious surrogates listed in **Table 1**, or coordinates their movements and operations (Knights, 2020; Mansour, 2021).

In principle, a surrogate with a higher similarity score is then theorised to be less unique, and subsequently, is less likely to be *authentic*. In a general analysis of surrogates with an unknown status, this could then allude to the fact that the surrogate was fabricated by a muqawama paramilitary group seeking to conduct false flag terrorism. However, before making such a determination, this study will first aim to understand how successful similarity scores are in determining the *authenticity* of surrogates. Consequently, this study will employ an initial

sample of surrogates that are known to be *inauthentic.* The reasoning here is that similarity must be first established as a reliable metric before being applied to test data (out-of-sample surrogates). By establishing the possible range of similarity scores, particularly max and min values, the explanatory value of similarity can be expanded to describe where out-of-sample groups can be placed in relation to specific fictitious surrogates. The sample of cases selected to test the study's prediction will include texts collected from public paramilitary Telegram channels.[13] The 15 channels[14] selected for this project are listed in **Table 1,** with paramilitaries being split into a *muqawama group* and *new-wave surrogates* to allow for a determination of *authenticity*. Markedly, and as noted above, the study has also separately established that several new-wave surrogates are fictitious, these are highlighted in red.[15] These groups will be used to determine an array of reference similarity scores, while also acting as a preliminary test for the study's prediction, namely that the online discourse of (*inauthentic*) fictitious surrogates is likely to be highly similar to that of their sponsor. Groups that are not highlighted in red will be classified later in the analysis. Overall, the paramilitary groups shown in **Table 1** were chosen after a careful investigation of the Iraqi paramilitary landscape. Choices were based on specific criteria: a link to the PMF, a pattern of attacking the US army and coalition forces stationed in Iraq, an indication of allegiance to Iran or a Iran-backed muqawama group, an active account on Telegram, and most importantly, evidence of being created after March 2020. The last point is

[13] All channel data used for this study is freely available and can be accessed using the built-in Telegram data-export tool. The only requirement is a personal Telegram account that can be used to join the various paramilitary channels that exist on the application.

[14] Although this project selected 15 channels, the actual number of new-wave surrogates that appeared in Iraq is closer to 30 (Badawi, 2021). The channels not included in this study either no longer existed on Telegram (due to deletion), or did not have enough text content at the time of analysis to be eligible for the sample. The missing data could be attributed to lack of text-based activity or the Telegram feature that allows for the complete deletion of content for all viewers by its publisher (Hamburger, 2014). However, this should not exclude them from future efforts as such groups have a tendency of re-activating when necessary (Knights, 2020), potentially indicating their singular purpose of participating in false flag terrorism operations.

[15] See note(s) 10 and 11.

quite crucial as this is when larger muqawama groups consolidated power and presumably started staging and publishing attacks using fictitious surrogates (Badawi, 2021).

4.2: Data Accessibility and Collection

As indicated in the previous sub-section, this paper will make use of a corpus of Telegram chat records produced by the 15 Iran-backed paramilitaries as a dataset, with appropriate divisions of this dataset being made where methodologically necessary. To have a full representation of channel activity, data used in this dataset includes all chats produced since a group's Telegram channel was first created.[16] Additionally, it should be reiterated that raw Telegram data is publicly accessible to anyone with a Telegram account, and does not include private conversations but rather the public discourse of surrogates. After joining a specified (public) surrogate channel, all chats can be freely viewed/downloaded.

Beyond accessibility, when using the term corpus, this study refers to an object which represents a table containing the full, raw text of the data-source (in this case, all the Telegram chats published by all groups) stored in one column, and the id of each entry stored in another, with further additions of metadata (e.g., author, year, source) being possible without restriction (Arnold et al., 2019, p. 10). To generate this corpus, the channels belonging to groups listed in **Table 1** will be scraped using a built-in data export tool on the Telegram app to collect all text-based content [chats and announcements] published by the groups since their formation. Given that the scope of this paper does not allow for image/video data, these files will be excluded. It should be mentioned that similar research has employed data from Twitter, and although this is a suitable source of data, many VNSAs similar to those operating in Iraq have

[16] The final data point (chat) was collected on the 6th of February. Most groups were formed in March/April 2020, which means that the dataset contains 23 months worth of chat records.

switched to Telegram due to the lower risk of account termination and increased security (Bloom et al., 2019, p. 1242; Prucha, 2016, p. 51). The popularity of Telegram has the added benefit of allowing for the collection of a large amount of up-to-date data as groups continue to actively post (violent) content on their Telegram channels. The study seeks to collect at least 50[17] observations (text messages) from every group. In summary, all data used by the study can be downloaded by any Telegram user (as long as the specified channel was not deleted), and the unit of observation in this study will be text-based Telegram content, while the unit of analysis will be Iraqi paramilitary surrogates and their muqawama sponsor.

4.3: Overview of the Conducted Method

This study will employ textual analysis, network analysis, and machine learning research methods within the R[18] programming environment to analyse and classify the Telegram content of Iran-backed paramilitary groups. A collection of interoperable libraries/packages will be used to conduct the analysis in the next section, these include: “quanteda”, “readtext”, “tidyverse”, “textplot”, “quanteda.textmodels”, “quanteda.textplots”, “quanteda.textstats”, “seededlda”, “igraph”, “stm”, “textreadr”, “class”, “stopwords”, “plyr”, “caret”. Overall, and with a recognition of the role played by the aforementioned R packages, the methodology behind this study can be outlined in several steps.

First, a corpus of all the published text content, grouped by paramilitary, will be produced and then used to create a document feature matrix (DFM). A DFM represents the frequencies of different features (in essence words) within documents (here Telegram chats) in the form of a

[17] This metric was chosen after conducting several tests. 50 observations appear to be a useful minimum for inferring authenticity, however this measure can be disputed when the method is applied to a different case-study.
[18] R is a user-friendly and high-level programming language geared towards the analysis of data (Ihaka & Gentleman, 1996).

matrix. Consequently, within this DFM, the words published by each group can be ranked in terms of frequency and visualised.[19] Here, to ensure analytical accuracy, the study will engage in pre-processing in order to remove redundant text (stop words, chat symbols, emojis, terms of a low/high frequency, etc.). Following this, the study will test the earlier prediction by computing the textual similarity between the sample *muqawama group* and *fictitious new-wave surrogates.*[20] This will allow the study to determine the validity of similarity as a measure of *authenticity.* Subsequently, groups will be placed along a spectrum of *authenticity*. As stated before, this placement is not aimed at being a pure classification, but rather a functional distinction that can take the analysis further. Generally, the study assumes that the arrangement of groups along this spectrum can allow for some form of binary division, with the likely outcome being a 50/50 split.

The binary division of fictitious surrogates based on their similarity to their muqawama sponsor will allow for the development of a machine learning model, and subsequently a structural topic model (STM). Specifically, the project will use the k-nearest neighbours (KNN) machine learning algorithm to classify groups whose *authenticity* has not yet been uncovered (as demarcated in **Table 1**). In brief, the KNN algorithm is a supervised learning algorithm which uses a set or sample of labelled training data to classify new test data (Bijalwan et al., 2014; Qian et al., 2004). This classification model looks at a specified number of neighbours (denoted as k) to classify new data points. Practically speaking, if k=3, the algorithm will examine the three (training-data) neighbours that are the closest to the new (test-data) point in order to decide how it will classify the supplied (test-data) point. Beyond being used for the KNN model, the binary division of groups will be treated as a covariate for an STM. For clarity, an STM (when used in

[19] The analysis will produce a DFM with Arabic-language features. In order to make the output more accessible, the researcher will translate the Arabic terms when necessary for data visualisation.
[20] This study will employ the cosine metric to initially compute similarly due to significant differences in the number of observations (published text) between select groups.

the R programming environment) allows for the discovery of topics and the estimation of their relationship to document metadata, with outputs of the model then allowing for hypothesis testing (Roberts et al., 2019). This model will be produced after out-of-sample groups are classified. In practice, an STM would allow the study to estimate the position of surrogate groups in relation to key (muqawama-relevant) topics.

As a final note, the robustness of this research study should be briefly addressed. It is important to state/note that measuring paramilitary *authenticity* will be fully based on Telegram text(s). To some extent, this may limit the generalisability of data when studying other militia groups, as the validity of this measure may be sample-specific.[21] Nonetheless, recent scholarly work has shown that Telegram is becoming a popular choice for VNSAs (Wan Mohd Nor, & El-Muhammady, 2021; Walther & McCoy, 2021), with many such groups also engaging in proxy tactics similar to Iran-back paramilitaries. Beyond generalisability, this study also recognises that similarity in published content may not provide a robust metric for *authenticity*. To this end, this study does not contend to provide a definite model for classification, rather a tool that can provide reliable first impressions and supporting metrics. Key to this reliability is the easily available data and replicable method. Given similar conditions and a hierarchy comparable to the one espoused by paramilitaries in Iraq, analogous research could provide conclusions in line with those found in this study.

5 Analysis: Results, Discussion, Implications

This section will elaborate on, analyse, and interpret the results collected after applying the above method and expand on the steps taken by the study to investigate the *authenticity* of

[21] Major Iran-backed paramilitaries have openly demonstrated the use of fictitious surrogates as a tactic in the past (Badawi, 2021; Elias 2020;, Knights et al., 2021a), but a tendency to use false flag terrorism may not be applicable to all cases.

new-wave surrogates. The first subsection will broadly outline the results of the conducted method while also providing additional explanations and interpretations where necessary. The following subsection will then offer a discussion of key insights provided by the results, and link data-based conclusions to the theoretical framework described in an earlier section of this study. This subsection will also practically apply data-based insights to the Iraqi security-political environment to (cautiously) offer realistic policy recommendations. Finally, limitations will be addressed as to delimit what similarity in discourse can say about paramilitary *authenticity,* and establish how reliable the author-constructed KNN machine-learning algorithm is in its current iteration.

5.1: Results

Table 2: Breakdown of Telegram Content by Group

Name (translated)	Abbreviation	Number of Telegram (text) messages
Kata'ib Hezbollah	KH	10984
Ashab al-Kahf	AK	3354
Usbat al-Tha'ireen	THA	117
Fariq Fatemiyoun al-Maidani	FFM	5022
Liwa Tha'ar al-Muhandis	LTM	338
Rab' Allah	RAB	916
Al-Wehda 10,000	WEH	9085
Qasim al-Jabarin	QAS	1674
Quwat Thu al-Faqar	FUQ	10803
Abu Jadaha	JAD	685
Al-Majame'e al-Khasa	SPE	1001
Saraya al-Thawra Al-Eshreen al-Tahniya	ISH	83
Saraya Tha'ar al-Shuhada	SHU	104
Liwa al-Shahid Ahmed Dar'am	DAR	423
Kata'ib Abul-Fadl al-Abbas	ABS	212

***Abbreviations are not official and were used for easy reference. Data in the table was compiled using the R-based 'Readtext' package, and base functions.**

Prior to performing the similarity analysis and producing the KNN machine-learning classifier, some preparatory steps were taken. As displayed above in **Table 2**, the very basic step of collecting all the (text) data belonging to each group was accomplished. The texts belonging to all surrogates are displayed for convenience. As was demarcated earlier in **Table 1**, groups highlighted in red are already known to be fictitious, and will be the basis of this analysis. The remaining groups will be treated as out-of-sample data, and used in the KNN machine-learning classifier. Markedly, the above table also serves to demonstrate the degree of variance in paramilitary group content-publication. Nevertheless, this is not a fully representative illustration

as it does not include the images or videos the groups may have published in the time-frame chosen for the analysis.[22] The implications of such variance will be further discussed in a later part of this section, but it can be briefly stated as a comparative limitation, and a factor that may constrain the reliability of the KNN classifier. However, it is of note that surrogates that are known to be fictitious are quite active online, as inferred from the number of Telegram messages they have published since their inception. To further investigate what predictions this collection of texts can provide on the *authenticity* of paramilitary surrogates, the data was processed to allow its conversion into a corpus, and then a DFM.

