|  |  |
| --- | --- |
| ../Desktop/static1.squarespace.png | Getting Started with Media Cloud  *Mastering the basics of analyzing news media attention, language, and representation* |

**TABLE OF CONTENTS**

[What is Media Cloud? 1](#_Toc515569916)

[What can you do with Media Cloud? 2](#_Toc515569917)

[Questions Media Cloud has helped answer 3](#_Toc515569918)

[Some technical notes on data collection 4](#_Toc515569919)

[Two research approaches: general searches and question & answer 5](#_Toc515569920)

[General search approach 5](#_Toc515569921)

[Question and answer approach 8](#_Toc515569922)

[Filling out your research worksheet 10](#_Toc515569923)

[Writing your query search string 11](#_Toc515569924)

[Examples of Media Cloud queries with their corresponding research questions 12](#_Toc515569925)

[Determining the right sources 13](#_Toc515569926)

[How many sources are enough? 14](#_Toc515569927)

[A note about U.S. collections 14](#_Toc515569928)

[Need more help with sources? 14](#_Toc515569929)

[Setting up a comparative query in Explorer 15](#_Toc515569930)

[Analyzing your results 15](#_Toc515569931)

[What to try if you’re getting no results, or “odd” results 15](#_Toc515569932)

[Attention 15](#_Toc515569933)

[Language 17](#_Toc515569934)

[**Representation** 18](#_Toc515569935)

[Refining your query 18](#_Toc515569936)

[Augmenting your source data 18](#_Toc515569937)

[Supporting your conclusions 19](#_Toc515569938)

[Other Media Cloud tools 19](#_Toc515569939)

# What is Media Cloud?

Media Cloud is an online news media research platform that allows you to analyze news stories from over 60,000 online sources, or publications, from around the world, at the local, regional, and national level, and in about [t**wenty different languages**](https://mediacloud.org/support/languages). The stories available to search are mostly digital-only text stories or the digital version of printed news stories, dating back as far as 2009, though most sources were added very recently (2016 or later).

Some television, radio, and podcast media are also available in the form of their text transcripts published by their online sources. Media Cloud’s technology allows you to search the full texts of any publicly available news article, and in some cases, some introductory text of stories behind paywalls (in journals and by subscription-only sources). While Media Cloud includes some social media analytics features and integration capabilities, it is not a tool meant to deeply analyze social media; Media Cloud is for analyzing published online news media available in text form.

Media Cloud includes two tools—Explorer and Topic Mapper—as well as a sources library (Source Manager) you can browse. We recommend you use Explorer to learn the basics of researching news media online with Media Cloud.

# What can you do with Media Cloud?

Media Cloud can be used by anyone interested in doing in-depth research into news media coverage around a topic of interest. Nonprofits and foundations, content creators, journalists, activists, and more have used Media Cloud to support their work. Media Cloud has helped them track coverage of their organization, track media around crisis events, identify media partners, assess impact of programs, and evaluate coverage of topics prior to creating new media.

Media Cloud can be used to answer research questions dealing with news media **attention, language, representation,** and **influence**.

* **Attention** refers to the amount of coverage given to a topic over time (i.e., *how much* the media is covering a topic).
* **Language** refers to the most frequently used terms in news media content on a topic (i.e., *what* is being discussed by the media and *how* they are discussing it—which words they use most).
* **Representation** refers to the most frequently mentioned people, organizations, and countries in media for a topic (i.e., *who* comprise the most representative figures and countries in the media’s coverage)
* **Influence** refers to which media producers are doing most of the reporting on a topic, narratives that spread from one media producer to another, as well as the amount of referencing, citing, or sharing of the topic’s media by the general population (i.e., which media are covering a topic the most, and how often do other sources and readers feel compelled to reference or share the media)

