

RESEARCH NOTE

# An Easy-to-Use Search Tool for the Abbottabad Compound Material Archive

Janna Mantua,\* Carrington Metts

Volume XVII, Issue I  
March 2023

ISSN: 2334-3745

Osama Bin Laden's compound in Abbottabad, Pakistan was raided following his death in 2011. The documents retrieved from his compound became known as the Abbottabad Compound Material (ACM). The majority of ACM documents were declassified and released to the public in 2017. However, these documents have not been thoroughly mined by academic researchers despite potentially important information that may exist in the ACM. Academic researchers may be deterred from using the ACM because the documents are not easily accessible or because the documents are predominantly in Arabic. This paper presents a search engine specifically built for researchers that will enable the mining of ACM documents. The ACM search tool may facilitate critical terrorism- and extremism-related research related to al-Qaeda and related terrorist groups.

**Keywords:** al-Qaeda, Abbottabad Compound Material, data mining, search tool

---

\*Corresponding author: Janna Mantua, Institute for Defense Analysis (IDA), email: [jmantua@ida.org](mailto:jmantua@ida.org)

## Introduction

Al-Qaeda is a violent non-state extremist group that was founded by Osama Bin Laden and Abdullah Azzam in 1998. After its founding, the Saudi-led group quickly garnered religious extremist followers across North Africa and Central Asia and eventually headquartered in Afghanistan and Pakistan. The group is most notably known for its September 11<sup>th</sup> (9/11) attacks against the US, but the group committed smaller terrorist attacks prior to 9/11 and continues to commit attacks in North Africa and Central Asia. Under Taliban rule in Afghanistan, al-Qaeda and extremist groups that have stemmed from al-Qaeda (e.g. the Islamic State) have made a resurgence. Consequently, al-Qaeda and affiliated groups remain a threat to those who oppose them.

In 2011, Osama Bin Laden was killed by US forces in a bunker in Abbottabad, Pakistan. All documents from the bunker, including computer files, flash drives, images, and videos, were archived. These documents included mundane items, such as software manuals and newspaper clippings, but they also included potentially important items, such as correspondence between Bin Laden and his followers and propaganda-related documents. In 2017, the Central Intelligence Agency (CIA) and the Office of the Director of National Intelligence (ODNI) declassified a large portion of the documents found in the bunker and released them in an archive called the Abbottabad Compound Material (ACM).<sup>1</sup> Tens of thousands of documents from the ACM are now available for public use.

Despite a great wealth of information that may exist in this collection of documents, few academic researchers have utilised these data to quantitatively or qualitatively explore issues of critical importance (despite a few exceptions).<sup>2</sup> It is possible that researchers have been deterred from using this material because the documents are not easily accessible. The documents are in individual files on the CIA/ODNI website, and there are no tools available for searching through the documents. To use the files, one must either manually search through them or attempt to use improvised search tools (e.g. the often-poor search feature built into a computer's operating system). Mining the documents in this manner is impractical for researchers and may be a slow process. A search tool that facilitates this process would make the ACM more accessible to researchers.

Furthermore, researchers may be deterred from using the ACM because the documents are predominantly in Arabic. Individuals who do not have knowledge of Arabic may need to rely on translations to determine the contents of each document. In theory, this could be accomplished by feeding the documents through a translation portal (e.g. Google or Amazon Translate), searching for keywords or phrases in translated output, and trying to extract data from documents that seem relevant. However, this approach is not advisable, as it has recently been shown that feeding a large amount of text into freeware translators can result in "critical" information gaps in the translated output.<sup>3</sup> A more practical approach would be to identify a subset of documents that contain specific Arabic keywords and to subsequently perform a translation. This approach would ensure keywords of interest are not "translated away." A search tool that can screen all documents for specific Arabic keywords would enable this approach.

This article presents a newly developed ACM search tool (ACMST) that enables researchers to quickly and systematically mine Arabic ACM documents for keywords or topics of interest.