Table 3: Summary of the Top 10 Features in the 'militia' DFM

Feature (Arabic)	Feature (English)	Frequency	Docfreq
الله	God	1584	1376
المقاومة	The Resistance	733	702
الشهيد	The Martyr	550	543
الاحتلال	Occupation	520	494
الامريكي	The American	489	482
العراق	Iraq	352	303
محافظة	Governorate	329	325
ابو	Father (/Sir)	314	309
رتل	Column/Convoy	299	298
استهداف	Targeting	293	290

*Total number of features: 20, 051 | Total number of documents: 42257
***Data in this table was collected by the author using the R-based 'Quanteda' and 'Stopwords' package(s).**

In brief, by converting the Telegram text-content of all paramilitary groups into a corpus, character strings and variables were saved as a data-frame. This data-frame was then

[22] Some previous studies have specifically worked to investigate video/image content (Prucha, 2016; Zenn, 2020), but this does not fit the scope of this study. Nevertheless, there are techniques that can be used to extract text from images (i.e. optical character recognition). This sampling limitation will be further discussed later in the paper.

subsequently converted into a DFM. To ensure that the collected Telegram data was accurately processable at the next stage of text analysis, this paper also stemmed the selected texts and removed stopwords, symbols, numbers, and punctuation.[23] The study additionally removed any Latin characters/strings located in the DFM. Moreover, taking into account the tendency of paramilitary groups (either muqawama or known surrogates) to repeatedly refer to themselves in their content, the study also manually removed any mention of militia names from the DFM.[24] This culminated in a total number of 20,051 features.

A summary of the top 10 translated features, their overall frequency, and how often they appear in at least one document (Telegram chat) is presented in **Table 3**. Noticeably, a certain tone can be derived from the top features within the DFM. From a general observation, the collection of features (words) seen in **Table 3** adheres to the context of surrogate operations in "Iraq", namely their "targeting" of, and "resistance" against "the American" "Occupation". This is of course an estimation of what the words mean and their sentiment, as without their position in the sentences they have been extracted from, inferring conclusions is rather difficult and can be erroneous. However, taking into account the nature of Iran-backed paramilitary operations in Iraq, the objectives of the muqawama and its surrogates, and recent political/security events, the frequency of these terms and their potential overall meaning can be anticipated.

Top features (words) seen in **Table 3**, align with the central mission expressed by paramilitaries belonging to the muqawama, namely resisting the so-called US occupation in Iraq, and evicting foreign forces (Knights et al., 2021a; Nada & Rowan, 2021). This type of rhetorical framing has been viewed as a sort of rallying cry, and a unifying sentiment that has assisted in

[23] This study made use of the "Marimo" multilingual stopword/wordstem collection to pre-process the militia DFM (Watanabe, 2020). An alternate approach would involve the creation of a personalised dictionary or collection that is tailored to the texts in question. This recommendation is based on the necessity to manually edit out certain strings and terms that external collection does not capture or cover.

[24] This required a manual removal of specific words. R does fully interpret the right-to-left nature of Arabic texts, and as such, the study employed a UTF-8 encoder to convert the text into values R is more effective at interpreting.

the re-mobilisation of local paramilitary forces (Al-Hura, 2020; Knights, 2020; Reuters, 2020). Prior to re-orienting their focus to the US, Iran-backed paramilitaries in Iraq served the primary purpose of defeating IS clusters around the country (CEDOCA, 2019; Nada & Rowan, 2021). However, with the ongoing cessation of large-scale military operations against IS, KH and other leading groups in the PMF faced a burgeoning governance and therefore legitimacy deficit (Knights et al., 2021a; Podder, 2017; Suchman, 1995). Antagonistic rhetoric and actions towards the 'US occupation' were a manner by which Iran-backed paramilitaries (and the PMF) could realistically maintain their authority, influence and popularity. This antagonistic mentality soon became official policy following the death of two iconic paramilitary/PMF leaders, Qassem Soleimani and Abu Mahdi al-Muhandis (Knights et al., 2020; Knights, 2022).

With such contextual factors in mind, the dominance of these specific features in **Table 3** is predictable, as the multitude of fictitious surrogates under KH are likely to duplicate this sentiment in their content to gain public sympathy, recruit new followers, and put pressure on the Iraqi government. Nevertheless, despite this being a realistic inference, it should be made with care. **Table 3** does not fully distinguish which groups dominate in terms of contributions, and as such, it is possible for a small portion of groups to influence the frequency of features. Another valid -related- consideration is variance in published content. Given that the **Table 3** simply counts top features in all provided documents, it may ignore the stance of smaller surrogates that are less active or focus on different topics. In essence, although the features in **Table 3** may indicate that a sizable portion of the Telegram-chat corpus is aligned with beliefs posited by the muqawama, it should still be viewed as a superficial expression of broader surrogate beliefs. An STM, which will be produced and interpreted later in this subsection, can more clearly indicate the stance of groups in relation to certain topics, but it is first necessary to designate groups using

a specific division or covariate, namely degree of *authenticity.* This can be done using a similarity matrix.

Table 4: (Cosine) Similarity Matrix

Group	KH	AK	THA	FFM	LTM	RAB	WEH	QAS	FUQ
KH	0	0.70	0.20	0.35	0.33	0.57	0.56	0.14	0.52
AK	0.70	0	0.20	0.28	0.35	0.58	0.48	0.26	0.35
THA	0.20	0.20	0	0.07	0.07	0.16	0.11	0.03	0.09
FFM	0.35	0.28	0.07	0	0.14	0.24	0.30	0.07	0.29
LTM	0.33	0.35	0.07	0.14	0	0.24	0.24	0.14	0.16
RAB	0.57	0.58	0.16	0.24	0.24	0	0.24	0.05	0.33
WEH	0.57	0.48	0.11	0.30	0.24	0.24	0	0.40	0.31
QAS	0.14	0.26	0.03	0.07	0.14	0.05	0.40	0	0.07
FUQ	0.52	0.35	0.09	0.29	0.16	0.33	0.31	0.07	0

***Data in this table was calculated by the author using the R-based 'Textplot' package.**

Table 5: Cosine-Similarity Matrix (Ordered)

Group	KH	AK	RAB	WEH	FUQ	FFM	LTM	THA	QAS
KH	0	0.70	0.57	0.57	0.52	0.35	0.33	0.20	0.14

***Data in this table was calculated by the author using the R-based 'Textplot' package.**

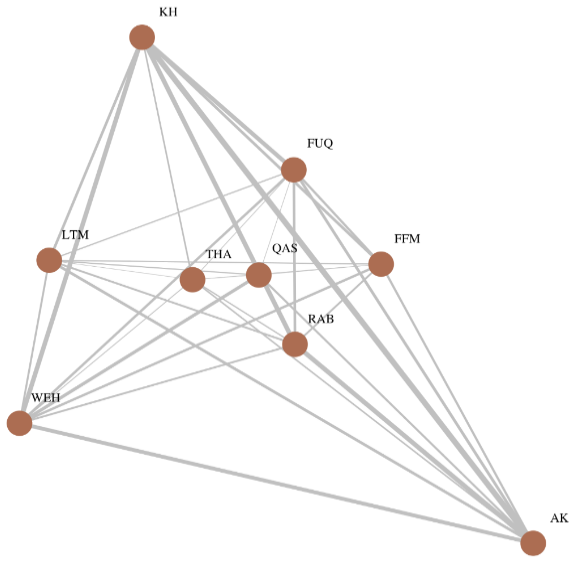
The previously constructed DFM (spliced to only include fictitious surrogates and KH) can be used to construct an applicable similarity matrix. This can be seen in **Table 4**. Results in the table indicate a baseline similarity of at least 3% (0.03*100), and a maximum of 70% (0.7*100) between all selected paramilitary documents. The study employed the Cosine metric at this stage to differentiate between known fictitious surrogates and ascertain how their online discourse compares against that of the chosen muqawama group: KH. More significantly, this matrix works to test the prediction posited by this study, and determine whether similarity can be used as a valid measure of *authenticity.*

Table 6: Closeness of Vertices

Paramilitary	Closeness
THA	1.072
QAS	0.852
LTM	0.596
FFM	0.573
FUQ	0.469
RAB	0.417
WEH	0.377
AK	0.312
KH	0.296

***Data in this table was collected by the author using the R-based 'iGraph' package.**

Figure 1: Cosine-Similarity Network of Militia Discourse



Principally, the results in **Table 4** indicate that the prediction that fictitious surrogates are likely to be highly similar in their online discourse to a muqawama sponsor is proven to be accurate, albeit with the added understanding that there is a spectrum along which fictitious surrogates can fall. **Table 5** takes a snapshot of the broader matrix, and demonstrates the distinct range in similarity scores seen when comparing the Telegram discourse of fictitious surrogates in relation KH. Some surrogates show an exceptionally high level of similarity (0.7), while others show a more muted, but still notable, similarity score (0.14). Likewise, **Figure 1** reinforces this metric in its network-based representation of the matrix presented in **Table 4**. Plainly, **Figure 1** provides a visual representation of the cosine-similarity matrix, with the thickness of the edges (connecting lines) indicating the level of similarity between vertices (militia groups), possibly showing how the surrogate online space is coordinated, and who the major players are. This interpretation can be further supplemented using a calculation of closeness centrality, being the measure of the average shortest distance from each vertex to each other vertex (Disney, 2020;

Golbeck, 2013). In principle, looking at the closeness measures listed in **Table 6,** it can be deduced that the high level of similarity between KH and most other surrogates, translates into it having the lowest closeness measure.

In practice, having the lowest closeness score can mean that KH is best placed to quickly influence the flow of information in the network. Referring back to the insights derived from the top 10 features in **Table 3,** the closeness measure associated with KH may give credence to the idea that the muqawama group imposes a specific narrative that is duplicated (with some adjustment) amongst all other surrogates. In a manner of speaking, the content published by KH is "just a hop away" from the content of all other vertices (surrogates) in the network (Hansen, Shneiderman, Smith & Himelboim, 2020, pp. 31-36). Furthermore, surrogates with closeness scores that are analogously low [AK, WEH, RAB, FUQ] could be interpreted as key vertices that expansively push a near duplicate version of KH content to the rest of the network.

Taking the above results into account, it can be inferred that KH plays a pronounced role as a coordinator of fictitious paramilitaries, particularly in relation to the following surrogate groups: AK, WEH, RAB and FUQ. This is a critical result, as the four groups have been shown to be key players in attacks against US coalition forces, and the publication of widely popular messages related to recruitment, resistance, and religious values.[25] Moreover, the degree of similarity between KH Telegram texts and the referenced groups is quite high, exceeding 50% (0.5*100). Such scores may relate to their performative standing as surrogates. AK is one of the most, if not the most, well-known surrogate(s) in Iraq, and has been shared amongst various paramilitaries in the muqawama to act as a front for operations against US-coalition convoys and targeted missile attacks on US bases in Iraq (Knights et al., 2021e). FUQ has served a similar purpose, albeit for smaller more targeted operations that do not earn as much publicity as those

[25] See Knights (2021) for a full breakdown of each group's profile and general activity.

conducted by AK (Knights et al., 2021a). Likewise, RAB and WEH have played prominent roles in false credit-claiming and the dissemination of the KH worldview, with the former being the 'mothership' / organiser for vigilante and counter-protest groups that operate within the KH network, and the latter being KH's premier propaganda, hacking, and intelligence gathering medium (Knights et al., 2021f; Knights et al., 2021g). In some ways, these groups do appear autonomous in their particular tasks, but their fictitious status and similarity to KH posits them as the arms to the octopus-like muqawama sponsor.

Table 7: Euclidean-Distance Matrix (Ordered)

Group	KH	AK	RAB	WEH	FUQ	FFM	LTM	THA	QAS
KH	0	524	612	627	661	690	712	727	841

***Data in this table was calculated by the author using the R-based 'Textplot' package.**

Taking this insights further, the range in similarity scores seen in **Table 5** can be practically applied to create a split between the fictitious surrogates: those with a 0.5+ discourse similarity score [AK, WEH, RAB, FUQ], and those with 0.5- score [FFM, LTM, QAS, THA]. These two groups can be functionally interpreted as a binary covariate [1: Over 50% | 0: Under 50%], and used for an STM and as target classes for KNN classification. Out of sample groups listed in both **Table 1** and **Table 2** will be classified into one of these target classes based on their proximity to the features of neighbouring surrogate groups (Bijalwan et al., 2014), and then arranged on an STM to investigate potential differences in preferred (discourse) topics.