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Media Cloud’s Explorer helps you analyze Attention, Language, and Representation, whereas Topic Mapper lets you analyze all four of the above categories (including Influence).  At the core of Media Cloud is a powerful search engine—much like any popular web search engine that provides a simple search, or query, field—but Media Cloud also provides insightful data visualization outputs, plus these distinct advantages:   * Media Cloud is designed specifically for news analysis. It searches *only* news media, so you don’t have to filter out other search results. * Unlike web search engines, Media Cloud does not prioritize or personalize your search results based on your search history or other content optimization tactics. As a result, you can perform a more objective analysis of data. * You can create side-by-side comparisons of media coverage by country or region, different topics, communities of readership, and more.   Media Cloud was initially developed to explore whether the internet actually enabled new voices and perspectives to enter public conversations, or if larger news sources remained the most influential. Helping people discover lesser-heard and diverse perspectives is still a key part of the value Media Cloud provides. | |  |  | | --- | --- | | **Key Media Cloud terms  to remember:** | | | **Sources (or Source Media)** | Online news media publications available to search in Media Cloud | | **Attention** | *How much* the media is covering a topic | | **Language** | *What* is being discussed and *how* the media is discussing it—words they use the most | | **Representation** | *Who* comprise the most representative figures and countries in media coverage | |

**Questions Media Cloud has helped answer**

**Media Cloud Question:**Who were the key information authorities for coverage of the Ebola epidemic, and did that impact the global health response?

|  |  |
| --- | --- |
| **../ebola-549471_960_720.jpg** | **Answer:** *Mainstream media did not frequently reference international health authorities, and so these authorities were largely unsuccessful in directing the narrative about the Ebola outbreak. Digital media sources shared scientific information infrequently. Further, significant public engagement and sharing was directed towards stories about Ebola infections in the U.S., rather than in  West Africa.* |

**Media Cloud Question:**Which key themes has the documentary, *Life on Parole* (featured on *Frontline*)*,* added to the news media conversation about re-entry into society after incarceration?

|  |  |
| --- | --- |
| OTO.jpeg | **Answer:** *An analysis of news media about prisoner  re-entry and parole prior to the documentary’s release revealed a focus on individual incidents of parole violation and recidivism.* Life on Parole *added new dimensions to this conversation: the relationship between early childhood trauma and difficulties with re-entry, links to specific policies, such as methadone treatment in prisons, and a focus on job and education opportunities as part of the re-entry process.* |

**Media Cloud Question: What language and framing does the Indian media use in discussing gender issues? And what gaps can be identified as potential impact areas for journalists or nonprofits working in the area?**

|  |  |
| --- | --- |
| ../../Downloads/india-416777_1280.jpg | **Answer:** *A study of various gender-focused topics in Indian media (rape, child marriage, selective abortion, and dowry) found that the Indian media focused on presenting the issues primarily as crimes. The stories attributed blame to unsafe public spaces or criminal identities, rather than drawing attention to root causes such as social biases, patriarchy, preference for sons over daughters, and so forth. Journalists and nonprofits could focus on shifting the frame of conversation towards the role of social norms.* |

|  |
| --- |
| **More Media Cloud question and answer examples,  with analysis details, are available by clicking here.** |

Some technical notes on data collection

Media Cloud uses story-based indexing for search and analysis purposes. This means that when Media Cloud searches for a query match, it searches an entire story. For example, the query [health AND wellness] will return all stories including the words “health” and “wellness” anywhere in the entire story.

In Explorer, many outputs in the analyses are based on a sampling of stories within the search results. Details about the sampling methodology are available throughout Explorer, by clicking the Learn More links near each analysis area.

# Two research approaches: “general searches” and “question & answer”

|  |  |  |  |
| --- | --- | --- | --- |
| There are two ways to approach searching for the right news media data to analyze in Media Cloud. The first is to just input one or more keywords to search, and explore the results—a “general search.” It’s easy to input anything in Media Cloud—but calling this approach “easy” overall would be misleading, as you may find it challenging to know what to look for in Media Cloud’s outputs when you search this way. For beginners, the general search approach is best for searches with a limited topical focus, such as tracking news coverage on your organization, or on your own media or published content.  If you’ve never used Media Cloud, follow the instructions in the next section, **General search approach**. This section will help you get familiar with the basics of the Explorer interface and set up your first search.  General search approach  To try a general search for tracking news coverage on your organization:   1. **Register for a Media Cloud account.**   On the landing page at [**https://explorer.mediacloud.org**](https://explorer.mediacloud.org), scroll down to the registration/ log in box, and click **REGISTER NOW**. Registering allows you to save your searches and other work. | |  | | --- | | **Overview of the Media Cloud research process** | | **1.** Choose a research approach: “general search,” or “question and answer.”   1. If you’ve chosen a “question and answer approach,” prepare a research worksheet. The worksheet will help you finalize your research question and query or queries. 2. In Explorer, enter the query you’ve prepared, dates, and select the appropriate media sources to search. 3. Run your query. 4. Analyze the results. 5. Refine your query. 6. Repeat steps 4-6 as often as needed, until you’ve answered your question or satisfied your curiosity. | |

After submitting the registration form, you’ll be e-mailed an activation link. Click the link to finish your registration.

