Executive Summary

This study provides an overview of the current knowledge on Al Qa'ida and an initial assessment and characterisation of the Bin Laden Archive

In 2017, the United States Central Intelligence Agency (CIA) disclosed approximately 470,000 files recovered in Abbottabad (Pakistan) during the 2011 raid on Osama Bin Laden's compound ('the Bin Laden Archive').[1] According to data provided by the CIA on its website,[2] the Bin Laden Archive ('the Archive' henceforth) comprises a wide array of original files from devices collected during the Abbottabad raid that are presumed to have belonged to Osama Bin Laden and other occupants of the compound he lived in. Files included in the Archive include (clustered according to file type and content):

- More than 72,000 image files accounting for approximately 7 GB of data;
- More than 18,000 text files accounting for approximately 16 GB of data;[3]
- More than 24,000 Microsoft Offices files converted into PDF format, accounting for approximately 12 GB of data;
- More than 11,000 audio files accounting for approximately 30 GB of data; and
- More than 10,000 video files accounting for approximately 162 GB of data.

Despite the potential insights that this Archive stands to offer, limited research into its data and materials has been published to date in the public domain. More broadly, publicly available research conducted so far has entailed a qualitative review of only a selected number of files and documents. In December 2018, the Research and Documentation Centre (*Wetenschappelijk Onderzoek- en Documentatiecentrum*, WODC) of the Dutch Ministry of Justice and Security (*Ministerie van Justitie en Veiligheid*) commissioned RAND Europe to conduct a study (WODC Project Number: 2995) aiming to:

1. Produce an inventory of current knowledge on Al Qa'ida and of completed and ongoing research on the Bin Laden Archive (Phase I); and

2. Conduct an initial assessment and characterisation of the Bin Laden Archive (Phase II).

---

[1] Collins, L. (2012) 'The Abbottabad Documents: Bin Ladin's Security Measures.' *CTC Sentinel* 5 (5).

[2] For further information please see Central Intelligence Agency (2017) 'November 2017 Release of Abbottabad Compound Material.' As of 3 July 2020: https://www.cia.gov/library/abbottabad-compound/index.html

[3] This includes a mix of Microsoft Word and PDF files.

Phase I of the study focused on conducting a literature review and interviews with experts on the study topics. It also entailed a Phase II feasibility assessment, to assess the extent to which the methods and approaches identified during the project planning stage could be employed in Phase II.

Phase II of the study focused on undertaking human-based and machine-based analysis of the different file types comprising the Archive, namely image, audio, video and text files, as well as the so-called Bin Laden journal - a handwritten journal included in the Archive that was reportedly written by Osama Bin Laden and other occupants of the Abbottabad Compound.

Both Phases sought to answer a set of research questions, which are presented in Table S.1 alongside the project phase in which they were primarily addressed.

Table S.1 Study research questions for Phases I–II

	Research questions	Study phase
	What is the state-of-the-art understanding as discerned from academic and grey literature of:	
1a	Al Qai'da's ideology and motives	I
1b	Al Qai'da's organisation, including its relations with external branches and affiliated groups	I
1c	Al Qai'da' strategy, tactics and modus operandi including as regards political, military, and propaganda activities	I
1d	The phenomenon of *Jihadi* terrorism	I
2a	What research efforts have been conducted or are ongoing to analyse the Bin Laden Archive?	I
2b	To what extent have these efforts exhausted the potential for insights and findings to be generated through an analysis of Bin Laden Archive data and files?	I
2c	Would it be feasible to apply the proposed Phase II research methods on data included in the Bin Laden Archive?	I
3	How can data and files included in the Bin Laden Archive be characterised, categorised and clustered?	II
4a	To what extent can different clusters and subsets of the Bin Laden Archive yield relevant insights on Al Qa'ida's ideology, organisation, and strategy?	II
4b	To what extent can clusters and subsets of the Bin Laden Archive offer relevant insights on the phenomenon of *Jihadi* terrorism and the threat this poses to the West in general and the Netherlands in particular?	II

The study mapped out the existing knowledge regarding Al Qa'ida's historical trajectory, ideology, strategy and organisation

Al Qa'ida's ideology is based on Salafism, however the importance of the group's ideology should not be overstated

Al Qa'ida's ideology revolves around the concept of *Jihad* and is characterised by a strong anti-US and anti-Western sentiment. The key ideological elements and principles undergirding Al Qa'ida have reportedly remained stable throughout the years. However, a number of nuances and conflicting views can be observed within the Salafi Jihadist movement itself and within the organisation that have had practical implications for the ways in which Al Qa'ida has structured itself throughout the years as an organisation; and in the ways in which it has operated in different contexts and areas of operations.