Notably, unlike with the cosine-based similarity scores computed in **Table 4** or **Table 5**, the KNN algorithm uses euclidean distance to estimate the position of data points (here surrogate groups) in relation to one another (Bijalwan et al., 2014; Qian et al., 2004). To ensure that the split chosen for distinguishing out-of-sample groups was still applicable, another similarity

matrix was produced using a euclidean-distance metric. This is shown in **Table 7.** When comparing **Tables 5 and 7,** the order of similarity[26] is observed to be the same.[27] Alternatively, the degree of difference is not as pronounced in **Table 7** as it is in **Table 5**. This is not unexpected, as the manner by which cosine and euclidean measurements approach text data is rather different (Qian et al., 2004). Cosine similarity establishes the proximity of documents irrespective of their size (Singhal, 2001), while euclidean distance is fundamentally oriented around calculating the magnitude of similar features (words) in a document, which makes it rather useful for classification tasks (Guo et al., 2003; Khamar, 2013). The cosine measure is more oriented towards finding specific or unique words and using them to determine the orientation of a document, which often makes it more suitable for textual analysis/mining and similarity testing (Singhal, 2001; Wang & Dong, 2020). There are a number of additional proximity measures that go beyond the scope of this study, but for the given analysis, cosine similarity and euclidean distance allow for satisfactory (albeit bounded) insights.

Taking into account the binary covariate described above, this study then produced a KNN model. This model arranged groups into the following two target classes: those with a 0.5+ discourse similarity score [AK, WEH, RAB, FAQ], and those with 0.5- score [FFM, LTM, QAS, THA]. This required splicing the dataset into three segments for processing, and labelling groups with a score of above 0.5 with a (1), those below 0.5 with a (0), and the out-of-sample group with a random character as a placeholder until the KNN model determined its target class. This is summarised in **Table 8.**

[26] It is important to keep in mind that 'distance' and 'similarity' are not exactly the same. This study may use the terms interchangeably in its application of such methods for the measurement of proximity between vectors in a vector space, but such interchangeability is not always applicable.
[27] Some studies have shown that results retrieved using euclidean measurements can be similar to those retrieved with cosine measurements (Qian et al., 2004).

The training model used for classification included all the labelled groups, but to ensure that the model was accurate, four groups from the training sample (FUQ [1], WEH [1], FFM [0], THA [0]) were also included in the test sample. The aim behind this inclusion is to ensure that the model is correctly classifying the out-of-sample surrogates. If the KNN model can successfully classify groups with a known target class, then the classification of out-of-sample groups is more likely to be accurate. This approach, which combines a train/test split with a train-test procedure which encompasses the entire dataset, encounters a distinct dilemma of overfitting the model (Meng et al., 2007, pp. 153-155). This issue is further exacerbated by the rather small training dataset, and the ensuing low K value entailed by the limited number of neighbouring data points (Guo et al., 2003). Overfitting the model to the provided training data would mean that it is overly good at predicting/ classifying training data, but as a result unable to generalise well for future out-of-sample test data. An additional consideration is the amount of 'noise' (meaningless/repetitive information) in the provided training data (Sha'abani et al., 2020, pp. 555-565). This was noted when discussing **Table 2**. The large variance in paramilitary Telegram content can make particular data points (surrogates) more 'noisy' and influential in determining the class of a test data point. Such limitations will be expanded on in the research limitations subsection, but should still be acknowledged when examining the KNN classification results presented in **Table 9.**

Table 8: KNN Classification (Test/Training Samples)

Militia Class	Militia Name (Abbreviation)		
0.5+ discourse similarity score [1]	AK	RAB	FUQ
	WEH		
0.5- discourse similarity score [0]	THA	LTM	QAS
	FFM		
Test sample [to be classified]	JAD	SPE	ISH
	SHU	DAR	ABS

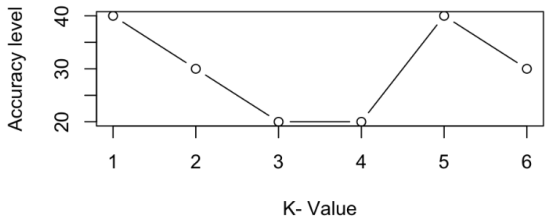
Table 9: Final KNN Classification (K=5)

Militia Class	Militia Name (Abbreviation)		
0.5+ discourse similarity score [1]	AK	RAB	FUQ
	WEH		
0.5- discourse similarity score [0]	THA	LTM	QAS
	FFM	JAD	SPE
	ISH	SHU	DAR
	ABS		

***Data in this table was produced by the author using base R functions, and the R-based 'Caret', 'plyr', and 'class' package(s).**

Keeping issues of overfitting in mind, and in order to potentially optimise the model, several values of K were tested in a loop and plotted. This can be seen in **Figure 2.** K=5 and K=1 are clear winners, as an accuracy of 40[28] is equivalent to a 100% correct prediction of the class of the four groups belonging to the training sample (FUQ

Figure 2: K-Value Accuracy Levels



[28] Given that the accuracy test estimated how well the model predicted the class of the four groups also located in the training sample, and the KNN model included ten groups, of which 6 were of an unknown class, the model indicated a success of 40%.

[1], WEH [1], FFM [0], THA [0]). Again, this result should not be taken as a reliable indication of the validity of the model given the above mentioned concerns. Evidence of this can be seen in the equivalence between K=1 and K=5, which is likely a result of the generally small training sample provided. Although each data point (here surrogate group) contains hundreds of features, the KNN algorithm focuses on the proximity of data points to data points, or specifically, surrogates to surrogates. Consequently, 'noise' and data variation can strongly influence classification results. In essence, the produced model is likely to be very sensitive to how much pre-processing was done to paramilitary data (after it was converted into a DFM) in each target class. This issue diminishes as the algorithm is provided with more test data. Thus, as noted earlier, when applied beyond the given case of Iran-backed Iraqi paramilitaries, the model is expected to be less sensitive to changes in data and be able to more consistently predict the classes of out-of-sample groups without huge variations based on small increments of K.

To ensure that the next stage of the analysis can produce interpretable outcomes that do not suffer from an uncertain foundation, a small test can be conducted to determine whether the euclidean-distance based KNN classification concurs with a cosine-based measure of document similarity. Initially, this study used cosine similarity to classify fictitious groups into two target classes. The computed similarity results are shown in **Table 5**, while the division into the 0.5+ [1], and 0.5- [0] classes is seen in **Table 8.** Credibly, and working backwards, conducting the same, but expanded, cosine-based measure of all surrogate Telegram content similarly should match the results seen in **Table 9** if the KNN model provided an accurate prediction of which classes out-of-sample groups would belong to. Results of this test can be seen in **Table 10**. As demarcated by the lack of a red highlight and the bolded values, out-of-sample groups do indeed

fit the [0] class, thus proving the KNN classifier to be accurate. Despite this, the limitations noted above should not be overlooked as they would weaken any similar classification model.

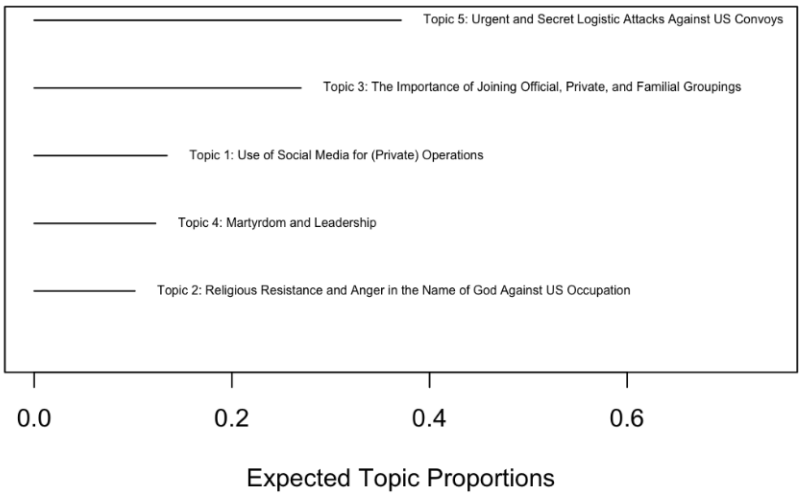
Table 10: Cosine-Similarity Matrix (Ordered) [All Surrogates]

Group	KH	AK	RAB	WEH	FUQ	SPE	JAD	FFM	LTM	DAR	ISH	ABS	THA	QAS	SHU
KH	0	0.70	0.57	0.57	0.52	**0.37**	**0.37**	0.35	0.33	**0.33**	**0.29**	**0.29**	0.20	0.14	**0.04**

***Data in this table was calculated by the author using the R-based 'Textplot' package.**

Taking the classification seen in **Table 9** a step further, an STM can be constructed to establish whether groups belonging to the 0.5+ discourse similarity score class [1] have a different topic focus as compared to those belonging to the 0.5- discourse similarity score class [0]. This expectation of difference in online discourse is based on a theoretical grounding. **Figure 3** offers some first impressions on this question by illustrating the distribution of all the collected features (sourced for a combined DFM) along 5 Topics. The topics produced below were based on the author's own interpretation of how the top words came together to form a concrete theme or subject and should not be taken as an absolute representation of surrogate discourse patterns. Nevertheless, **Figure 3** does offer some interesting insights that extend what was seen in **Table 3**. Two of five topics seem to revolve around resisting the US occupation or targeting US forces [Topics 5 + 2]. This further reinforces the framing approach taken by KH,

Figure 3: Top (5) Topics Composing Surrogate Discourse



and shows its prevalence in the online discourse of its surrogates. Topic 3 [“The Importance of Joining Official, Private, and Familial Groupings”] is also rather interesting, as it appears to almost attempt to mask the fact that most surrogate groups in the DFM are fictitious. The ideas seen in this topic lean in various directions, but recruitment and self-promotion are crucial aspects. Surrogates seem to both promise a sense of belonging, and advertise their status as an official group or gang. Given the urge of such surrogates to appear genuine in the eyes of their enemies, this discursive pattern matches what a normal group might post. Nevertheless, appearing credible to move suspicion away from muqawama groups like KH is not necessarily a consistent practice, and this pattern of discourse may simply be a duplication of the content posted by KH.

Table 11: STM Regression Table (Topics 3 + 5)

Topic	Coefficients	Estimate	t-value	p-value
(5) Urgent and Secret Attacks Against US Convoys	(Intercept)	0.342	2.234	0.044*
	Authenticity [0]	0.156	0.542	0.597
(3) The Importance of Joining Official, Private, and Familial Groupings	(Intercept)	0.283	2.045	0.062*
	Authenticity [0]	-0.041	-0.153	0.881

***Data in this table was produced by the author using base R functions, and the R-based ‘stm’ and ‘seededlda’ package(s).**

This then leads to the next inquiry, where the study tested whether the discourse of groups in class [1] differs from that of groups in class [0], when compared through the lens of the top topics listed in **Figure 3**. The immediate answer to this can be seen in **Table 11**. Prior to interpreting the concrete differences between the two classes, a quick comment on the significance of the results is worthwhile. It is clear that only some of the results in **Table 11** are statistically significant, and only one result is significant at ($p < 0.05$). Given that many groups in

the sample show a moderate-to-high degree of similarity to one another, particularly when this similarity is computed through a euclidean-distance measure as seen in **Table 4**, an STM regression may struggle to demonstrate major differences in the tested samples. As stated in other parts of this subsection, a wider and more diverse dataset could serve to correct this, and other issues. Nevertheless, these initial results do show some statistical significance, and as will be explained below, point towards the expected outcome. This can be an indication that the method does offer a valid measure of surrogate *authenticity,* and that results produced in the earlier similarity matrix and KNN model do allow for a satisfactory answer to the RQ.