1. **Enter your organization’s name into the search box on the Explorer landing page, and click** **Search**.  
     
     
     
   The page that appears is a combination *full query pane* (for search query refinement), and analysis page. The full query pane looks like this:  
     
   

The analysis section shows results for the last month of stories published by sources in U.S. Top Online News 2017, a collection of top news websites of August 2017 in the U.S., according to data from comScore, Activate and Alexa. This collection, and the date search range of the last month, are Media Cloud defaults that you can see selected in the screenshot above.

**NOTE:** You can also get to the full query pane by clicking the **Search** button on the Explorer landing page, without typing anything into the Search box. You’ll get results without any query language filtering—that is, you’ll see results for analyzing everything in U.S. Top Online News 2017 from the last month.

1. If your organization is not based in the United States, or is known mostly on a local or regional level, use the **Select media** function to select different source media.   
     
   If you’re interested in a longer or shorter timeframe, you can edit the default date range under   
   **For dates** too.



For help on selecting the right sources for your search, see the section on [**Determining the right sources**](#_Determining_the_right_1).

After you click **Add Media**:

* Choose a listed Featured collection, if appropriate, by clicking the “+” button to the right of the description; or
* Search for a collection or sources by:

1. Choosing a filter for your source search (Geographic, All, or individual Sources) on the left,
2. Entering a keyword for the collection or source name on the right; and
3. Clicking **Search**.

When you find a source or collection you want to add, click the **+** button to the right of the description. You can edit your selections (by clicking the **X** button to remove unwanted selections) in the lower left corner of the dialog window. Click the **OK** button to add your source selections to your query.

Back on the query page, you can click the **X** button to remove unwanted media selections.  
  
Click the **Search** button to run the search again with your edited source media and dates.

1. **Click the Save Search button to save your search.**   
     
   You’ll then be able to use the **Load Saved Search** button to load this search again any time.
2. **Analyze your results.**Use the Attention, Language, and Representation tabs to view the three categories of outputs. More information about using the data visualization for analysis are available in the online help by clicking **Learn more** under each section’s introduction on the left, and are also available in the  
   [**Analyzing your results**](#_Analyzing_your_results) section of this guide.   
     
   Again, this search method is helpful just for analyzing late-breaking news coverage of a specific entity, and looking for anything out of the ordinary—words used, high attention peaks, unexpected people and places mentioned. This method will give you a general overview of how the media is discussing the entity that matters most to you.



Question and answer approach

The more interesting analyses in Media Cloud tend to be ones conducted to investigate and answer specific news media questions. The question and answer approach also helps you determine what to focus on in the many slices of analysis available in your Media Cloud results.

To determine if this approach is right for you, it first helps to get an idea of the types of questions Media Cloud can help answer. Examples are mentioned in the [**Questions Media Cloud has helped answer**](#questions)section, and below.

There are two basic types of research questions Media Cloud can help answer: **comparative questions**, and **questions focused on one aspect of analysis**.

Comparative questions

Comparative questions are probably the most common type of research question: investigating news coverage using two or more different points of comparison, such as different countries, different timeframes, different entities (or subjects), different U.S. partisan readership sources, or different subtopics. Some examples:

**Media Cloud question:** How does media “hype” from the last two years compare with the truth and reality around artificial intelligence in the U.S.?

* ***To answer this:*** *You might run two queries: one using the U.S. Top Online News collection, which searches mainstream, “general interest” articles and op-eds, most likely to include emotionally provocative (and sometimes deceptive) content, and one query on scientifically-focused publications.*

**Media Cloud question:** Which kinds of issues, angles and slants related to my topic has the target audience for my documentary been exposed to?