## Features of the ACMST

- **Speed.** The CIA/ODNI website houses 18,500 ACM documents. The documents range from one page to several hundred pages in length, totalling over 20GB of data. The ACMST can search all documents for keywords in a matter of seconds to minutes.
- **Savable search results.** Thousands of documents may match a desired search term. The ACMST includes a feature that enables the user to download a list of document titles that matched search terms. This feature supports the systematic processes necessary for rigorous research.
- **Searches for exact and similar search terms.** Most Arabic words have a “root” or “base” word that represents a core concept. Derivatives of the word are different versions of the root word. For instance, in Arabic, “he wrote” and “book” are derivatives of a common root word that relates to writing. The ACMST allows the user to search for their exact search term and derivatives related to the chosen word. In other words, searching the Arabic word for “he wrote” will also identify documents with “book” and “I wrote.” Researchers who are not familiar with Arabic syntax and sentence structure may benefit from searching derivatives of their primary search term. Doing so would reduce the likelihood of missing potentially important documents.
- **No coding required.** Many political or social scientists do not have the opportunity to develop coding skills. Although the ACMST is based on Python coding, the user does not need to have any knowledge of coding in order to use the search tool.

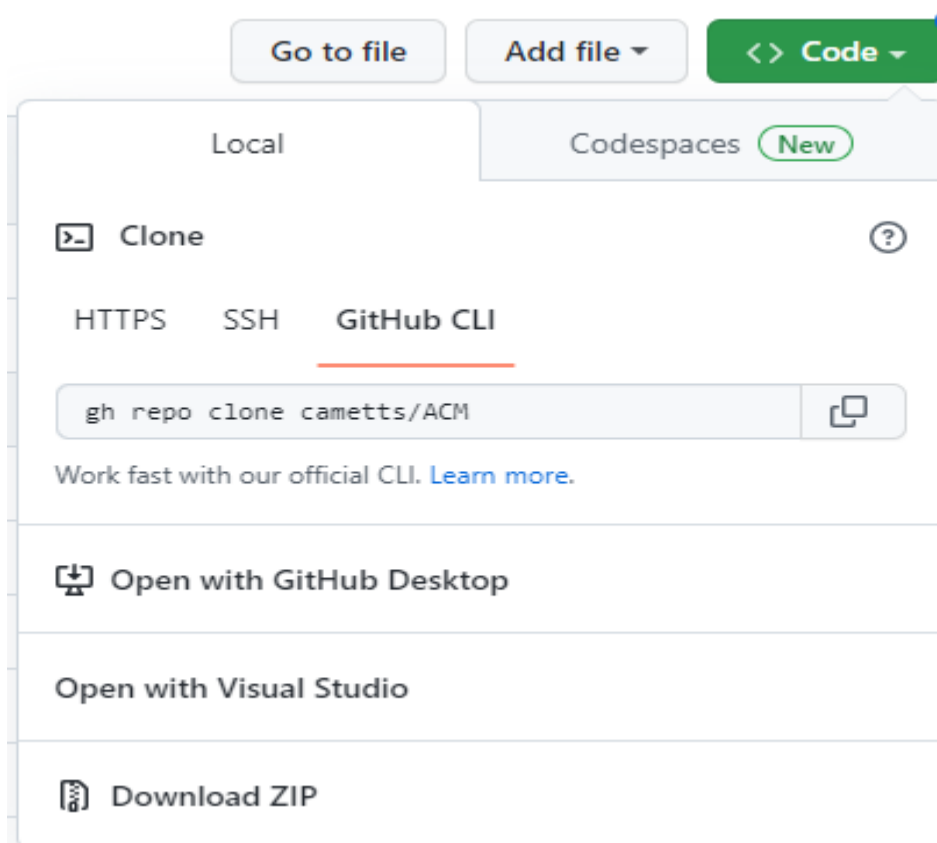
## How to use the search engine

The instruction manual for the search engine is in Appendix 1.