Figure 4: Topic (5) Regression Plot

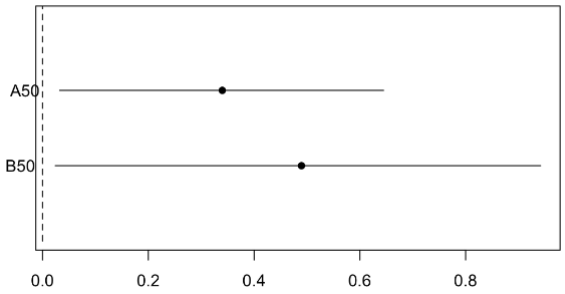
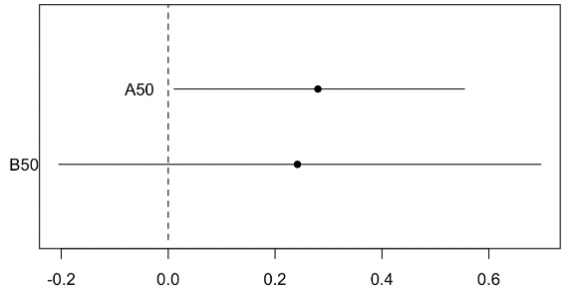


Figure 5: Topic (3) Regression Plot



To put this in more concrete terms, **Table 11** will be properly discussed with an understanding that insights and inferences are not significant enough to draw unconditional conclusions. Looking at Topics 3 [“The Importance of Joining Official, Private, and Familial Groupings”] and 5 [“Urgent and Secret Attacks Against US Convoys”], which occupy the highest proportions in the DFM, it can be observed that there is some level of difference in online discourse. This can also be visually seen in **Figure 4** which represents Topic 5, and **Figure 5** which represents Topic 3. Altogether, it appears that groups that have a similarity score of 0.5+ (represented in the two figures as A50), mention a combination of terms related to Topic 5 about 34.2% of time in their Telegram discourse, while groups that have a similarity score of 0.5- (represented in the two figures as B50) do so nearly 50% of the time. This is an interesting

result, as it is in-line with the idea that KH employs fictitious surrogates to avoid being linked to attacks against US convoys and maintain its public image. Groups with a similarity score above 50% are predictably duplicating most of the content published by KH, which may advocate against US occupation or congratulate groups on their successful attacks, but will not advertise such actions often lest it earn public and political ire. Nevertheless, as was discussed earlier, this conclusion should be tempered as although this pattern of behaviour appears to be statistically significant for A50 groups ($p < 0.05$), it is not for B50 groups ($p > 0.001$). This generally means that a distinct difference in behaviour is unlikely given the provided data. Alternatively, it appears that A50 groups will mention Topic 3 more often in their discourse (28%), as compared to B50 groups (24%). Evidently, the distinction is not major, and only marginally significant for A50 groups ($p < 0.1$) and not significant for B50 groups ($p > 0.001$).

5.2: Discussion and Implications

The previous section sought to methodologically test the theory-based prediction that online discourse could be an indicator of a paramilitary group's *authenticity,* and that similarity in published (Telegram) content between a sponsor and a surrogate could mean that the latter was *inauthentic*, and thus a fictitious front created for false flag terrorism operations. Additionally, the goal of this determination was to have a functional spectrum that could be divided into two classes for the later classification of out-of-sample surrogate paramilitary groups. All in all, this study did succeed in determining that similarly could be used as a valid metric for *authenticity*. This can be seen in either **Table 4** or **Table 5**. Moreover, the additional calculation of closeness centrality, seen in **Table 6**, and interpreted in combination with the distribution of top features

(words) in **Table 3**, worked to show the prominent role KH, the sample muqawama sponsor, holds as a distributor of information within the sample network of fictitious surrogates.

Nevertheless, and with reference to theories and concepts on credit-claiming discussed in the literature review, the formulation of KH online content (e.g. in the tone of threats of violence, denying organisational involvement, denying principal intent), and to what extent that formulation of ideas was duplicated by surrogate groups is not something that can be established using the proposed methodology. Although the general composition of information is evident in **Table 3**, and an assumption on the sentiment it contains can be made based on the nature of Iran-backed paramilitary operations in Iraq, the objectives of the muqawama and its surrogates, and recent political/security events, formulation remains an important aspect that can reveal more about the character, worldview, and intentions of a surrogate, with variations in formulation allowing for a better overall characterisation of surrogates, both fictitious and out-of-sample.

Despite such benefits, the dimension of similarity in content formulation, specifically in the context of credit-claiming, was not within the scope of analysis. Although the STM sought to include it in part, this dimension is an expansion to this study that could take the basic idea that there is a link between *authenticity* and content similarly a step further. Historical examples have shown that VNSAs are not one-minded in their violent operations (Mroszczyk & Abrahms, 2021, p. 427), and in the instances where credit is claimed under the premise of denying principal intent, it can manifest as an apology (Matesan & Berger, 2017). This can be seen in the case of al-Qaeda leaders issuing apologies after harm was done to civilians in terrorist attacks in Afghanistan and Pakistan (CNN, 2009), or Hezbollah leader Hassan Nasrallah apologising for a rocket attack that caused the death of two children in Nazareth (CNN, 2006).

Substantially, and with options for future research in mind, it should be stated that although scores in **Tables 5** and **7** support the study's prediction on the link between similarity and *authenticity*, it is not an indisputable conclusion. In support of the prediction, key fictitious groups with distinguishable roles like AK, WEH, FUQ and RAB were observed to have a relatively high similarity score of 50% and over (0.5*100), as well as a clearly low range of closeness centrality scores (0.31-0.47). The remaining fictitious surrogates, namely FFM, LTM, QAS, and THA, demonstrated lower similarity scores ranging from 14% to 35%, and consequently high closeness scores (0.57-1.0). To some degree, this range in scores may indicate that similarity cannot be used as a unitary and fully objective metric of *authenticity* estimation. Surrogates exist on a relatively wide spectrum of *authenticity,* with smaller fictitious surrogates [FFM, LTM, QAS, THA], who are not leading players in the paramilitary surrogate environment (Knights, 2020; Knights et al., 2021a; Nada & Rowan, 2021), being noticeably dissimilar in their online discourse when compared to KH. In practice, this does not mean that similarity should be rejected as a measure of *authenticity,* but rather that the computed results should be put in context and interpreted as part of a larger puzzle. As stated earlier, the methodology used by this study is proposed as a first step in a longer sequence of inquiries, research, and analysis.

For instance, the distinction between [AK, WEH, FUQ, RAB] and [FFM, LTM, QAS, THA], cannot be understood without knowledge of surrogate profiles and how influential they are in the Iraqi security-political environment. Furthermore, given lack of data on genuine surrogates, and the subsequent need to create a binary division based on the spectrum of fictitious group *authenticity,* visible in **Table 5**, the KNN classifier could only be functionally trained with the understanding that not all fictitious group are born equal, or at the very least, that not all fictitious groups are prominent players in the KH surrogate network. Knowing that [AK,

WEH, FUQ, RAB] are the major fictitious fronts preferred by KH for the majority of its activities (Knights et al., 2021a; Knights et al., 2021e; Knights et al., 2021f; Knights et al., 2021g), can then allow for the construction of target classes with the following logic: 0.5+ (50% and over similarity to KH) [1] and 0.5- (under 50% similarity to KH) [0]. The basic breakdown, and group allocation, can be seen in **Table 8.**

Overall, despite this functional division, and a generally successful KNN model (K=5), the limitations underlying the constructed model should not be omitted when considering its validity, and general reliability in classifying out-of-sample groups. A general lack of additional paramilitary groups to train the KNN model, conducting an accuracy test which combined a train/test split with a train-test procedure which encompassed the entire dataset, and issues of 'noise' in the dataset lead to a prominent risk of model overfitting which would considerably bound the ability of the KNN model to generalise, and predict out-of-sample surrogates. Given the intention of this study to conduct a STM to supplement prior results, and the reliance of the STM on the classification of out-of-sample groups into class [0] or [1], a small test was conducted using a cosine-similarity matrix. In some ways, this test reasoned backwards. Given that the KNN model was built around the original cosine matrix in **Table 5** which placed groups along a spectrum of *authenticity* and lead to the binary division associated with class [0] and [1], creating a similar matrix that is expanded to include out-of-sample groups rather than only fictitious surrogates should produce results that align with the KNN model (if it is correct). The matrix presented in **Table 10** did indeed allow for this verification, and did demonstrate that the classification of out-of-sample groups into the [0] class, as shown in **Table 9**, was accurate as their similarity in relation to KH was under 50%, with it ranging between 4% and 37%.

Verification of the accuracy of the accuracy KNN model in this scenario should nonetheless not justify the existence of design limitations. In any analysis, such factors are likely to influence the validity of conclusions and any inferences made on their basis. Yet, it should also be appreciated that this analysis of Iran-backed paramilitary groups in Iraq and their surrogates is fully reliant on the reality on the ground (in addition to the virtual reality on Telegram), to conduct the analysis. Analogous research that aims to study groups like IS (in Iraq/Syria), the increasingly active Islamic State of Khorasan Province's (ISIS-K), Al-Shabaab or Hezbollah, is likely to encounter similar issues when it comes to data collection, variety, and as seen in the STM, significance.

When looking at the STM, and specifically the results presented in **Table 11** and **Figures 4** and **5**, it is worth concisely discussing statistical significance. This study did not necessarily set out to produce statistically significant results, but it did aim to investigate a certain expectation. Generally, groups falling into the presumably different classes of [0] and [1], as presented in **Table 9**, are expected to show differences in how they approach key (muqawama-relevant) topics, especially given the variance they show in Telegram-content similarity (to KH). This expectation can be additionally grounded in theory. Although this study assumes groups in class [1] to be somewhat uniform in their online discourse (given their high similarity and profile as prominent KH fronts), groups in class [0], particularly out-of-sample groups, have shown moderate similarity scores which can indicate a different mode of discourse. Literature has been shown that no single narrative exists amongst, admittedly, genuine VNSAs (Prucha, 2016, pp. 48-50). Each VNSA enforces a coherent worldview which advances personalised religious, nationalist, ethnic, or ideological causes (Etaywe & Zappavigna, 2021). A distinction between class [0] and [1] might not fully allow for the implementation of such a theory with the inclusion

of known fictitious groups in class [0], but given that such smaller fictitious surrogates [FFM, LTM, QAS, THA], are not leading players in the paramilitary surrogate environment (Knights, 2020; Knights et al., 2021a; Nada & Rowan, 2021), they can still offer comparative insights.

Altogether, it should be noted that this particular analytical focus is somewhat beyond the scope of the study, particularly in its lean in the direction of studying similarity in content formulation. Despite this, the STM model does work to supplement previous results, and reinforce the KNN classification as a relevant division not only based on the factual profiles of the presented groups, but also on the basis of discourse similarity.

Practically, the aim of the STM was to investigate if different target classes were also statistically different in their dissimilarity. This investigation did not yield a significant result, but it nevertheless provided an outcome that met the methodological and theoretical expectation of a marked difference in topic preference between surrogates in class [0] and class [1]. This study specifically focused on Topic 3 [“The Importance of Joining Official, Private, and Familial Groupings”] and 5 [“Urgent and Secret Attacks Against US Convoys”], which occupy the highest proportions in the DFM.

Upon observation, and study of results in **Table 11,** there is a notable difference in how the two classes discuss the aforementioned topics. Altogether, it appears that groups in class [1] mention a combination of terms related to Topic 5 about 34.2% of time in their Telegram discourse, while groups in class [0] do so nearly 50% of the time. This is an interesting result, as it is in-line with the idea that KH employs fictitious surrogates to avoid being linked to attacks against US convoys and maintain its public image. Groups with a similarity score above 50% are predictably duplicating most of the content published by KH, which may advocate against US occupation or congratulate groups on their successful attacks, but will not advertise such actions

often lest it earn public and political ire. Nevertheless, as was discussed earlier, this conclusion should be tempered as although this pattern of behaviour appears to be statistically significant for class [1] groups ($p < 0.05$), it is not for class [0] groups ($p > 0.001$).