* ***To answer this:*** *You might run separate queries for your topic on different media types, such as traditional print publication versus digital-native publications, or source collections by communities, if applicable (for example, African American, Tech blogs, Parenting blogs). Similarly, looking at coverage in ideologically diverse media such as Left versus Right partisan collections in the U.S. can help you compare differences in agenda and framing driven by varying audiences. You can often get more information on audience data for different publications from resources outside Media Cloud (listed under* [Augmenting your source data](#_Augmenting_your_source) *in this guide).*

**Media Cloud question:** How does the coverage of climate change compare between conservative media in the U.S. and the rest of the world?

* ***To answer this:*** *You might write several queries with different keyword aspects of climate change, run them using U.S. partisan-affiliated collections, and run identical ones using different major media sources around the world. It might be interesting to compare and contrast media in nations affected most by recent climate-related disasters versus ones who haven’t, or developing vs. developed nations.*

Questions focused on one aspect of analysis

Other kinds of questions focus on one aspect of analysis offered in Media Cloud, such as Representation.   
  
**Media Cloud question:** Which people, organizations, and places are being covered most often in news media about poverty?

* ***To answer this:*** *You can run one query using the keyword “poverty” (possibly coupled with other search terms, which you’ll learn about in the section on* ***How to structure your query****) and focus on the results in the Representation tab.*

Formulating your own research question

The examples above may have given you an idea for a research question, or you may still need help formulating one that you can answer with Media Cloud. Either way, completing a research worksheet, as described in the next section, can help you finalize a question to answer in Media Cloud, and turn that question into a query that will help you get the best results. Writing good search queries, selecting the right sources, and knowing what to look for in the outputs are the most challenging aspects of using Media Cloud. The research worksheet can help.

# Filling out your research worksheet

A research worksheet can help with the [**question and answer approach**](#questionanswer) to Media Cloud. You’ll be able to get much more meaningful outputs when you put a bit of preparation into your search beforehand, by brainstorming and listing details and key terms on various aspects of the topic you’re interested in. These details will also help you create or finalize a research question, and translate that question into a query you’ll enter in Media Cloud.

**1. [Open the research worksheet.](https://docs.google.com/spreadsheets/d/1fsU3ZM9Hnln76Yf7WEnuYeYFv2R7XxlMBLVttR7T52E/edit" \l "gid=286595416)** [You’ll see an example filled in for you in the second column.](https://docs.google.com/spreadsheets/d/1fsU3ZM9Hnln76Yf7WEnuYeYFv2R7XxlMBLVttR7T52E/edit" \l "gid=286595416)

1. Start filling out the worksheet at the top of the third column by naming your topic of interest.
2. **In the next row down, type in a question, if you have one in mind.** It can be broad or vague, or you can skip this step if you’re not sure what might be most interesting to focus on yet.
3. **Follow the instructions in the left column of the worksheet for each row.**   
     
   In row 10, you’ll probably want to browse Source Manager for interesting sources and source collections that you’ll want to select for your query. See the section on [**Determining the right sources**](#_Determining_the_right) for details.
4. **Use the brainstorming you did in rows 4-10 to finalize the research question you’d like to answer, and type it in row 11.** Questions consist of some combination of people, places, locations, time ranges, media types, readership types (such as U.S. partisan-focused readerships), events, or responses. Read through the **question and answer examples** and combine different elements of what you’ve brainstormed.
5. **In row 12, you’ll translate this question into a query search string that you’ll enter into Media Cloud.**Note that part of your query inputs will include date ranges (row 8) and sources (row 10). Row 12 refers to the query search string only, which you’ll enter in the query text box in Explorer.   
     
   **Important:** see the section on [**Writing your search string**](#_Writing_your_query) for instructions on how to form your query.
6. **Set up and run your query in Media Cloud.** Then, proceed with [**analyzing your results**](#_Analyzing_your_results_1) and [**refining your query**](#_Refining_your_query).
7. **(Optional) After you’ve answered your research question using Media Cloud, you can revisit this worksheet to type it in row 13, and add supporting evidence and screenshots of the visualizations in row 14.**

# Writing your query search string

You have probably have heard of Boolean logic, which defines connectors like AND, OR, and NOT to focus a search on retrieving the results you really want. Crafting the right query using Boolean standards is critical in Media Cloud, and part of the process of getting the best results is trying different queries with different combinations of keywords using Boolean logic. You’re looking to be as broad in your search as possible without picking up too much “noise”—that is, unrelated content.