## Appendix 1: Installation Instructions

### *First-Time Installation*

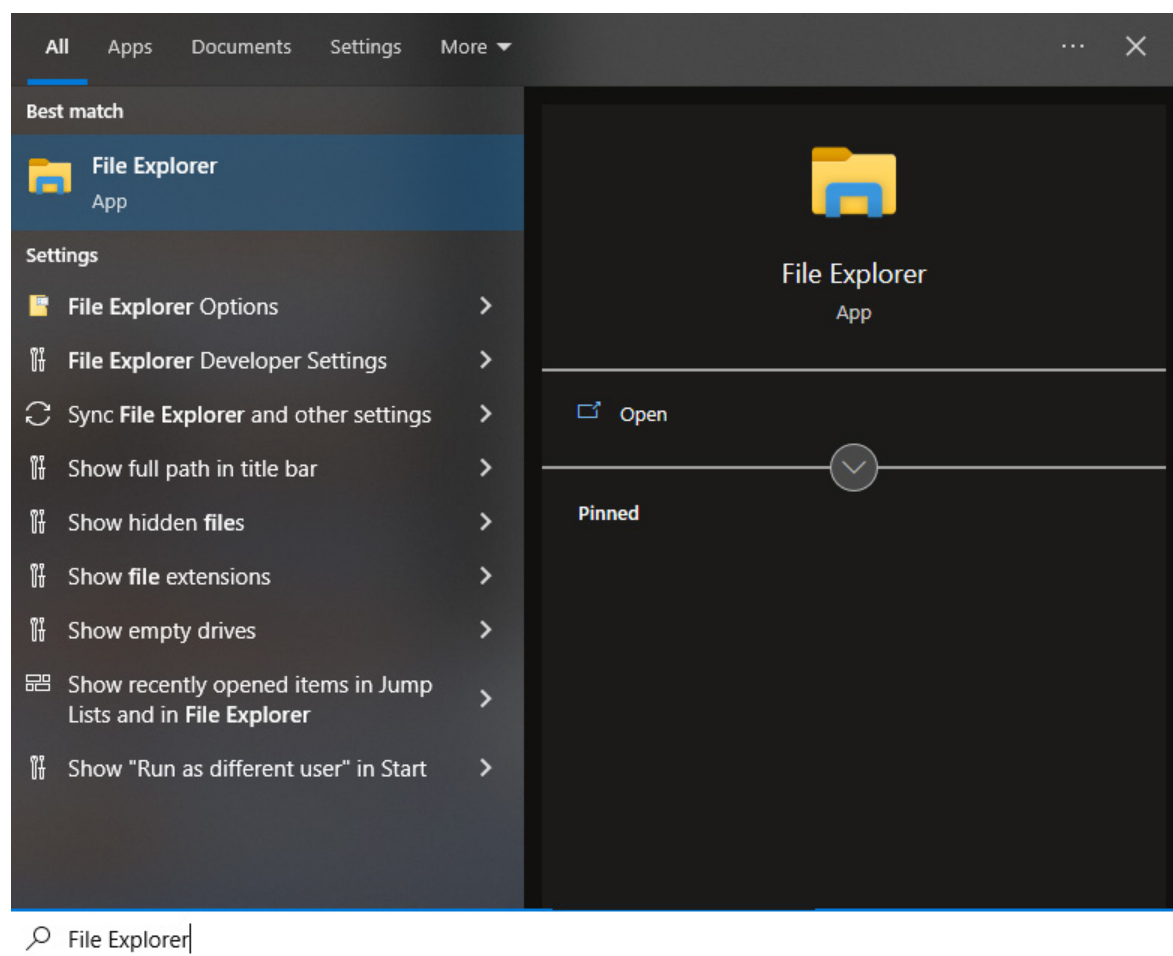
1. Be sure your computer is connected to the Internet for the duration of this process.
2. Open your Internet browser and navigate to the file location at <https://github.com/cametts/ACM>. Click the green “<> Code” button (near the top right corner) and click “Download ZIP”. Save the files to your Desktop, or another convenient location. Depending on the speed of your internet connection, this may take several minutes.