Alternatively, it appears that class [1] groups will mention Topic 3 more often in their discourse (28%), as compared to class [0] groups (24%). Evidently, the distinction is not major, and only marginally significant for class [1] groups ($p < 0.1$) and not significant for class [0] groups ($p > 0.001$). This topic is admittedly more vague, but its fixation on self-promotion and recruitment indicates it as a rather natural topic of focus amongst paramilitary surrogates, be they major or minor. Nevertheless, examining the topic from the perspective of class [1] groups, content published in relation to this topic might almost be perceived as an attempt to mask the fact that such groups are actually fictitious. Looking at the top features of the topic, and keeping in mind that any interpretation is subjective, surrogates seem to both promise a sense of belonging, and advertise their status as an official group or gang. Given the urge of such surrogates to appear genuine in the eyes of their enemies, this discursive pattern matches what a normal group might post. It can be seen as a clever form of subterfuge. Nevertheless, appearing credible to move suspicion away from muqawama groups like KH is not necessarily a consistent practice. Iran-backed paramilitaries are quite clever, but this does not mean that every move, statement or message on surrogate channels is calculated to minimise links to KH. Overall, this pattern of discourse may simply be a duplication of the content posted by KH. This is of course an assumption, but one based on the similarity of class [1] groups to KH, and their predicted tendency to duplicate KH content. Again, lack of analysis on similarities in content formulation makes an interpretation of the STM quite subjective.

Combining the above insights, it can be deduced that Telegram content can provide a number of insights on paramilitary *authenticity,* especially when similarity in content is used as a basis and paired with contextual knowledge. Moreover, similarity in online content can be a useful starting point for producing a KNN classifier that, depending on the provided dataset, is able to classify groups along a spectrum of *authenticity*, with target classes relating to the intention of the study. Practically speaking, the implications of this conclusion in Iraqi security policy is the rather immediate ability of front-line researchers, policy-makers, and regional academics to more clearly understand the second-order proxy network composed by factions like the muqawama, and potentially be able to attribute responsibility to the muqawama following the occurrence of false flag terrorism. Significantly, the tools offered by this study are, in a manner of speaking, primitive when considered without the broader context they are being used in.

As stated earlier, the methodology used by this study is proposed as a first step in a longer sequence of inquiries, research, and analysis. This first step may not provide robust conclusions in its immediate application, but it does offer the input required for concrete investigation. In addition to investigation, knowledge of the muqawama surrogate network, its key players, and the surrogates responsible for violent and propagandistic content can allow for security forces to choke the spread of messages and audio-visual demonstrations of violent content through critical nodes (like popular Telegram channels) (Mroszczyk & Abrahms, 2021). On the ground, this would mean limiting access to, and removing content from channels like AK, WEH, FUQ, RAB. Companies hosting media platforms like Telegram have customarily not been liable for user content published on their services (Ortutay, 2020), however recent years have shown a reassessment of this attitude (Klein & Flinn, 2017; White, 2020). There has been clear acknowledgement by social media companies that their platforms are being used by VNSAs like

the muqawama and its surrogates to spread extremist content/violence, leading to initiatives like the Global Internet Forum to Counter Terrorism (Mroszczyk & Abrahms, 2021). A concrete step would be for the Iraqi security forces to use this initiative, and others like it, to corner surrogate accounts and choke their online reach. However this is not an easy move, and given evidence of resistance to government interference by such platforms (Ortutay, 2020), a concrete case is necessary for effective action to be taken in regards to the muqawama, and VNSAs more broadly. To this end, applying the methodology proposed in this paper, and using it as a starting point for bigger, consolidated, and region-specific research is a legitimate manner by which the online influence of VNSAs can be tethered.

Looking beyond the case study of Iran-backed paramilitaries and their surrogates, it can be argued that the outlined methodological roadmap in this study can be viably applied to the content of various VNSAs operating around the world. As deduced earlier, Telegram content can provide a number of insights on paramilitary *authenticity,* especially when similarity in content is used as a basis and combined with contextual knowledge. This can be done to more clearly understand credit-claiming mechanisms, and to a larger extent, reveal the orientation of an online VNSA network, its information hubs, key influencers, and internal interactions. In essence, the method applied, and the conclusions reached, in this study are generalisable to other militarised environments, but with the assumption that limitations in sample size and data processing can be overcome, or sanctioned, as part of the conducted research process.

5.3: Research Limitations

In principle, this study did successfully establish that similarity could be used to determine *authenticity,* and was therefore also able to build a KNN classifier using insights from

the similarity matrix. Nevertheless, both the initial determination of similarity as a valid measure of paramilitary *authenticity*, and the accuracy of the constructed KNN algorithm and its classification of out-of-sample groups should be viewed as moderate successes. The noted tools only offer a basic explanation that grows even less robust without relevant context, and suffer from prominent limitations in their design.

For one, as addressed in the methodology section and demonstrated in the results/discussion, this study cannot use the above tools to reliably determine whether an out-of-sample surrogate is genuine. Overall, fictitious surrogates can be mapped on a spectrum of *authenticity,* and based on this spectrum, out-of-sample surrogates can be classified into one target class or another. Such classification can then allude to the likely status of out-of-sample surrogates, particularly when viewed through the 0.5+ (50% and over similarity to KH) [1] and 0.5- (under 50% similarity to KH) [0] lenses. Groups seen in the [1] target class are well-known players in KH's and the muqawama's operations, and as such can be seen to duplicate and propagate content that is highly similar to the content produced by KH. Groups in the [0] class are small paramilitaries that are not as prominent and seem to serve more specific aims that position them away from mainstream KH discourse. All of the out-of-sample groups fell within the [0] class. This classification indicates that out-of-sample groups have a similar discursive behaviour to smaller fictitious paramilitaries in the [0] class, but it does not go beyond this conclusion.

This leads to a secondary limitation, namely, sparsity of data. Given circumstantial factors in the Iraqi paramilitary landscape, a sample of genuine surrogates was not available. Methodologically, this means that the KNN classifier could not be trained to distinguish between genuine and fictitious surrogates, and that broader conclusions on a potential similarity (for

Telegram content) threshold, or range within which genuine surrogates are likely to fall, cannot be satisfied. This issue is a reflection of real circumstances, and can potentially reoccur in other 'real' cases. Keeping this in mind, future research can still employ the method delineated by this study to reach cursory insights or verify case-relevant theories, particularly when it comes to investigating credit-claiming or *authenticity* more broadly. Furthermore, paramilitary surrogates in Iraq show variance in Telegram activity, specifically seen in the disparate number of Telegram chats published by each group since its inception. This is another circumstantial factor that cannot be bypassed, but its effect on result robustness and reliability can be minimised.

Specifically, when looking to improve the reliability and accuracy of the KNN classifier, and prevent overfitting in the model, a bigger data sample is advisable. Increasing the amount of training data available to the KNN model will allow for higher K values, and prevent 'noisy' data points from strongly influencing the classification result. In that vein, a stricter approach to data-cleaning is another recommendation for improvement, and limitation faced by study. Given the author's reliance on external R packages, libraries, and dictionaries for pre-processing, some manual effort was necessary when constructing the DFM. Moreover, the trimmed (cleaned) DFM still included a considerable amount of 'noise'. With a dataset as small as the one used in the study, pre-processing can significantly influence the outcome of analysis. This issue is unlikely to persist with larger datasets, but cleaner and more focused data is always a good practice prior to classification, and when using techniques such as a STM.

In that regard, similar closing recommendations can be made for the STM in terms of sample size and data variety. As noted earlier, genuine surrogates were not available for the analysis, and groups that have already been established to be quite similar were used throughout. These two factors were shown to influence the ability of the model to distinguish between target

classes, and could conceivably prevent the detection of statistically significant regression results. With the above in mind, a larger sample size, cleaner data, better adapted statistical tools for data processing, and an attempt to balance the size of data points within the sample are concrete recommendations that can be proposed for the limitations of this study, and future research that may employ a similar methodological approach.

6. Conclusion

In sum, social media platforms like Telegram can be recognised as an effective tool for VNSAs to conduct strategic credit claiming operations like false flag terrorism, and simultaniously, an effective tool for researchers to examine such operations. VNSAs have demonstrated a distinct aptitude to disseminate propaganda to a large public audience; in doing so, they have exerted an outsized impact on public perception by employing a combination of extremist rhetoric and graphic violence (Berger & Morgan, 2015; Ward, 2020). Explicit examples of this include the Islamic State's (IS), Al-Qaeda's or Boko Haram's use of social media and online platforms [Twitter, Telegram, Youtube] (Chatfield et al., 2015; Pieslak et al., 2021; Weimann, 2010; Zenn, 2020). Strikingly, in a manner that mirrors the above groups, Iran-backed paramilitaries in Iraq and their surrogates have produced a decidedly similar effect on public audiences using instances of false flag terrorism facilitated by the Telegram social media platform.

With the above in mind, this study has sought to investigate this method of credit-claiming using the noted case-study of Iran-backed paramilitaries in Iraq, specifically the muqawama faction, and its deliberate network of proxies in the form of allegedly 'real' surrogates. To be precise, this study chose to investigate a sample of 14 new-wave surrogate groups that continue to be active on Telegram since their original creation in early 2020, and who

have been discursively credited with anti-US coalition operations, and attacks targeting domestic rivals after posting evidence on social media (Seligman, 2021). This choice is also on account of the belief that a portion of the aforementioned groups is fictitious (Badawi, 2021; Elias, 2020; Knights et al., 2021a), and evidence of the muqawama using social media outlets like Telegram or Twitter to create 'shell' profiles for the purpose of publicising their attacks (Badawi, 2021; Knights et al., 2020). With this in mind, this study elected to investigate the *authenticity* of such surrogates, being the degree of autonomy said actors have in representing themselves through relatable rhetoric, competent action, and attitudinal evaluations of their relationships (Ryan & Deci, 2000; Tracy & Robles, 2013, pp. 22-23; Van Leeuwen, 2001).

Concretely, in investigating the aforementioned surrogates, this study sought to answer the RQ: what can published Telegram content reveal about the *authenticity* of the new wave of Iran-backed paramilitary groups in Iraq? Underlying this RQ is a combination of theoretical insights based on prominent literature on representation, proxy sponsorship, credit-claiming, identity-building, and online discourse. In premise, *authentic* paramilitaries are understood as those who have the capacity to pursue an independent agenda detached from the interests of a sponsor (Ryan & Deci, 2000; Tracy & Robles, 2013, pp. 22-23; Van Leeuwen, 2001). The sponsor, in this case being a muqawama group, would use this surrogate for political objectives related to consolidating power, authority, and political-social legitimacy (Podder, 2017; Schuman, 1995). This view of the sponsorship-proxy arrangements between a VNSA like the muqawama and its paramilitary surrogates can then be extended beyond the use of genuine proxies as political assets (Moghadam & Wyss, 2020), to also encompass the use of fictitious proxies for the purpose of strategic credit-claiming (Kearns et al., 2014).

The rationale behind using fictitious proxies for strategic credit-claiming adheres to the same principle of employing genuine surrogates as a political asset, specifically in the sense that a VNSA must constantly balance between generating fear and attracting sympathy (Crenshaw, 1981; Lake, 2002). In a general sense, VNSAs are only likely to claim credit when the political return from the exercise is positive (Abrahams & Conrad, 2017, p. 281). Nevertheless, generating fear through performative feats of violence is a vital communicative aspect of VNSA operations (Pape, 2008), and as such, a balance or compromise is necessary. Fictitious surrogates fit neatly into this theoretical interaction by acting as a political asset through their utility as a cover for the violent operations of their sponsor. This can be particularly seen when civilian targets/casualties are involved (Findley & Young, 2007). As observed in the behaviour of muqawama groups like KH who have notable political ambitions and seek to present a moderate front (Abrahms & Conrad, 2017; Knights et al., 2021a), this can manifest as the false attribution of attacks to a rival third-party, or more specifically, a fictitious (*inauthentic)* group. Altogether, this can be described as false flag terrorism (Kearns et al., 2014, p. 425).