In addition to elements of Salafi thought, Al Qa'ida's ideology draws upon selected parts of Islamic scholarship and traditions to build narratives and rhetoric reaching into long-standing and deeply held views and grievances among Muslim communities. Furthermore, the group is seen as building on the teachings of several scholars whose works and writings are used to provide legitimacy in the eyes of potential supporters and recruits.

Finally, scholars focusing on Al Qa'ida and other Jihadist groups also emphasise the importance of not overstating Al Qa'ida's ideological basis and coherence given the tendency that this and other *Jihadist* groups have to adjust theological and ideological arguments (i) to fit operational realities and needs, rather than the other way around; and (ii) in light of changing context, needs, and surviving members and scholars active in the organisation.

Al Qa'ida has evolved significantly since the early 2000s moving towards a decentralised, networked and global movement

Al Qa'ida was established in the 1980s in Afghanistan by Osama Bin Laden and a close cohort of battle-hardened jihadists who had originally fought the Soviet invasion of Afghanistan, forming what was initially a strongly hierarchical organisation. Al Qa'ida's core leadership has been based in Afghanistan and in the areas formerly known as the Federally Administered Tribal Areas (FATA) of Pakistan. The US invasion of Afghanistan in 2001 led to a critical erosion and destruction of Al Qa'ida's physical infrastructure and organisational capacity.

Within the literature, there appears to be a general consensus that Al Qa'ida's organisation has evolved significantly since the early 2000s, moving from being a single, centralised entity to a more decentralised, networked and global movement. In particular, literature reviewed suggests a strong consensus for conceptualising Al Qa'ida's organisational structure as comprising of a central core and of a global network of affiliate groups connected to the core by different relations and arrangements.

In this regard, a number of debates have been identified within the literature regarding, for example, the level of control that Al Qa'ida Central exerts over its affiliates, and the extent to which the formal command and control structure linking Al Qa'ida Central and its affiliates has remained effective following the death of Osama Bin Laden in 2011. Political instability and uprisings occurring throughout the Greater Middle East region from 2011 onward offered opportunities for Al Qa'ida and like-minded organisations to exploit

power vacuums and political transition periods to establish stronger footholds and presence. In Syria, an unprecedented internal split emerged between Al Qa'ida Central and the leadership of the Islamic State. This ultimately led to the Islamic State's separation from Al Qa'ida.

The existing literature captures the development and transformation of Al Qa'ida's organisation; while discrepancies exist in some places, there is wide consensus on the organisation's historical trajectory. Many sources view this transformation to have been borne out of necessity (due to losses incurred and a changing strategic environment), whilst others view it as part of a deliberate long-term strategy for ensuring the organisation's global reach.

No clear consensus exists among scholars concerning Al Qa'ida's strategic timelines and short-to-medium term objectives

Within the literature, scholars have formulated different approaches to analysing and conceptualising Al Qa'ida's strategy. Overall, some consensus emerges from the literature around the overarching goals and objectives which include: awakening consciences and inspiring Muslims across the world to join *Jihad*; targeting apostate regimes across the Muslim world and leading to their downfall; confronting Western countries and their allies to weaken their standing and solidity; and, re-establishing a global Caliphate and achieving final victory. Most notably, however, documents retrieved from Al Qa'ida so far, and data available in the public domain, do not provide a comprehensive understanding of the ways in which different strategic objectives and goals are prioritised and pursued by the group.

Available literature on Al Qa'ida's strategy also reveals a tension within the group's approach to prioritising targets for the 'near' and 'far' enemies. Some sources provide empirical evidence suggesting that in recent years, Al Qa'ida has been mainly focused on operations against the so-called near enemy and that its strategy has shifted towards localised goals and population-centric approaches, rather than high-profile attacks designed to instil fear in the so-called far enemy. Other scholars suggest that the line between these two targeting approaches has become increasingly blurred in the 21st century as Al Qa'ida and other *Jihadist* groups have begun to attack both near and far enemies in conjunction with and dependent on opportunities.

As regards other strategic enablers and aspects of Al Qa'ida's strategic and operational work:

- Training camps have played a significant role in the history of the organisation, and particularly during its early years leading to the formation of a cadre of members and supporters. Limited information is available about current Al Qa'ida's training practices and facilities and these appear to be primarily the concern of local and regional chapters, rather than an element of Al Qa'ida Central's work.