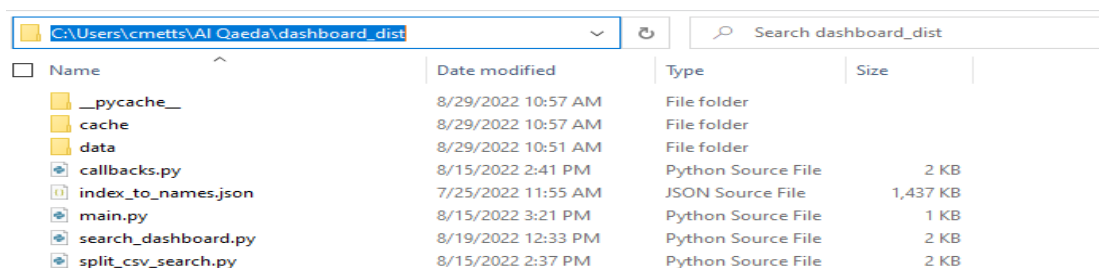
3. On your Desktop, find the file that was just downloaded and double-click it. When prompted, unzip the files to your Desktop. Depending on your system configuration and other previously installed software, you may be prompted to “unzip the files to Desktop”, “Extract to folder”, or similar command. Wait several minutes while the files are unzipped.
4. On your desktop, find the folder that was just created, which should be represented with a manilla folder icon. Hover the mouse over it and right-click. Select “Rename”, and rename the folder to “dashboard\_dist” (without quotation marks).
  - a. If you like, you may move the dashboard\_dist folder to a new location. However, do not remove any files from within the folder. Doing so will cause the search engine to become inoperable.
5. If you do not have Python installed on your system, download and install it from <https://>

[www.anaconda.com/products/distribution](http://www.anaconda.com/products/distribution). Follow the instructions to install “Anaconda.” Installing Anaconda may take up to 20 minutes.

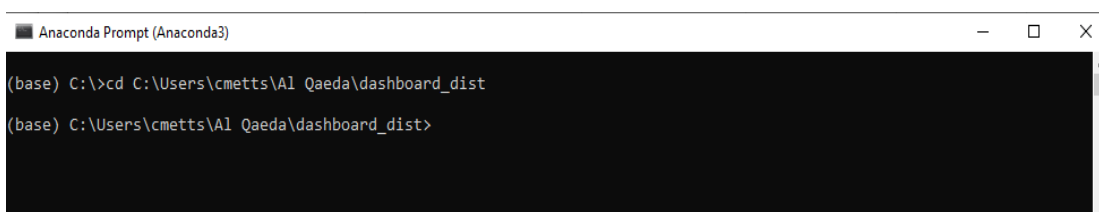
- NOTE: the search engine has been tested on Windows with Python versions 3.8 and 3.9. Other operating systems and Python versions may work, but they have not been tested.
6. Click on the magnifying glass icon in the lower left corner of your desktop (directly next to the Windows icon), or press the Windows key on your keyboard, which will pull up the search bar automatically. Enter “File Explorer” into the search bar (outlined in red in the image below). Open the File Explorer app.



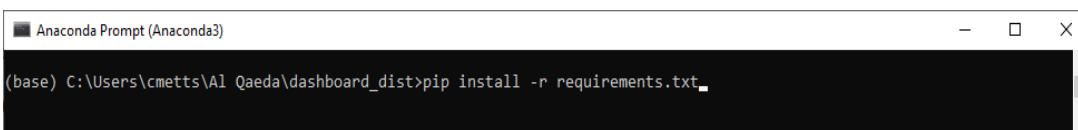
7. Navigate to where you saved the dashboard\_dist folder within File Explorer. If you followed steps 1-4 above and did not move the folder, it should be saved to your Desktop. Click anywhere inside the address bar (highlighted in red in the image below). The path to the folder will appear highlighted in blue. Copy the full path to the folder. This path should have the format of “C:\...\dashboard\_dist.”