Principally, false flag terrorism is a rather apt view of how the muqawama operates in Iraq. However, key to this process of fabricating fictitious proxies for the purpose of taking credit for an attack or violence actually committed by the muqawama group, is Telegram. This is not to say that Telegram is pivotal to the general process of false flag terrorism, but rather that it is a vital part of false flag terrorism operations in Iraq. The ease of account creation, ability to share instantiations audio-visual evidence of violence (Purcha, 2016), user anonymity, and encryption of private content (Bloom et al., 2019), in addition to the freedom to delete all content for all users when necessary (Hamburger, 2014), are all highly conducive features to the creation of fictitious and temporary online identities. Furthermore, Telegram allows for the creation of a

multipurpose digital environment which both latently and actively connects surrogates to widely accessible paramilitary information hubs (Shehabat & Mitew, 2018). These hubs are a vital communication asset, and beyond coordinating activities and sharing centralised propaganda messages, they also allow a muqawama group to concretely broadcast its worldview through easily accessible online content (Prucha, 2016).

To a large extent, the value of media channels in how they allow muqawama groups to broadcast their worldview also serves to circle back to the idea that an *authentic* paramilitary is one that has the capacity to pursue an independent agenda detached from the interests of its sponsor. Specifically, it can be argued that online discourse is a manner by which surrogates can express/pursue an independent agenda, and thus demonstrate *authenticity.* Each group enforces a consistent worldview which carries original underpinnings of sociocultural meaning and advances religious, nationalist, ethnic, or ideological causes (Etaywe & Zappavigna, 2021). This coherent expression then serves to produce an authentic representation of a paramilitary's identity, as its repetition and demonstration on platforms like Telegram reinforces and embeds it in the minds of the audience. With this, it can be inferred that a genuine proxy, like its sponsor, will express an *authentic* worldview in online discourse, while a fictitious proxy will only offer a duplication of the worldview conveyed by its 'sponsor'.

Taking the above into account, and for the purpose of an effective investigation, the aforementioned RQ was restated as a testable prediction. The study posited that increased Telegram-content dissimilarity from a muqawama group would indicate higher levels of *authenticity*, while similarity may allude to the fact that suspect groups are fictitious, *inauthentic*, surrogates. To effectively test this relationship between *authenticity* and discourse similarity, *authenticity* was operationalised as a continuum/spectrum along which the Telegram text-content

of surrogates can be positioned. Placement on this spectrum corresponded to how distant or unalike a surrogate group's content is from the selected muqawama sponsor, with higher or lower percentages in text divergence corresponding to a surrogate's *authenticity.* Notably, as was shown by the results, this placement did not assume to determine the *authenticity* of surrogates in absolute terms, but rather worked to test similarity as a metric of *authenticity,* test the overall methodology in its development of a KNN model for the classification of out-of-sample surrogates, and provide preliminary insights on out-of-sample surrogates.

In reference to the above, it can be further stated that the purpose of this research was to provide an exploratory study that would lay the groundwork for future explorations of VNSA virtual networks and/or the use of false flag terrorism on social media platforms. Taking this into account, it should be noted that the author's attempt to build a strong theoretical background was aimed at positioning the analysis and its broader value, and although references concepts and theories were used to interpret results relating to Iran-backed paramilitaries and their surrogates, the intention was not to make definitive claims on this specific case. In essence, the provided case which investigated the similarity between documents published by new-wave surrogates and the muqawama group (KH), was a template to test the reliability, validity and potential generalisability of the proposed method.

In the end, it can be stated that the study made a modest but useful contribution by providing a methodological roadmap by which groups can be classified on a spectrum of *authenticity,* and mapped in terms of their role and value within a broader online (paramilitary) network. In this regard, using similarity between the text-based Telegram content of a paramilitary surrogate and a prominent sponsor as a measure of *authenticity,* was a crucial metric. In plain words, it can be argued that Telegram content, viewed through the lens of

similarity, can reveal preliminary impressions on the *authenticity* of a surrogate group. Practically, this meant that through a (cosine) measurement of document similarity, this study was able to establish a binary division that could be used for a KNN classifier and an STM. The KNN model successfully classified out-of-sample groups along an *authenticity* spectrum, while the STM supplemented findings on surrogate classification by exploring whether surrogates in either classification demonstrated a significant difference in topic preferences. Results from the STM were not significant, but pointed towards the expectations posited by the study and facts on the ground. Nevertheless, it should be reinforced that the method also faced conspicuous limitations that influenced its overall reliability, and the validity of inferences derived from it.

Altogether, it can be concisely stated that key issues faced by this study in its attempt to indicate that document similarity was a valid measure of *authenticity,* related to real-world circumstances. Nevertheless, to be explicit, these circumstances can be also translated into concrete limitations in the research design. When looking at the results gathered by the study, and the approach taken to their interpretation, which relied on contextual information, it can be understood that the methodology used for the study of Iran-backed paramilitary groups, and proposed by the author, is one that is pragmatic. Given the primary intention of this study's methodology, namely providing first impressions on real VNSAs operating in the online space and engaging in strategic credit-claiming/false flag terrorism, limitations are part and parcel. From risks of overfitting the KNN model, to more basic issues related to uneven datasets or limited data points, such factors are an expected concern when dealing with a topic of this nature and particularly actors that operate on a constantly changing online platform. Despite this, and as shown by the study, conclusions can still be reached, and expectations met. In essence, the method applied, and the inferences made in this study are generalisable to other militarised

environments, but with the assumption that limitations in sample size and data processing can be overcome, or sanctioned, as part of the conducted research process.

References,

Abrahms, M. (2013). The credibility paradox: Violence as a double-edged sword in international politics. *International Studies Quarterly*, 57(4), 660-671. https://doi.org/10.1111/isqu.12098

Abrahms, M., & Gottfried, M. S. (2016). Does terrorism pay? An empirical analysis. *Terrorism and Political Violence*, 28(1), 72-89. https://doi.org/10.1080/09546553.2013.879057

Abrahms, M., & Conrad, J. (2017). The strategic logic of credit claiming: A new theory for anonymous terrorist attacks. *Security Studies*, 26(2), 279-304. https://doi.org/10.1080/09636412.2017.1280304

Abrahms, M. (2018). *Rules for rebels: The science of victory in militant history*. Oxford University Press.

Abrahms, M., Ward, M., & Kennedy, R. (2018). Explaining civilian attacks: Terrorist networks, principal-agent problems and target selection. *Perspectives on Terrorism*, 12(1), 23-45. https://www.jstor.org/stable/26343744

Ahram, A. (2011). *Proxy Warriors*. Stanford University Press.

Al-Hura (2020). Sources reveal details of an Iraqi entity directly linked to Iran. *Al-Hura.* https://www.alhurra.com/iraq/2020/04/26/%D9%85%D8%B5%D8%A7%D8%AF%D8%B1-%D8%AA%D9%83%D8%B4%D9%81-%D8%AA%D9%81%D8%A7%D8%B5%D9%8A%D9%84-%D9%83%D9%8A%D8%A7%D9%86-%D8%B9%D8%B1%D8%A7%D9%82%D9%8A-%D9%85%D8%B1%D8%AA%D8%A8%D8%B7-%D9%85%D8%A8%D8%A7%D8%B4%D8%B1%D8%A9-%D8%A8%D8%A5%D9%8A%D8%B1%D8%A7%D9%86

Arnold, T., Ballier, N., Lissón, P., & Tilton, L. (2019). Beyond lexical frequencies: using R for text analysis in the digital humanities. *Language Resources and Evaluation*, 53(4), 707-733. https://link.springer.com/article/10.1007/s10579-019-09456-6\

Badawi, T. (2021). Iraq's Resurgent Paramilitaries. *CARNEGIE*. https://carnegieendowment.org/sada/84368

Bakke, K. M., Cunningham, K. G., & Seymour, L. J. (2012). A plague of initials: Fragmentation, cohesion, and infighting in civil wars. *Perspectives on Politics*, 10(2), 265-283.

Bale, J. M. (1996). The may 1973 terrorist attack at Milan police HQ: Anarchist 'propaganda of the deed'or 'false flag'provocation?. *Terrorism and Political Violence*, 8(1), 132-166. https://doi.org/10.1080/09546559608427337

Bar-Siman-Tov, Y. (1984). The Strategy of War by Proxy. Cooperation and Conflict, 19(4), 263–273. https://doi.org/10.1177/001083678401900405

Berger, J. M., & Morgan, J. (2015). The ISIS Twitter Census: Defining and describing the population of ISIS supporters on Twitter [Analysis Paper No. 20]. The Brookings Project on U.S. Relations with the Islamic World.

Berrebi, C., & Klor, E. F. (2008). Are voters sensitive to terrorism? Direct evidence from the Israeli electorate. *American Political Science Review*, 102(3), 279-301. https://www.cambridge.org/core/journals/american-political-science-review/article/are-voters-sensitive-to-terrorism-direct-evidence-from-the-israeli-electorate/B1FE65A2EA22B126F63B48E25DBB09D2

Bijalwan, V., Kumar, V., Kumari, P., & Pascual, J. (2014). KNN based machine learning approach for text and document mining. *International Journal of Database Theory and Application,* 7(1), 61-70. https://www.researchgate.net/profile/Bhavna-Reddy/post/how_text_classification_is_based_on_rocchios_method/attachment/59d623706cda7b8083a1e0d1/AS%3A331938180157440%401456151637583/download/knn+document+classification+%281%29.pdf

Bloom, M. (2005). *Dying to kill: The allure of suicide terror*. Columbia University Press.

Bloom, M., Tiflati, H., & Horgan, J. (2019). Navigating ISIS's preferred platform: Telegram. *Terrorism and Political Violence*, 31(6), 1242-1254. https://doi.org/10.1080/09546553.2017.1339695

Brown, J. M. (2020). Force of words: the role of threats in terrorism. *Terrorism and political violence*, 32(7), 1527-1549. https://doi.org/10.1080/09546553.2018.1486301

Bruce, S. (2001). Terrorism and Politics: The Case of Northern Ireland's Loyalist Paramilitaries. *Terrorism and Political Violence*, 13(2), 27–48. https://doi.org/10.1080/09546550109609679

Bruce, S. (2004). TURF WAR AND PEACE: LOYALIST PARAMILITARIES SINCE 1994. *Terrorism and Political Violence*, 16(3), 501–521. https://doi.org/10.1080/09546550490509829

Byman, D., Chalk, P., Hoffman, B., Rosenau, W., & Brannan, D. (2001). *Trends in outside support for insurgent movements*. Rand Corporation.

Byman, D. (2005). *Deadly connections: States that sponsor terrorism*. Cambridge University Press.

Chatfield, A. T., Reddick, C. G., & Brajawidagda, U. (2015). Tweeting propaganda, radicalization and recruitment: Islamic state supporters multi-sided twitter networks. In *Proceedings of the 16th Annual International Conference on Digital Government Research* [239–249]. https://dl.acm.org/doi/pdf/10.1145/2757401.2757408

CEDOCA. (2019). *Iraq: Targeting of Individuals* [EASO Country of Origin Information Report]. https://coi.euaa.europa.eu/administration/easo/PLib/Iraq_targeting_of_individuals.pdf

Clarke, C. P. (2019). *After the Caliphate*. Polity Press.

CNN. (2006). Hezbollah leader apologizes for attack's child victims. *CNN*. https://edition.cnn.com/2006/WORLD/meast/07/20/nasrallah.interview/index.html

CNN. (2009). Al Qaeda offers 'condolences' for innocent victims. *CNN*. http://edition.cnn.com/2009/WORLD/asiapcf/12/12/afghanistan.alqaeda/index.html

Conrad, J., Greene, K. T., Phillips, B. J., & Daly, S. (2021). Competition from Within: Ethnicity, Power, and Militant Group Rivalry. *Defense and Peace Economics*, 32(6), 757-772.

Conrad, J., & Greene, K. (2015). Competition, differentiation, and the severity of terrorist attacks. *The Journal of Politics*, 77(2), 546-561.