- Propaganda has played, and continues to play, a pivotal role for Al Qa'ida, facilitating the recruitment of Jihadists across the globe and boosting the organisation's ranks.

- Limited conclusive information is available with regard to the role of women in Al Qa'ida. The available sources of primary data suggest that women are expected to support and encourage their husbands in pursuit of jihad and raise their children to fight for the same cause in the future.

Prior to the emergence of the Islamic State, Al Qa'ida was an unrivalled dominant actor in the global *jihadi* movement

Since its inception, Al Qa'ida has transformed from a small, relatively unknown and regionally focused organisation, to a global organisation with a quasi-monopoly over the Jihadist space, and then to one that is faced by a competitor, the Islamic State, which has proven more effective in generating resources, recruiting and inspiring fighters and uniting affiliates through the use of technology.

The literature reviewed over the course of Phase I indicates that the context in which Al Qa'ida as an organisation operates has changed drastically from its early years. In connection to discussions of Al Qa'ida's global influence relative to that of the Islamic State, many sources reflect on the organisation's changing structure and relations with its affiliates and its struggles to ensure that the wider *Jihadist* movement continues to pursue a coherent strategic direction overseen by Al Qa'ida's leadership.

Looking at the present context, the literature reveals disagreements among scholars regarding the extent and nature of Al Qa'ida's relative decline since the emergence of the Islamic State. Some hold that since its establishment, the Islamic State has made rapid progress in dismantling Al Qa'ida's leadership of the global *jihadi* movement and has thus come to dominate the ideological and strategic space. Others suggest that the core organisation and wider network of Al Qa'ida have remained resilient and its 'brand appeal' continues to resonate with and influence extremist groups worldwide.

The Bin Laden Archive represents a resource of partially untapped potential

The study team conducted a systematic search for publications and research focusing on Bin Laden Archive data. The majority of publications identified made only limited or passing references acknowledging the existence of the Bin Laden Archive, but did not conduct any analysis of its data. In six instances,[4] the study team identified relevant publications which were further reviewed due to their having data included in the Archive. In all such instances, research conducted on data included in the Bin Laden Archive and on the Bin Laden Bookshelf, a previous, incomplete release of the Archive, entailed a qualitative review of a limited sample of files and documents. Overall, the small number of articles and publications identified, combined with the methodological limitations of endeavours undertaken so far on the Bin Laden Archive, suggest that there is significant scope for further research and analysis of these data.

---

[4] Lahoud, N., S. Caudill, L. Collins, G. Koehler-Derrick, D. Rassler & M. al-`Ubaydi (2012) 'Letters from Abbottabad: Bin Ladin Sidelined?' Combating Terrorism Centre at Westpoint; Collins, L. (2012) 'The Abbottabad Documents: Bin Ladin's Security Measures.' *CTC Sentinel* 5 (5); Roggio, B. (2017) 'Analysis: CIA releases massive trove of Osama bin Laden's files.' *The Long War Journal*; Gartenstein-Ross, D. and N. Barr (2018) 'How Al-Qaeda Works: The Jihadist Group's Evolving Organizational Design.' *Current Trends in Islamist Ideology* 23: 66–138; Grace, E. (2018). 'A dangerous science: psychology in Al Qaeda's words.' *Dynamics of Asymmetric Conflict* 11(1): 61–71; Grace (2018) 'Lex talionis in the twenty-first century: revenge ideation and terrorism.' *Behavioral Sciences of Terrorism and Political Aggression* 10(3): 249–263.

The study sought to characterise the Bin Laden Archive

File types were used as the characteristics around which to build macro clusters of the Archive to further investigate and characterise

A total of four clusters were identified: images, audio, video, and text. The study team adopted a sample-based analysis approach for the image, audio and video clusters that leveraged stratified samples to ensure that different file types and sub-clusters would be considered in the work. As for the text cluster, the study team adopted a machine-enabled approach to conduct a mapping and characterisation of the entirety of text files available, as well as a qualitative analysis of the 'Bin Laden journal', a handwritten journal included in the Archive that was reportedly written by Osama Bin Laden and other occupants of the Abbottabad Compound.

The image sample categorised over the course of Phase II activities suggests that the image cluster might have limited relevance to the study objectives. Two exceptions are worth highlighting. First, within the image cluster, a subset of images appears to originate from devices in use in the Archive, or from individuals who then shared photos and images with the compound inhabitants. However, these images appear to be of limited relevance in the context of the study due to their focus primarily on children, animals and the spaces and landscapes in and around the compound, rather than topics connected to the study. Second, a number of handwritten and printed letters and documents appear as scans within the image cluster. However, a review of the letters and documents included in the sample did not yield significant novel insights relevant to the study.

