

## Experiencing the September 11 Digital Archive, Using Omeka's Ngrams and Text Analysis Plugins

By: Jannelle Legg, December 2017

### Summary

*"This is not about piecing together what 'actually happened.' The whole world saw on their television screens how the buildings collapsed and were shattered into fragments. The stories indeed are themselves fragments that, when assembled, form a collected communal memory of the event, one that gathers and flows under and around and through the famous photographic and video images that have somehow become solid and clear in visual memory across the world. Precisely how the accounts of September 11, 2001 and similar archives actually represent communities when the only evidence of these communities is in the online records they create, is a current concern of archivists." [Paul Arthur, "Trauma Online"](#)*

The volume of the approximately 150,000 items in the *September 11 Digital Archive*, presents a challenge to researchers interested in learning about the experiences of 9/11 as told by the project's contributors. Exploring collections of this size is an overwhelming task. Even narrowing the content using the search box to filter items can yield thousands of results and undertaking a close reading of the entire archive is an unworkable project.

A distant reading approach allows scholars to look across a collection to question and make sense of a large number of objects. For this case study, I used two new Omeka plugins, [Ngram](#), and [Text Analysis](#), as research tools to explore the [September 11: Bearing Witness to History: Stories Submitted Online](#) collection, which includes 7,420 reflections submitted by visitors to the Smithsonian National Museum of American History "Bearing Witness to History" online exhibition. Both of these plugins operate at a macro-level of analysis, highlighting avenues for closer reading and interpretation. The Ngram plugin graphs the instances of a word or phrase in a corpus and the Text Analysis plugin delivers information on the corpus content. I used these tools to explore the ways that contributors described the experience of 9/11.

At the outset, I wanted to identify what might be learned about 9/11 by examining how people *made sense* of these events. Using the Ngram and Text Analysis plugins together, I found the following:

- Many contributors' experiences of 9/11 were mediated by communications technologies, like television and radio.
- Contributors conveyed information about the contexts of these experiences by referencing educational environments (schools, teachers) that offer some insight into the visitor demographics to the "Bearing Witness to History" online.
- The plugins were useful in facilitating a deeper, close reading of submissions by providing a manageable entry point to the exceptionally large collection.

Working from these results I was able to identify items that described or addressed mediated- or classroom-experiences, for closer study and examination.

### **Using the Ngram Plugin to Create a Corpus**

Before I could begin my analysis, first I created a corpus. Both the Ngram and Text Analysis plugins rely on a pool of texts referred to as a corpus. The corpus is defined by selecting a particular text field, the Text Element, and further filtered by a search query, date, and range.

After spending some time examining items from across the extensive archive, I focused my efforts on the *September 11: Bearing Witness to History: Stories Submitted Online* collection of reflections submitted by visitors to the Smithsonian National Museum of American History “Bearing Witness to History” online exhibition beginning in 2002. The collection provided a large pool of items (7421) gathered over a period of time (2002-2007). Each online submission was created digitally allowing the text to be read by the plugins, and the questions on the online form used by contributors provided a structure to the narratives. In particular, the first question, “*How did you witness history on September 11th? Share your experience*” prompted visitors to reflect and describe their experience as witnesses, offering an opportunity to engage with contributor experience (The online submission form included four prompts, [retrieved via Internet Archive](#)).

I configured the Ngram plugin to run on “NMAH Story:Story,” the appropriate text element for the *September 11: Bearing Witness to History: Stories Submitted Online* collection. Then, I created an item pool by defining a search query (in this case I did not want to limit my results and did not enter additional text to my search query string), sequence element (date entered), sequence type (date by year), and sequence range (2001-2007). After completing that process, my corpus of 7,420 items was available to analyze.

### **Using a Corpus with the Text Analysis Plugin**

The Text Analysis plugin relies on the [IBM Watson Natural Language Understanding](#) service to analyze the text in the corpus created. The service includes a number of features for exploring a corpus; entities, keywords, categories, and concepts. These are selected from the Analyze Corpus page. For this initial exploration, I wanted a broad overview of the text and I did not limit the features included. [It should be noted that the Watson NLU does incur a cost per analysis. To estimate the expense it is possible calculate the NLU item cost on the Add a Corpus page (for more information on how the cost is calculated see [documentation](#)). In this case, a corpus of this size with four features, was considered 2,693 items according to the IBM NLU service, and with a Standard Plan rate of \$0.003, incurred a cost of \$8.08.]