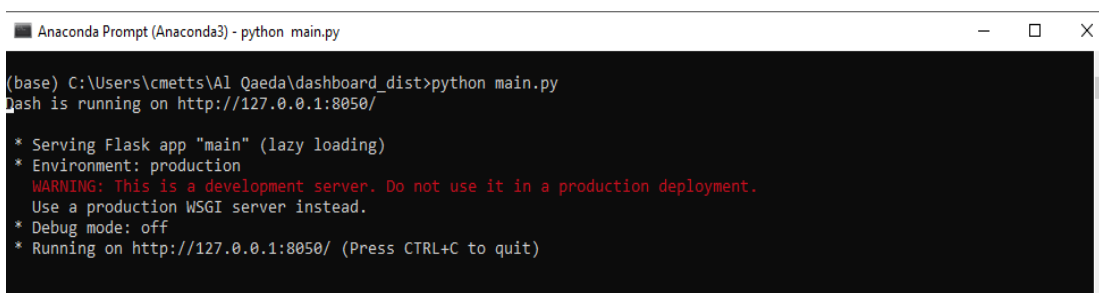
8. Click on the Start/Windows button (indicated with the Windows icon) and find the newly installed “Anaconda3” folder. Open the folder and click “Anaconda Prompt (Anaconda3).” Other Anaconda programs in that folder (e.g. “Anaconda Powershell Prompt”) will not work.
9. Find your cursor in the Anaconda program. Enter the text “cd” and paste the path that you copied in Step 2. Be sure there is a space between “cd” and the path that you are pasting (pictured in the top line below). Press Enter. A new line will appear (pictured in the bottom line below).



10. Enter or paste the text “pip install -r requirements.txt” (see picture below). Press Enter. After a short time, a number of new lines will appear on the screen as the packages install. This may take up to a few minutes. When installation has completed, your cursor will reappear.



11. It is now time to launch the search tool. To do so, enter or paste the text “python main.py”. Press Enter. After several minutes, the window will print a message, including a web address (pictured on the bottom line below). This address will begin with http://. Copy the web address and paste it into a web browser. Press Enter to navigate to the web page.



12. The search engine will appear in your browser window. There will be a blue status bar moving back and forth. The search engine is now ready to use. You can enter or paste any Arabic-language search term into the search bar (see image below). Names of files containing the word you searched (or related root words) will appear in a list below the search bar. You must leave the Anaconda Prompt window open while you are using the search engine. Closing it will cause the search engine to quit.
- You will know the search engine is working when the “Search” button is faded grey. It may take several minutes for a search to complete. If the search engine appears frozen for an extended period of time (more than 10 minutes), you may restart it by refreshing your browser.
  - If there are no files matching the search terms, a blank row will appear rather than a list of file names. If there is a long list of results, you can toggle through multiple pages using the arrows on the bottom left side of the page.
  - File names can be exported by clicking the “Export” button. The file names will automatically download into an Excel spreadsheet.

### ACM Searchable Database

A searchable database of Arabic-language documents from the ACM repository.

الأسرة      Search

Export

Files
00BC0F57427AF11646ABD51AB5AA90FB0f- نؤ Σ-¼ f06âøí
01161FA8E078CSEFEF172E03C2977DD0_الإغتيال_سنة_إحياء_سنة_الإغتيال
0170E29FBDA73657B26504D84097A49C_ALLETHAM
01959F7441B554AFCE8844F807421E7_0r0i f0ánf6 - f0n²² 027
028AEBE7F166BF4BAE3577427B465F7Cøfy

13. When you are finished searching, you may close your browser window and the Anaconda Prompt.
14. Now that you have identified files that contain your search words or phrases, return to the CIA/ODNI website and download those files using the identified title of the document.

## Launching the search engine during subsequent sessions

Once you have completed the steps above, your computer is configured to run the search engine. To use it at any future time:

- Follow Steps 6–9 in the instructions to navigate to the dashboard\_dist folder.
- Enter the command `python main.py` to launch the search engine (Step 11 in the instructions above). Navigate to the web address that is printed in the Anaconda Prompt window.