The following table is a guide to Boolean logic query terms. “String” refers to series of search terms you enter.

|  |  |  |  |
| --- | --- | --- | --- |
| ***If you want to retrieve stories with:*** | ***Use:*** | ***Explanation:*** | ***Example:*** |
| any single word in your string | **OR** | Useful when interchangeable terms will yield an equally valid result. | **producer OR director** returns results for stories in which either term might be used |
| all of the words in your string | **AND** | Useful if the words in your string don’t always appear in the same order. | **workplace AND harassment AND sexual** returns results for stories in which these keywords are used, but not necessarily next to one another |
| phrases (words that must appear in a certain order) in your string | quotation marks (**“”**) | By default, Media Cloud assumes the equivalent of the Boolean **OR** between all words unless otherwise specified; quotation marks are useful for searching on an exact phrase. | **“climate change”** returns results for stories in which that exact phrase is used, rather than stories mentioning climate and change separately |
| hyphenated words in your string | quotation marks around the hyphenated word | Hyphenated words must be placed within quotation marks if it’s important that the hyphenated form only is searched upon; otherwise, Media Cloud will return results for the hyphenated form and two consecutive words without a hyphen | **“well-being”** returns results for “well-being,” but **well-being** returns results for well-being and well being |
| a mixture of the above | nested phrases/ parentheses | If your query is somewhat complex you will probably need to use parentheses to structure it and nest search terms. | **("key issue") AND (politics OR economy OR campaign)** will retrieve stories containing the phrase “key issue” AND any of the other three terms |
| several versions of a word with different fragments based on a core fragment | **\*** | Use the wildcard **\*** to represent more than one letter | - **key\*** retrieves results that include key, keys, keyboard, keystone, keynote, keynesian, keyword  - **\*spell\*** retrieves results that include spell, misspell, misspelling, spelling |
| several versions of a word with one different character based on a core fragment | **?** | Use the wildcard **?** to represent a single letter | **kee?** retrieves results that include kee, keen, keep, keel (but not keeping or keenly) CAN WE LIST A BETTER EXAMPLE? |
| punctuation | N/A | Media Cloud does not archive any other punctuation. | **@POTUS** and **#POTUS** will retrieve results for POTUS |

**Note:** searches are not case-sensitive, so capitalization will not affect your search results. (However, most people capitalize the Boolean terms **AND**, **OR**, and **NOT** to help distinguish them visually from keywords.)

Examples of Media Cloud queries with their corresponding research questions

The following queries demonstrate how Media Cloud research questions can translate into simple or complex Boolean search strings. Much of the complexity is dependent on the breadth of terms that may be used in isolation when referring to your topic. For example, the word “cancer” is always used in a discussion of cancer—no matter the type of cancer being discussed. It also is a core fragment of words like cancerous, cancer-causing, and so on. But sexual and reproductive health stories may include a wide range of terms, so you may need to list many to find all the stories you want on this topic.

|  |  |
| --- | --- |
| ***Research question:*** | ***Media Cloud query:*** |
| Which forms of cancer receive the most attention in news media? | Cancer\* |
| Which media sources cover the topic of vaccine-hesitancy the most? | "anti-vaccine" OR "anti-vax" OR ((anti OR against OR hesitant OR reluctant) AND vaccine\*) |
| Which organizations are referenced the most frequently in conversations about sexual and reproductive health in the United States? | Contraception:(((unintend\* OR unwanted OR unplan\* OR mistim\* OR teen\*) AND pregnan\*) OR (contracept\* OR condom OR "birth control" OR "abortion pill" OR "plan b" OR "e pill" OR "ec pill" OR "ella" OR "levonelle" OR "morning after pill" OR IUD\* OR vasectom\* OR spermicid\* OR (steriliz\* AND sex\*)) OR (sex\* AND (safe OR unsafe OR "without protection" OR unprotected)) OR ((STD OR STI OR UTI OR (disease AND transmi\*)) AND sex\*) OR ("sex education" OR "sex ed" OR "sex chat") OR ("family planning")) |

I WOULD SUGGEST DOING A SCREENSHOT/ VISUAL OF EXAMPLE 3, AND ADDING CALLOUTS TO DESCRIBE THE PURPOSE BEHIND DIFFERENT CONNECTORS/ PARTS, WHAT THEY ARE MEANT TO RETRIEVE, TO HELP ILLUSTRATE THIS ADVANCED USE OF BOOLEAN CONNECTORS