Owing to the differences in submission text (length and vocabulary, for instance) some features, like, sentiment and emotion scores, did not offer clear results. In these cases a close reading of the text, guided by the distant-reading results produced by the plugin, would be more useful for investigation. Additionally, many of the results produced by the Text Analysis plugin confirmed

that contributions focused on the subject of 9/11 (in other words, the results with the highest relevance were entities, concepts and terms associated with 9/11, like “Twin Towers”).

With this in mind, I approached the results of the Text Analysis plugin as an opportunity to surface themes and subjects pertaining to the experience of 9/11. I looked for those entities or concepts that offered insight on experience and occurred across multiple years. For instance, the entities and keywords features of the Text Analysis plugin unexpectedly highlighted media and education as significant subjects in the corpus. These results were further reinforced in the lists of categories and concepts.

The entities feature of the Text Analysis plugin offers a chance to look across a corpus, identifying people, events, places, and other entities mentioned. The tool is able to identify these entities and offer some insight on how they are described (emotion, sentiment) and the relevance of the entities to the entire corpus. As expected, the entities with the highest relevance in this corpus are associated with the places and events linked to 9/11: world trade center, new york, pentagon, twin towers, Bin Laden. However, among the results identified were entities that gave some indication of the contexts in which contributors experienced 9/11.

Entity	Type	Emotion	Sentiment	Count	Relevance
Twin Towers	Facility <ul style="list-style-type: none"> <li>Location</li> <li>TouristAttraction</li> <li>Building</li> <li>Skyscraper</li> </ul>	<ul style="list-style-type: none"> <li>Sadness: 0.626316</li> <li>Joy: 0.084401</li> <li>Fear: 0.624346</li> <li>Disgust: 0.415314</li> <li>Anger: 0.188017</li> </ul>	-0.369236	13	0.87485
New York City	Location <ul style="list-style-type: none"> <li>PoliticalDistrict</li> <li>GovernmentalJurisdiction</li> <li>PlaceWithNeighborhoods</li> <li>WineRegion</li> <li>FilmScreeningVenue</li> <li>City</li> </ul>	<ul style="list-style-type: none"> <li>Sadness: 0.582067</li> <li>Joy: 0.460527</li> <li>Fear: 0.536694</li> <li>Disgust: 0.094889</li> <li>Anger: 0.146463</li> </ul>	-0.539313	19	0.842733
Pentagon	Organization	<ul style="list-style-type: none"> <li>Sadness: 0.625662</li> <li>Joy: 0.077341</li> <li>Fear: 0.548359</li> <li>Disgust: 0.123106</li> <li>Anger: 0.171298</li> </ul>	-0.514324	12	0.774268
World Trade Center	Facility <ul style="list-style-type: none"> <li>Location</li> <li>BuildingComplex</li> <li>Skyscraper</li> </ul>	<ul style="list-style-type: none"> <li>Sadness: 0.629771</li> <li>Joy: 0.546108</li> <li>Fear: 0.126401</li> <li>Disgust: 0.067336</li> <li>Anger: 0.223667</li> </ul>	-0.493842	11	0.683296
America	Location	<ul style="list-style-type: none"> <li>Sadness: 0.618308</li> <li>Joy: 0.621988</li> <li>Fear: 0.558674</li> <li>Disgust: 0.489468</li> <li>Anger: 0.105665</li> </ul>	-0.652561	8	0.596203

Media entities were named in the corpus across multiple years. For instance, news sources CNN (2002, 2003, 2005, 2006) New York Times (2002, 2003), CNBC (2002), ABC (2007), and MTV (2003) were associated with the submissions. As well as news reporters Bryant Gumble (2002), DJ (2002), reporter (2004), Dan Rather (2006), and Tom Brokaw (2007).