## Troubleshooting

1. I was able to download Anaconda and install all requirements, but running “python main.py” is not working.

First, make sure you have navigated to the correct folder. The file path shown in Anaconda Prompt to the right of the “>” should contain the file path where you have saved dashboard\_dist. If that still does not work, try replacing “python” with “py”, “python3”, or “py3”. Different versions and system configurations have slightly different syntax.

2. I entered a command into Anaconda Prompt, but nothing is happening.

Try pressing Enter. You will know that your command finished running successfully when you see your cursor blinking to the right of the > symbol. If nothing happens after pressing Enter, be patient! It may take several minutes to install the requirements and launch the search engine.

3. The search engine appears to be frozen.

If you see the blue status bar moving back and forth and the “Search” button is faded, that means the search engine is working. Be patient! It may take several minutes to find all results for a given word. If that is not the case, try refreshing the web page in your browser.

*Janna Mantua is a research staff member at the Institute for Defense Analyses (IDA). She has a PhD in behavioural neuroscience from the University of Massachusetts - Amherst and is completing an additional degree in military studies with a concentration in irregular warfare. At IDA, she contributes to and leads projects on counter-terrorism, influence operations, human-system interaction, and wargaming.*

*Carrington Metts is a data science fellow at the Institute for Defense Analyses (IDA). She has a Master of Science degree in business analytics from the College of William and Mary. Her work at IDA encompasses a wide range of topics, including modelling and simulation, natural language processing, and demographic analyses.*

## Endnotes

1 Online at [https://www.cia.gov/library/abbottabad-compound/index\\_documents.html](https://www.cia.gov/library/abbottabad-compound/index_documents.html).

2 For example: Jacopo Bellasio, Sarah Grand-Clement, et al., "Insights from the Bin Laden Archive," RAND Corporation (2021), [https://www.rand.org/pubs/research\\_reports/RRA109-1.html](https://www.rand.org/pubs/research_reports/RRA109-1.html); Emma Grace, "A Dangerous Science: Psychology in Al Qaeda's Words," *Dynamics of Asymmetric Conflict* 11, no. 1 (2018): 61–71; Nelly Lahoud, "Al-Qa'ida's Contested Relationship with Iran." *New America Foundation* (September 7, 2018), <https://www.newamerica.org/international-security/reports/al-qaidas-contested-relationship-iran/>; Bryce Loidolt, "Al-Qaeda's Iran Dilemma: Evidence from the Abbottabad Records," *Studies in Conflict & Terrorism* (2020): 1–28, <https://www.tandfonline.com/doi/abs/10.1080/1057610X.2020.1780011?journalCode=uter20>; and Bryce Loidolt, "Were Drone Strikes Effective? Evaluating the Drone Campaign in Pakistan Through Captured al-Qaeda Documents," *Texas National Security Review* (Spring 2022), <https://repositories.lib.utexas.edu/handle/2152/97115>.

3 Zakaryia Almahasees, Samah Meqdadi, and Yousef Albudairi, "Evaluation of Google Translate in Rendering English COVID-19 Texts into Arabic," *Journal of Language and Linguistic Studies* 17, no. 4 (2021): 2065–2080, <https://files.eric.ed.gov/fulltext/EJ1325432.pdf>.

# About

---

## Perspectives on Terrorism

Established in 2007, Perspectives on Terrorism (PT) is a quarterly, peer-reviewed, and open-access academic journal. PT is a publication of the International Centre for Counter-Terrorism (ICCT), in partnership with the Institute of Security and Global Affairs (ISGA) at Leiden University, and the Handa Centre for the Study of Terrorism and Political Violence (CSTPV) at the University of St Andrews.

## Copyright and Licensing

Perspectives on Terrorism publications are published in open access format and distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License, which permits non-commercial reuse, distribution, and reproduction in any medium, provided the original work is properly cited, the source referenced, and is not altered, transformed, or built upon in any way. Alteration or commercial use requires explicit prior authorisation from the International Centre for Counter-Terrorism and all author(s).

© 2023 ICCT

## Contact

E: [pt.editor@icct.nl](mailto:pt.editor@icct.nl)

W: [pt.icct.nl](http://pt.icct.nl)



Universiteit  
Leiden