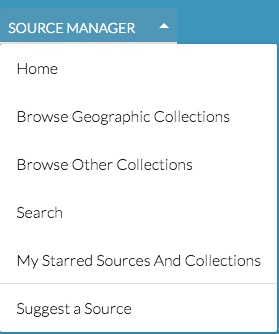
# Determining the right sources

Being confident that your search query is broad enough without including too much “noise” (unrelated aspects of a broad topic) will get you the best results in Media Cloud’s Explorer. You accomplish this partly by writing a good query, and partly by selecting the right sources. Source Manager lets you browse Media Cloud’s full database of sources and collections, which are groups of sources.

Using your research worksheet, you may have brainstormed events that take place in locations whose media you’d like to research. You may also like to compare media conversations by communities or special interest (such as Indigenous, parenting, U.S. partisanship). Use the brainstorming you’ve done on your worksheet to narrow your sources selection by thinking about considerations around dates covered, countries of publication, countries of readership, state and local perspectives, and community perspectives, or media type (like digital native or print).

To browse sources:

1. [**Open Source Manager**](https://sources.mediacloud.org/#/home) **by clicking its name in the menu bar at the top of the Explorer application window.**
2. **Click the triangle next to the Source Manager menu to reveal menu options:**



* + - **To browse collections of sources by country, use the Browse Geographic Collections menu option.** These source collections are based on country of publication.
    - **To browse specially curated collections, based on topics, communities, blogs, and more, use the Browse Other Collections menu option.** Note that popular, curated U.S. media collections are listed in the Other Collections section; these are based on U.S. readership and include sources published in countries outside the U.S.
    - **If you’re looking for sources in languages other than English, sources of a certain media type, or are interested in other advanced search parameters, use the Search menu option.**

1. **(Optional) After you’ve completed any of the step options under (b) above, click the name of any listed source or collection to review its details (such as sources included, date ranges, top words).** Then, if you want to save the source or collection to your own list of Starred Sources and Collections, click the star next to the name.

How many sources are enough?

It’s better to select multiple collections for your search if you have doubts about being comprehensive enough with your sources. However, including collections that don’t have enough relevance could introduce “noise” into your findings. So selecting sources is a balancing act. Major national collections are usually the best start for your search. For any other collection, make sure you click the name of the collection to open that collection’s page and read about the kind of sources and content you’ll be searching.

A note about U.S. collections

Some countries, such as the U.S., will have an especially wide variety of collections, many resulting from Media Cloud partnerships. These include the national and state level collections that you’ll find for the other countries in our database, as well as some U.S. partisanship-affiliated collections, rural collections, and more.

U.S. partisan-affiliated collections are of two types: partisan sets defined as right, center or left-leaning through partners (such as BuzzFeed), and “retweet partisanship collections,” categorized based on relative number of tweets of election story links by Clinton vs. Trump followers.

Need more help with sources?

If you’re still not finding what you want, click the **Support** menu item to get in touch with Media Cloud. You can even request the addition of a source to the Media Cloud database, if it’s especially important to your research.

# Setting up a comparative query in Explorer

If you’ve set up a [**general search query**](#generalsearch) in Explorer, you know the basics of query crafting. But if you have a comparative question you’re trying to answer, you’ll want to set up two or more comparative queries that run simultaneously, so you can compare and contrast the outputs easily.   
  
From the [**full query pane**](https://explorer.mediacloud.org/#/queries/search?q=[{%22label%22:%22%22,%22q%22:%22%22,%22color%22:%22%231f77b4%22,%22startDate%22:%222018-04-30%22,%22endDate%22:%222018-05-31%22,%22sources%22:[],%22collections%22:[58722749]}]&auto=true), use the **Add query** function to run multiple queries at the same time. The screenshot below shows two queries that have been set up, plus the **Add query** button.



When you set up and run multiple queries this way, you’ll get one analysis below the full query pane, with interesting compare and contrast features, such as the one below (in the Language tab), showing the top words of one query on the left, the other on the right, and top words common to both in the middle.



# Analyzing your results

After your results load, use Media Cloud’s categories of analysis to start exploring ways to interpret your results.