Additionally, results also included references to educational environments; C.O. Wilson Middle School (2002), Principal (2002, 2003, 2004, 2006 ), Jenks High School (2002), Mr. Heasley (2002), Mrs. Curtis (2002), East Grand Rapids Middle School (2002), Mr. Wilson (2002), High School (2002), Sussex Middle School (2002), Hofstra University (2004), Superintendent (2005), Morning Announcements (2005), Mr. Leichtling (2006), and Erskine College (2007).

The keywords feature mirrored these results. Another useful feature of the Text Analysis plugin, keywords highlights relevant terms or phrases in the corpus, offering some insight on how the keyword text is described (emotion, sentiment) and the relevance of the keywords to the entire corpus. As with the entities results, the plugin produced highly relevant keywords expected of a corpus that describes experiences during the events of 9/11; world trade center, plane, twin towers, new york, and pentagon.

Keyword	Emotion	Sentiment	Relevance
world trade center	<ul style="list-style-type: none"> <li>• Sadness: 0.623836</li> <li>• Joy: 0.242702</li> <li>• Fear: 0.116451</li> <li>• Disgust: 0.085524</li> <li>• Anger: 0.533785</li> </ul>	-0.526138	0.904258
plane	<ul style="list-style-type: none"> <li>• Sadness: 0.609829</li> <li>• Joy: 0.104241</li> <li>• Fear: 0.209915</li> <li>• Disgust: 0.484354</li> <li>• Anger: 0.19063</li> </ul>	-0.554314	0.508847
twin towers	<ul style="list-style-type: none"> <li>• Sadness: 0.654289</li> <li>• Joy: 0.149586</li> <li>• Fear: 0.148484</li> <li>• Disgust: 0.071359</li> <li>• Anger: 0.168989</li> </ul>	-0.455975	0.473816
new york	<ul style="list-style-type: none"> <li>• Sadness: 0.546165</li> <li>• Joy: 0.528809</li> <li>• Fear: 0.102606</li> <li>• Disgust: 0.095589</li> <li>• Anger: 0.113941</li> </ul>	-0.418177	0.463939
people	<ul style="list-style-type: none"> <li>• Sadness: 0.669406</li> <li>• Joy: 0.107717</li> <li>• Fear: 0.161496</li> <li>• Disgust: 0.501736</li> <li>• Anger: 0.189729</li> </ul>	-0.2185	0.443579

The results of the keywords analysis also reinforced the observation made about entities in the corpus (see extracted results in the chart below). Again, the tool surfaced terms pertaining to news sources and news formats.

<b>2002</b>	tv (relevance .6) news (.59) radio (.44) little portable radio (.38) television (.38)
<b>2003</b>	news (.42) tv (.39) phone (.34) radio (.34) fake news reports (.33) cable TV feed (.33) cell phone rang (.33)
<b>2004</b>	news (.49) tv (.46)
<b>2005</b>	news (.46) tv (.44) TV station (.38) TV news reports (.38) television (.37) news people (.37) TV stations (.36) big TV screens (.36) local TV stations (.35)
<b>2006</b>	news (.45) local television news (.37) Sesame Street (.36) local news anchors (.36) horrible news (.36) tv (.36) radio (.36)
<b>2007</b>	phone ring (.74) Tom Brokaw (.72) TV watching (.71) ABC news (.70) TV screen (.7) replay (.54)

Text Analysis plugin also analyzed these texts for the categories and concepts features. The results from these features mirrored the findings of entities and keywords. In both cases, education and media emerged from the corpus as significant across the texts.

The Categories feature of Text Analysis categorizes the text using a five-level classification hierarchy. The tool evaluates the text content and assigns a category that is relevant to the subject matter described in the corpus.



Combining the results of the category text analysis into a single table reveals the trends in contributor's texts.