Conway, M. (2017). Determining the role of the internet in violent extremism and terrorism: Six suggestions for progressing research. Studies in Conflict & Terrorism, 40(1), 77-98. https://doi.org/10.1080/1057610X.2016.1157408

Conway, M., Scrivens, R., & Macnair, L. (2019). *Right-Wing Extremists' Persistent Online Presence: History and Contemporary Trends*. The Hague: The International Centre for Counter-Terrorism.

Crenshaw, M. (1981). The causes of terrorism. *Comparative politics*, 13(4), 379-399. https://doi.org/10.2307/421717

Crenshaw, M. (2001). Counterterrorism policy and the political process. *Studies in conflict and terrorism*, 24(5), 329-337. https://www.tandfonline.com/doi/pdf/10.1080/10576100175043420 4?casa_token=0cuGYugXF6sAAAAA:gT9ERbnIbMxoYGyCSAx1Mfi51raCBbrykCBsWdUbZOHAr5gAYsRkWvVxz91ozyjNtNRAR5tjTHUQ1g

Crenshaw, M. (2007). The logic of terrorism. In S. Mahan & P. L. Griset (Eds.), *Terrorism in perspective*, (pp. 24-33). London: Sage Publications.

Cronin, A. K. (2009). *How Terrorism Ends: Understanding the Decline and Demise of Terrorist Campaigns.* Princeton University Press.

Cunningham, D. E., Gleditsch, K. S., & Salehyan, I. (2013). Non-state actors in civil wars: A new dataset. *Conflict Management and Peace Science*, 30(5), 516–531. https://doi.org/10.1177/0738894213499673

Dalton, M. G. (2017). Iranian Backed Militias: Destabilizing the Middle East (Statement Before the House Foreign Affairs Committee Subcommittee on Terrorism, Nonproliferation, and Trade). *Center for Strategic and International Studies* (CSIS), 4. http://csis-website-prod.s3.amazonaws.com/s3fs-public/congressional_testimony/ts171004_Dalton_testimony_TNT.PDF

Disney, A. (2020). Social network analysis 101: centrality measures explained. *Cambridge Intelligence.* https://cambridge-intelligence.com/keylines-faqs-social-network-analysis/

Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 139-182). Philadelphia, PA: John Benjamins.

Dunér, B. (1981). Proxy Intervention in Civil Wars. *Journal of Peace Research*, 18(4), 353–361. https://doi.org/10.1177/002234338101800404

Elias, F. (2020). "Katyusha Cells": The Long Arm of Iran-Backed Factions in Iraq. *Emirates Policy Center.* https://epc.ae/en/details/featured/katyusha-cells-the-long-arm-of-iran-backed-factions-in-iraq

Etaywe, A., & Zappavigna, M. (2021). Identity, ideology and threatening communication: An investigation of patterns of attitude in terrorist discourse. *Journal of Language Aggression and Conflict*. https://www.jbe-platform.com/content/journals/10.1075/jlac.00058.eta

Ezzeddine, N., & van Veen, E. (2021). Qassem Musleh and Iraq's Popular Mobilization Forces. *Clingendael.* https://www.clingendael.org/publication/qassem-musleh-and-iraqs-popular-mobilization-forces

Findley, M. G., & Young, J. K. (2007). Fighting fire with fire? How (not) to neutralize an insurgency. *Civil Wars,* 9(4), 378-401. https://doi.org/10.1080/13698240701699482

Forrest, C. (2009). Coercive engagement: a security analysis of Iranian support to Iraqi Shia militias. *Strategic Studies Quarterly*, 3(2), 99-123. https://www.jstor.org/stable/26268564

Fromkin, D. (1975). The strategy of terrorism. *Foreign Affairs*, 53(4), 683-698. https://www.jstor.org/stable/20039540

Gaibulloev, K., & Sandler, T. (2009). The impact of terrorism and conflicts on growth in Asia. *Economics & Politics*, 21(3), 359-383. https://doi.org/10.1111/j.1468-0343.2009.00347.x

Gaibulloev, K., Hou, D., & Sandler, T. (2020). How do the factors determining terrorist groups' longevity differ from those affecting their success? *European Journal of Political Economy*, 65, 101935. https://doi.org/10.1016/j.ejpoleco.2020.101935

Golbeck, J. (2013). Analyzing the social web. Newnes.

Groh, T. L. (2019). *Proxy War: The Least Bad Option* (1st ed.). Stanford University Press.

Grynkewich, A. G. (2008). Welfare as warfare: How violent non-state groups use social services to attack the state. *Studies in Conflict & Terrorism*, 31(4), 350-370. https://doi.org/10.1080/10576100801931321

Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). KNN model-based approach in classification. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"* (pp. 986-996). Springer.

Hamburger E. (2014). Why Telegram Has Become the Hottest Messaging App in the World. *The Verge.* http://www.theverge.com/2014/2/25/5445864/Telegram-messenger-hottest-app-in-the-world

Hansen, D. L., Shneiderman, B., Smith, M. A., & Himelboim, I. (2020). Social network analysis: measuring, mapping, and modeling collections of connections. In *Analyzing social media networks with NodeXL*, (pp. 31-51). Elsevier.

Habermas, J., & Burger, T. (1989). *The structural transformation of the public sphere : an inquiry into a category of bourgeois society*. Polity Press.

Hoffman, B. (1997). Reply to Pluchinsky and Rapoport comments. *Terrorism and Political Violence,* 9(1), 18-19. https://doi.org/10.1080/09546559708427384

Hughes, G. (2012). *My Enemy's Enemy: Proxy Warfare in International Politics*. Apollo Books.

Hughes, G. (2016). Militias in internal warfare: From the colonial era to the contemporary Middle East. *Small Wars & Insurgencies*, 27(2), 196–225. https://doi.org/10.1080/09592318.2015.1129171

Hughes, S., & Meleagrou-Hitchens, A. (2017). The Threat to the United States from the Islamic State's Virtual Entrepreneurs. *CTC Sentinel*, 10(3), 1-8.

Ihaka, R., & Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299–314. https://doi.org/10.2307/1390807

Jackson, P. (2003). Warlords as alternative forms of Governance. *Small Wars & Insurgencies*, 14(2), 131–150. https://doi.org/10.1080/09592310412331300716

Jenkins, B. M. (1974). *International terrorism: A new kind of warfare*. Rand Corporation.

Kearns, E. M., Conlon, B., & Young, J. K. (2014). Lying about terrorism. *Studies in Conflict & Terrorism*, 37(5), 422-439. https://doi.org/10.1080/1057610X.2014.893480

Kearns, E. M. (2021). When to take credit for terrorism? A cross-national examination of claims and attributions. *Terrorism and political violence*, 33(1), 164-193. https://doi.org/10.1080/09546553.2018.1540982

Kenney, M. (2007). *From Pablo to Osama : trafficking and terrorist networks, government bureaucracies, and competitive adaptation*. The Pennsylvania State University Press.

Khamar, K. (2013). Short text classification using kNN based on distance function. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(4), 1916-1919. https://www.academia.edu/download/38502879/knn.pdf

Klein, S., & Flinn, C. (2017). Social media compliance programs and the war against terrorism. *Harv. Nat'l Sec. J.*, 8, 53. https://heinonline.org/HOL/LandingPage?handle=hein.journals/harvardnsj8&div=4&id=&page=

Knights, M. (2019). Iran's expanding militia army in Iraq: The new special groups. *CTC Sentinel*, 12(7), 1-12. https://ctc.usma.edu/wp-content/uploads/2019/08/CTC-SENTINEL-072019.pdf

Knights, M., Malik, H., & Al-Tamimi, A. J. (2020). The Future of Iraq's Popular Mobilization Forces [Policy Analysis]. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/future-iraqs-popular-mobilization-forces

Knights, M., Malik, H., & Smith, C. (2021a). Discordance in the Iran Threat Network in Iraq: Militia Competition and Rivalry. *CTC Sentinel*, 14(8), 1-21. https://ctc.usma.edu/wp-content/uploads/2021/10/CTC-SENTINEL-082021.pdf

Knights, M., Malik, H., & Smith, C. (2021b, November). Sabereen News Criticizes Iran for Lack of Support [Policy Analysis]. *The Washington Institute for Near East Policy*.https://www.washingtoninstitute.org/policy-analysis/sabereen-news-criticizes-iran-lack-support

Knights, M., Malik, H., & Smith, C. (2021c, December). The Militia Assassination Surge (Part 1): Sabereen and Unit 10,000 Threaten Iraqi Forces [Policy Analysis]. *The Washington Institute for Near East Policy.* https://www.washingtoninstitute.org/policy-analysis/militia-assassination-surge-part-1-sabereen-and-unit-10000-threaten-iraqi-forces

Knights, M., Malik, H., & Smith, C. (2021d, April). Profile: Sabereen News [Policy Analysis]. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/profile-sabereen-news

Knights, M., Malik, H., & Smith, C. (2021e, April). Profile: Ashab al-Kahf [Policy Analysis]. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/profile-ashab-al-kahf

Knights, M., Malik, H., & Smith, C. (2021f, April). Profile: Raba Allah [Policy Analysis]. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/profile-raba-allah

Knights, M., Malik, H., & Smith, C. (2021g, April). Profile: Unit 10,000 [Policy Analysis]. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/profile-unit-10000

Knights, M. (2020). Back into the Shadows? The Future of Kata'ib Hezbollah and Iran's Other Proxies in Iraq. *CTC Sentinel*, 13, 1-22. https://ctc.usma.edu/back-into-the-shadows-the-future-of-kataib-hezbollah-and-irans-other-proxies-in-iraq/

Knights, M. (2021, March). How to Use Militia Spotlight: Profiles [Policy Analysis]. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/how-use-militia-spotlight-profiles

Knights, M. (2022, January). Muhandis Replaced by Not One Man, But Two? [Policy Analysis]. *The Washington Institute for Near East Policy.* https://www.washingtoninstitute.org/policy-analysis/muhandis-replaced-not-one-man-two

Knights, M., & Smith, C. (2022, April). Spring Awakening: A New Wave of Militia Strikes or Just a Blip? [Policy Analysis]. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/spring-awakening-new-wave-militia-strikes-or-just-blip

Kydd, A. H., & Walter, B. F. (2006). The Strategies of Terrorism. International Security, 31(1), 49–80. https://doi.org/10.1162/isec.2006.31.1.49

LaFree, G., & Dugan, L. (2004). How does studying terrorism compare to studying crime. *Terrorism and counter-terrorism: Criminological perspectives*, 5, 53-74. https://www.researchgate.net/profile/Mathieu-Deflem-2/publication/331306143_Terrorism_and_Counter-Terrorism_Criminological_Perspectives/links/5c71c205458515831f699682/Terrorism-and-Counter-Terrorism-Criminological-Perspectives.pdf#page=63

LaFree, G., Dugan, L., & Miller, E. (2014). *Putting terrorism in context: Lessons from the Global Terrorism Database*. Routledge.

Lake, D. A. (2002). Rational extremism: Understanding terrorism in the twenty-first century. *Dialogue IO*, 1(1), 15-28. https://www.cambridge.org/core/journals/dialogue-io/article/rational-extremism-understanding-terrorism-in-the-twentyfirst-century/BF1544EC14CAEB60B0118EFB9D0B6E5C

Lesser, I., Arquilla, J., Hoffman, B., Ronfeldt, D. F., & Zanini, M. (1999). *Countering the new terrorism*. RAND corporation.

Lohlker, R. (2016). Theology matters: The case of jihadi Islam. *Strategic Review*, 6(3), 92-105. https://www.baytarrahmah.org/media/2016/Strategic-Review_Theology-matters-The-case-of-jihadi-Islam_Jul-Sep-2016.pdf

MacKinlay, J. (2000). Defining warlords. *International Peacekeeping (London, England)*, 7(1), 48–62. https://doi.org/10.1080/13533310008413818

Malet, D. (2013). *Foreign fighters: Transnational identity in civil conflicts*. Oxford University Press.