What to try if you’re getting no results, or “odd” results

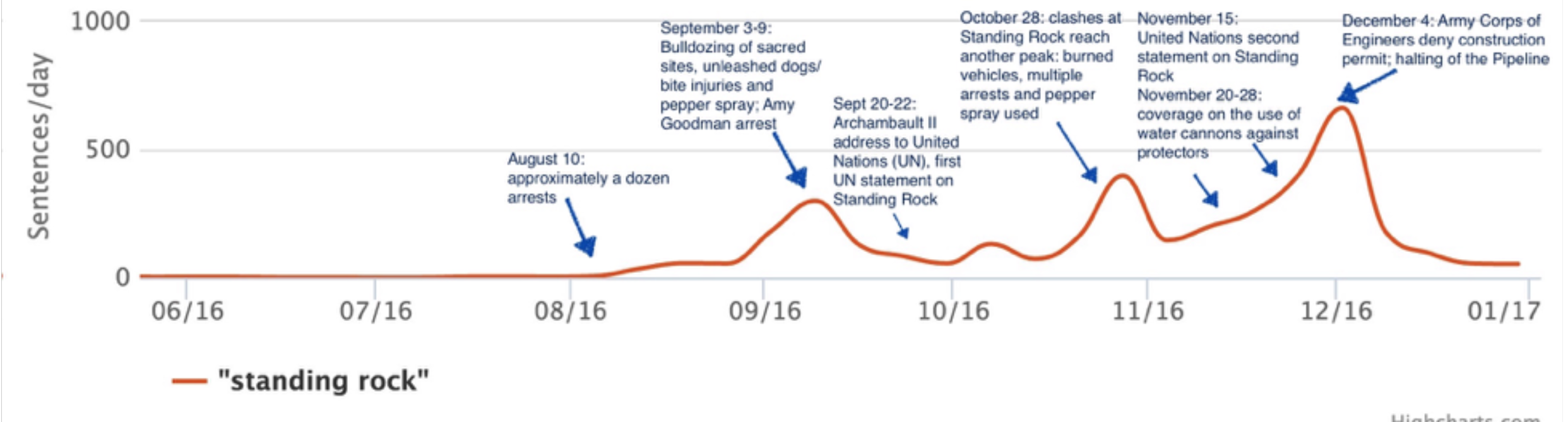
NATALIE

Attention

The Attention tab will be selected by default when your query finishes running, displaying results for the Attention analysis category.

Attention over time

Compare key event dates with amount of coverage. Some people even visually map dates over the Attention Over Time peak graph visualization (but you’ll need another graphic program to do this):



* Do you see the peaks you might expect in coverage around these dates?
* Try clicking a valley or peak before or after your key dates. Wait for new media details about that date to load below the graph upon your click (scroll down a bit to see these). Browse the stories and the word cloud, which shows the most frequently used words. What was the media saying before and after your key dates (whether they have corresponding peaks or not)?
* If you see a peak that’s not related to a key event, what else could be driving coverage? Again, to answer this, click the peak to browse stories and frequently used words to see if you can find clues in the ways the narrative was shaping, or discover other events that weren’t an obvious focus for your topic.

If you don’t get peaks or the level of peaks you might expect, it might be time to go back and brainstorm other key events that you may not be capturing—or simply widen your date range, if you’re unsure of key event dates related to your topic.

If you’re getting story samples that don’t seem to fit with your topic, try to think of ways to exclude that kind of content in your query using Boolean logic. You may also need to narrow your topic.

If you’ve run multiple queries to answer a comparative question, you’ll be able to view Attention Over Time peak lines with one another all in one graph, to compare and contrast amount of coverage in your defined time ranges.

Attention over time

Total Attention tells you the total number of stories matching your query in the sources and date range you’ve defined. View these numbers by mousing over the circles. They give you an idea of how important your topic was to the media in general during the timeframe you specified. Don’t be discouraged by lack of stories and data; often, this is a finding in itself.

If you’ve run multiple queries to answer a comparative question, you can compare Total Attention (total number of stories) visually for each query by looking at the size of the circles, and then reveal the exact number of stories for each by mousing over the circles.

Top themes

The results in this section are based on an automated categorical classifier based on a large sampling of stories in the *New York Times* (which is considered a broad resource). It auto-detects “themes,” or categories of similar content that exist in the group of stories you’ve found.