A qualitative review of audio transcripts generated under the study indicate that the majority of audio files included in the Phase II sample focus on religious topics, although not exclusively or predominantly from an extremist perspective. Recitations of the Qur'an, anasheed, lessons and sermons are particularly prevalent across all sample strata. More broadly, all of the recordings included in the study sample appear to be public and non-sensitive or private in their content; this also applies to a sizable proportion of randomly selected files excluded from the study sample due to the exclusion criteria employed.

A qualitative review of video transcripts generated under the study – as well as a review of a second sample of videos in their original format – indicate that the majority of video files included in Phase II samples focus on topics connected to religion and terrorism. This does not include, however, several videos that touch on topics or originate from sources that led to their exclusion from the study sample. In that regard, it is worth noting that a wide array of television broadcasts, cartoons and other non-relevant materials were identified during the preparation of the study video samples.

A machine-enabled corpus linguistics analysis of the texts included in the Archive led to the identification of clear patterns within the available data. In particular, the RAND-Lex[5] analysis of the Archive's text cluster led to the identification of four categories of content:

- First, a sizable segment of the text clusters was comprised of files that were edited by the CIA, as well as corrupt and unreadable files.

---

[5] RAND-Lex is a proprietary suite of analytic tools developed by RAND researchers to perform rigorous and complex text analytics and machine learning.

- Second, a number of text clusters identified in the Archive were comprised of newspapers copies, as well as copies of individual online articles and media publications from different periods of time.

- Third, a number of clusters identified in the text analysis were comprised of publications of different nature, covering religious themes and topics. This group of clusters includes both pedagogical materials, as well as publications covering advanced notions and topics within Islamic jurisprudence and theology.

- Fourth, a number of clusters that were identified comprised of documents covering issues related to jihadism, terrorism, extremism, politics and international affairs. Most interestingly, these clusters were comprised of public sources and documents downloaded from a wide range of websites and sources, but also documents that appear to be either of a confidential nature, or to have originated and been authored directly by occupants of the compound, or by individuals living in hiding who had been in contact with them.

A qualitative review and analysis of the Bin Laden journal highlights a number of themes discussed in the document with the main being on the developing situation across Arab countries during the so-called Arab Spring. Particularly in the first half of the journal, the entries are focused on detailing the latest developments of political turmoil across different countries in the Middle East and North Africa region,. The journal also suggested that Bin Laden was considering thinking about 'public relations' activities and creating content for wider release in response to the Arab Spring, including developing a recorded statement.

The journal also touches on wider ideological, strategic and organisational aspects of Al Qa'ida at the time at which it was written, although these are not the focus, and do not provide much additional information to what already known and discussed in existing academic literature.

The text cluster is the one most likely to yield relevant insights on Al Qa'ida's ideology, organisation and strategy and on the broader phenomenon of Jihadi terrorism

The image, audio and video clusters appear to have only a limited potential to help generate new knowledge and insights on Al Qa'ida and on the related phenomenon of Jihadi terrorism. The text cluster of the Archive appears to be the most promising cluster that could help generate novel insights and knowledge about Al Qa'ida. This is due to the cluster being comprised of personal, sensitive and private documents authored by individuals living in the Abbottabad compound or by other Al Qa'ida senior personnel living in hiding and in contact with those in the compound.

Little to no materials were found that had a specific relevance or reference to the Netherlands

The material with a connection to the Netherlands includes a picture of a former Member of the European Parliament at a time at which this individual was still an MEP. The image was captured from what appears to be a televised programme or interview. There was also a website banner with Dutch text about board games. Neither of these items were deemed to be of relevance.

The files within the text cluster should be further characterised and prioritised in future in-depth studies on the Bin Laden Archive

Based on the findings gathered in the initial assessment and characterisation of the Bin Laden Archive, the study team holds the view that analysing selected clusters and subsets of the Archive could generate novel insights and knowledge through access to primary sources so far not considered in scholarly literature. In particular, the study team recommends prioritising an in-depth analysis of selected segments of the text cluster. While, as noted above, there is material of interest in other clusters (notably the handwritten letters within the image cluster, and certain elements of the videos cluster), the overall proportion of potential material of interest within these clusters appears to be lower than that of the text cluster.

Furthermore, and in light of possible time and resource constraints that need to be taken account when conducting future studies on the Bin Laden Archive, clustering the text files would enable researchers to prioritise files of interest, and conduct a more in-depth analysis of the most relevant files in an efficient and time effective manner.