Malik, H. (2021). Qais al-Khazali's Show of Independence. *The Washington Institute for Near East Policy*. https://www.washingtoninstitute.org/policy-analysis/qais-al-khazalis-show-independence

Mansour, R. (2021). *Networks of Power: The Popular Mobilization Forces and the State in Iraq* [Research Paper]. https://www.chathamhouse.org/sites/default/files/2021-02/2021-02-25-networks-of-power-mansour.pdf

Mansour, R. (2018, April). More than militias: Iraq's popular mobilization forces are here to stay. *War on the Rocks.* https://warontherocks.com/2018/04/more-than-militias-iraqs-popular-mobilization-forces-are-here-to-stay/

Matesan, I. E., & Berger, R. (2017). Blunders and blame: how armed non-state actors react to their mistakes. *Studies in Conflict & Terrorism*, 40(5), 376-398. https://doi.org/10.1080/1057610X.2016.1210891\

McInnis, J. M. (2016). Iranian Deterrence Strategy and Use of Proxies. *Hampton Roads International Security Quarterly*. https://www.foreign.senate.gov/imo/media/doc/112916_McInnis_Testimony.pdf

Meng, Q., Cieszewski, C. J., Madden, M., & Borders, B. E. (2007). K nearest neighbor method for forest inventory using remote sensing data. *GIScience & Remote Sensing*, 44(2), 149-165. https://doi.org/10.2747/1548-1603.44.2.149

Moghadam, A. (2008). *The globalization of martyrdom: Al Qaeda, Salafi Jihad, and the diffusion of suicide attacks*. JHU Press.

Moghadam, A., & Wyss, M. (2020). The Political Power of Proxies: Why Nonstate Actors Use Local Surrogates. *International Security*, 44(4), 119–157. https://doi.org/10.1162/isec_a_00377

Mroszczyk, J., & Abrahms, M. (2021). Countering extremist organizations in the information domain. In A. Sheehan, E. Marquardt & E. Collins (Eds.), *Routledge Handbook of US Counterterrorism and Irregular Warfare Operations* (pp. 423-435). Routledge.

Mumford, A. (2013). *Proxy warfare*. John Wiley & Sons.

Nada, G., & Rowan, M. (2021, November). Profiles: Pro-Iran Militias in Iraq. *United States Institute of Peace: The Iran Primer.* https://iranprimer.usip.org/blog/2021/nov/10/profiles-pro-iran-militias-iraq

Neumann, P. R., & Smith, M. L. R. (2007). *The strategy of terrorism: How it works, and why it fails*. Routledge.

Palasinski, M., & Bowman-Grieve, L. (2017). Tackling cyber-terrorism: Balancing surveillance with counter-communication. *Security Journal*, 30(2), 556-568. https://link.springer.com/article/10.1057/sj.2014.19

Ortutay, B. (2020, October). AP Explains: The rule that made the modern internet. *AP*. https://apnews.com/article/what-is-section-230-tech-giants-77bce70089964c1e6fc87228ccdb0618

O'Shaughnessy, N. J., & Baines, P. R. (2009). Selling terror: The symbolization and positioning of Jihad. *Marketing Theory*, 9(2), 227-241. https://doi.org/10.1177/1470593109103069

Pape, R. A. (2008). Dying to win: The strategic logic of suicide terrorism. In M. Perry & H. E. Negrin (Eds), *The theory and practice of Islamic terrorism* (pp. 129-132). Palgrave Macmillan

Piazza, J. A. (2017). Repression and terrorism: A cross-national empirical analysis of types of repression and domestic terrorism. *Terrorism and Political Violence*, 29(1), 102-118. https://doi.org/10.1080/09546553.2014.994061

Pieslak, J., Pieslak, B., & Lemieux, A. F. (2021). Trends of anashid usage in Da 'esh video messaging and implications for identifying terrorist audio and video. *Studies in Conflict & Terrorism*, 44(4), 310-325. https://doi.org/10.1080/1057610X.2018.1545828

Pluchinsky, D. A. (1997). The terrorism puzzle: Missing pieces and no boxcover. *Terrorism and Political Violence*, 9(1), 7-10. https://doi.org/10.1080/09546559708427382

Podder, S. (2017). Understanding the Legitimacy of Armed Groups: A Relational Perspective. *Small Wars & Insurgencies*, 28(4-5), 686–708. https://doi.org/10.1080/09592318.2017.1322333

Prucha, N. (2016). Is and the jihadist information highway–projecting influence and religious identity via telegram. *Perspectives on Terrorism*, 10(6), 48-58. https://www.jstor.org/stable/26297705

Qian, G., Sural, S., Gu, Y., & Pramanik, S. (2004, March). Similarity between Euclidean and cosine angle distance for nearest neighbor queries. In *Proceedings of the 2004 ACM symposium on Applied computing* (pp. 1232-1237). https://dl.acm.org/doi/abs/10.1145/967900.968151?casa_token=ClIaxKR-wYIAAAAA:YYGqVr5f7Irwn8IhOnpZr--TMke5a2C1VkMwku02GMYA06iwWWTkllHZwhZ0Y-lUXhsR7dRFStJl

Ranstorp, M. (1996). Terrorism in the Name of Religion. *Journal of international affairs*, 41-62. https://www.jstor.org/stable/24357404

Rapoport, D. C. (1997). To Claim or not to Claim; that is the Question—Always!. *Terrorism and Political Violence*, 9(1), 11-17. https://doi.org/10.1080/09546559708427383

Reuters. (2020). Inside the plot by Iran's Soleimani to attack U.S. forces in Iraq. *Reuters*. https://www.reuters.com/article/us-iraq-security-soleimani-insight-idUSKBN1Z301Z

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, 91, 1-40. https://doi.org/10.18637/jss.v091.i02

Rogers, M. (2003). The psychology of cyber-terrorism. In A. Silke (Ed.), *Terrorists, Victims and Society* (pp. 72-99). Wiley.

Rorie, M. L. (2008). *Communicating through violence: An application of rational choice theory to terrorist claims of responsibility*. University of Maryland.

Rozema, R. (2008). Urban DDR-processes: paramilitaries and criminal networks in Medellín, Colombia. *Journal of Latin American Studies*, 40(3), 423–452. https://doi.org/10.1017/S0022216X08004392

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1), 68. https://doi.apa.org/doiLanding?doi=10.1037%2F0003-066X.55.1.68

Salehyan, I. (2010). The Delegation of War to Rebel Organizations. *The Journal of Conflict Resolution*, 54(3), 493–515. https://doi.org/10.1177/0022002709357890

Schlichte, K., & Schneckener, U. (2015). Armed groups and the politics of legitimacy. *Civil Wars*, 17(4), 409-424. https://doi.org/10.1080/13698249.2015.1115573

Schultz, K. A. (2010). The Enforcement Problem in Coercive Bargaining: Interstate Conflict over Rebel Support in Civil Wars. International Organization, 64(2), 281–312. https://doi.org/10.1017/S0020818310000032

Seligman, L. (2021, February). Rocket attack in Iraq marks third in one week. *Politico.* https://www.politico.com/news/2021/02/22/rockets-us-embassy-baghdad-470868

Sha'abani, M. N. A. H., Fuad, N., Jamal, N., & Ismail, M. F. (2020). kNN and SVM classification for EEG: a review. In A. N. K. Nasir, M. A. Ahmad, M. S. Najib, Y. Abdul Wahab, N. A. Othman, N. M. Abd Ghani, A. Irawan, S. Khatun, R. M. T. R. Ismail, M. M. Saari, M. R. Daud & A. A. M. Faudzi (Eds.), *InECCE2019*, (pp. 555-565). Springer.

Shehabat, A., Mitew, T., & Alzoubi, Y. (2017). Encrypted jihad: Investigating the role of Telegram App in lone wolf attacks in the West. *Journal of Strategic Security*, 10(3), 27-53. https://www.jstor.org/stable/26466833

Shehabat, A., & Mitew, T. (2018). Black-boxing the black flag: anonymous sharing platforms and ISIS content distribution tactics. *Perspectives on Terrorism*, 12(1), 81-99. https://www.jstor.org/stable/26343748

Shultz, R. H., Farah, D., & Lochard, I. V. (2004). Armed groups: a tier-one security priority (ADA435193). *INST FOR NATIONAL SECURITY STUDIES US AIR FORCE ACADEMY CO*.https://apps.dtic.mil/sti/citations/ADA435193

Smith, M. A., Shneiderman, B., Milic-Frayling, N., Mendes Rodrigues, E., Barash, V., Dunne, C., ... & Gleave, E. (2009). Analyzing (social media) networks with NodeXL. In *Proceedings of the fourth international conference on Communities and technologies* (pp. 255-264).

Singhal, A. (2001). Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.*, 24(4), 35-43. http://160592857366.free.fr/joe/ebooks/ShareData/Modern%20Information%20Retrieval%20-%20A%20Brief%20Overview.pdf

Staniland, P. (2012). Between a Rock and a Hard Place: Insurgent Fratricide, Ethnic Defection, and the Rise of Pro-State Paramilitaries. *The Journal of Conflict Resolution*, 56(1), 16–40. https://doi.org/10.1177/0022002711429681

Staniland, P. (2015). Militias, ideology, and the state. *Journal of Conflict Resolution*, 59(5), 770-793. https://doi.org/10.1177/0022002715576749

Suchman, M. C. (1995). Managing Legitimacy: Strategic and Institutional Approaches. *The Academy of Management Review*, 20(3), 571. https://doi.org/10.2307/258788

Thurber, C. (2014). Militias as sociopolitical movements: Lessons from Iraq's armed Shia groups. *Small Wars & Insurgencies*, 25(5-6), 900-923. https://doi.org/10.1080/09592318.2014.945633

Tracy, K., & Robles, J. S. (2013). *Everyday talk: Building and reflecting identities* (2nd edition). Guilford Press.

Van Leeuwen, T. (2001). What is authenticity?. *Discourse studies*, 3(4), 392-397. https://journals.sagepub.com/doi/pdf/10.1177/1461445601003004003?casa_token=N9agUblXrcEAAAAA:nvCY2hzcDsY26oCEf7Z4bYi58assYQaOyWXfvKuyyRLhs1Ert4IQU5JT0Cw7IuGztFCX6DpzPOEO

Walther, S., & McCoy, A. (2021). US Extremism on Telegram. *Perspectives on Terrorism* (Lowell), 15(2), 100–124.

Wang, J., & Dong, Y. (2020). Measurement of text similarity: a survey. *Information*, 11(9), 421. https://www.mdpi.com/2078-2489/11/9/421

Wan Mohd Nor, M., & El-Muhammady, A. (2021). Radicalisation and Paramilitary Culture: The Case of Wanndy's Telegram Groups in Malaysia. *In Militarization and the Global Rise of Paramilitary Culture* (pp. 95-122). Springer.

Ward, M. (2020). Walls and cows: social media, vigilante vantage, and political discourse. *Social Media+ Society*, 6(2), 2056305120928513.

Watanabe, K. (2020). Marimo multi-lingual stopwords collection. *Github.* https://github.com/koheiw/marimo

Weimann, G. (2004). www.terror.net: How Modern Terrorism Uses the Internet [Special Report]. https://www.usip.org/sites/default/files/sr116.pdf

Weimann, G. (2010). Terror on facebook, twitter, and youtube. *The Brown Journal of World Affairs*, 16(2), 45-54. https://www.jstor.org/stable/24590908

Weimann, G. (2015). Terrorist migration to social media. *Geo. J. Int'l Aff.*, 16, 180. https://heinonline.org/HOL/LandingPage?handle=hein.journals/geojaf16&div=24&id=&page=

Weinstein, J. M. (2006). *Inside rebellion: The politics of insurgent violence*. Cambridge University Press.

White, J. (2020). *Terrorism and the Mass Media* [Occasional Paper]. https://static.rusi.org/terrorism_and_the_mass_media_final_web_version.pdf

Williams, P. (2008). Violent non-state actors and national and international security. *International Relations and Security Network*, 25. https://www.files.ethz.ch/isn/93880/vnsas.pdf

Zenn, J. (2020). Chronicling the Boko Haram Decade in Nigeria (2010-2020): distinguishing factions through videographic analysis. *Small Wars & Insurgencies*, 31(6), 1242-1294. https://doi.org/10.1080/09592318.2020.1776582