These can be helpful when an unexpected, or perhaps, seemingly unrelated theme is detected. In this case, you could infer that there was a significant amount of attention paid to something that was slightly off-topic, which may warrant further investigation of the stories.

It may also be helpful to view themes you may have expected that *weren’t* detected. You can [**browse the full list of themes here**](https://mediacloud.org/support/theme-list)**.**

Mouse over the circles to view the percentage of stories you’ve found with that detected theme.

Sample stories

This section lists a random sampling of the stories for your query, giving you easy access to browsing stories right from the screen. If you want full access to all the stories that have been found in Media Cloud for your query, use the Download Options function—you’ll be able to view a CSV file with a URL for each story. You may find that browsing the headlines of the full set of stories in the CSV gives you enough clues to answer your research question. But often you’ll want to read some of the text of the story itself.

Language

Under the Language tab, you’ll find results for Top Words, Word Space, and a comparison of Top Words if you’ve completed comparative queries.

Top Words  
  
Browse the word cloud for unexpected terms, or terms that might reveal more clues to help answer your question. Are there words that seem to express conflicting biases or slants? Often, more “contentious” angles reveal themselves in the word clouds.

Clicking a word in the word cloud adds it to your query with the Boolean connector, **AND**. You can think of this modified query as a sort of subtopic or one aspect of your overall topic, and may want to save it separately for comparison.

Word Space

This visualization helps you analyze patterns in the use of top words in the found stories. Here’s a key to the visualization:

|  |  |
| --- | --- |
| *Size + opacity* = | frequency of word appearance |
| *Distance from the center* = | variance in semantic usage of word (for example, a distance far from the center would indicate a word is used in a wide variety of phrases with varying meanings) |
| *Orange/ cone* = | words that are used in similar contexts or narratives |

The Word Space visualization can help give you a sense for angles that may dominate the set of stories you’ve found.

### Representation

Top People and Top Organizations

Under the Representation tab, you’ll find results on the people and organizations being discussed the most for your query. These are often subjects of stories, or are being referenced or quoted in your area of interest. Note that some figures, such as the President of the United States, often appear at the top of any list, simply because there is so much news in general where this person’s name will be mentioned at least once.

Some users of Media Cloud like to use these results to make lists of accounts to follow on social media.

Geographic coverage

Mouse over the map to view the percentage of stories mentioning each country. (**Note**: this map does *not* refer to origin of story publication.) What inferences might you draw by the countries the media has focused on for your area of interest?

# Refining your query

The process of refining a query can include many steps we’ve already reviewed:

* Combining different elements of your worksheet data (date ranges for different events, topical keywords, sources, comparative issues, media types, key people, and so on) to create and explore different queries with different sources.
* Browsing stories for additional key terms to add to your query, once you’ve discovered different potential angles.
* Clicking a word in the word cloud, under Language, to add an interesting new term with the Boolean connector **AND** to your query.

# Augmenting your source data

If you want to take your analysis a step further, you can often find helpful datasets at these various sources:

[Pew Research](http://www.pewresearch.org/topics/state-of-the-news-media/) (helpful for researching audience data per source)

[United Nations (for data and statistics by country)](http://www.un.org/)

[Centers for Disease Control (for health data)](https://www.cdc.gov/)

[Internet Archive](https://archive.org/index.php) (great for adding TV, movie, music sources)

[Diversity in Newsrooms](http://asne.org/content.asp?contentid=121) census

You can also use a major web search engine site to search identical queries (with restricted date ranges, usually available under Settings), and compare the results for gaps or further indications of influence from search engine algorithm rankings.

# Supporting your conclusions

When you’ve settled on an answer to your research question, write it out. Use the data visualizations in Explorer and citations from your story results to support your conclusions.

You can also further elucidate your answers by creating your own layered Attention peak graphs with dates, or network maps using a tool such as [Gephi](https://gephi.org/).

# Other Media Cloud tools

Explorer is a great tool to practice searching and analyzing online news media. Once you’ve developed a basic level of skill and confidence with techniques of querying and analysis, and you have more than a few days to wait for results, you can try Topic Mapper, which provides a way to dive even more deeply into online data and work with more types of outputs for analysis, including the analysis of influence.