Label	2002	2003	2004	2005	2006	2007
/education/school	0.582805	0.55159		0.490977	0.38203	
/art and entertainment/movies and tv/television	0.419976	0.382951		0.399888	0.398203	0.427089
/art and entertainment/movies and tv/movies				0.396735		
/art and entertainment/radio	0.302997					
/law, govt and politics/espionage and intelligence/terrorism		0.338887				
/society/unrest and war					0.561007	
/family and parenting/children						0.405218
/sports/walking						0.358636

In the chart above, each of the categories from 2002-2007 is organized by year with the categorization score. In this case 0 means it is not confident, 1 means that it is confident. Note that the analysis does not produce confident results. Again, this is largely a result of the varied nature of submissions. However, as the table demonstrates, across time the submitted texts were categorized as containing subject matter relating to education and media. These findings reinforce the entities and keywords results.

Finally, the concepts feature of the Text Analysis plugin identifies high-level concepts not necessarily directly referenced in the text. In this case, the analysis appropriately identified that the corpus included indirect reference to the World Trade Center and the September 11 attacks.

**Text Analysis: Corpus Analysis** Q

Overview Entities Keywords Categories **Concepts**

### General Concepts

Identify high-level concepts that aren't necessarily directly referenced in the text.

#### Glossary

Concept  
Name of the concept

Relevance  
Relevance score for the concept ranging from 0 to 1. A 0 means it's not relevant, and a 1 means it's highly relevant.

Concept	Relevance
World Trade Center	0.958409
September 11 attacks	0.540865
High school	0.33737
Television	0.295509
Twin Towers	0.281996
The Two Towers	0.241463
Twin towers	0.236078
World Trade Center site	0.233549

Text Analysis by IBM Watson Natural Language Understanding

Looking across the results (from 2002-2207), the following table provides a comparison of all of the concepts identified in the corpora.

Concept	2002	2003	2004	2005	2006	2007
<u>World Trade Center</u>	0.958409	0.960324	0.96008	0.979958	0.941157	0.919261
<u>September 11 attacks</u>	0.540865	0.45329	0.643233	0.455474	0.948345	0.909722
<u>American Airlines Flight 11</u>		0.264269	0.349613	0.308315	0.580091	
<u>Twin Towers</u>	0.281996	0.192269	0.290345	0.42366		
<u>Empire State Building</u>		0.297824	0.323137			
<u>Television</u>	0.295509		0.363279	0.288862		
<u>Windows on the World</u>		0.257187			0.67416	
<u>The Two Towers</u>	0.241463			0.353833		
<u>Twin towers</u>	0.236078			0.332905		
<u>The Pentagon</u>					0.679224	0.785897
<u>High school</u>	0.33737					
<u>World Trade Center site</u>	0.233549					
<u>Keefe, Bruyette &amp; Woods</u>		0.231415				
<u>Minoru Yamasaki</u>		0.201367				
<u>What's Going On</u>			0.373345			
<u>Marvin Gaye</u>			0.357993			
<u>plane</u>				0.278373		
<u>World War II</u>					0.739047	
<u>World War I</u>					0.706687	
<u>Lehman Brothers</u>					0.556777	
<u>Attack</u>						0.724333
<u>Mind</u>						0.705025
<u>English-language films</u>						0.678803
<u>Attack!</u>						0.673933
<u>Tom Brokaw</u>						0.574268

Among the concepts identified by the plugin was “Television” and “High School.” Though the relevance scores are low (between .28 and .36), it does suggest, along with the other findings,

that it would be worthwhile to further explore media and education as contexts for how contributors experienced the events of 9/11.

The features of the Text Analysis plugin are useful for distantly reading a corpus. In this case they provide an overview of the content included in descriptions of contributor's experiences of 9/11. Of course, this overview should be interpreted and contextualized. First, the 7,420 texts included in the collection vary in length and number per year. This fact should be considered as results like relevance scores are evaluated. Similarly, the collection included submissions from contributors with different writing styles, who may have used different language to refer to similar experiences. For instance, a close reading of texts may collapse phrases like "twin towers," "two towers," "world trade center," "WTC," and "Trade Center" into a single entity, referring to the two World Trade Center Towers struck by airplanes during the events of 9/11. The text analysis tool, however, maintains these as distinct entities.

The emergence of news/media and educational contexts in these results is useful. They offer some insight into how and where contributors experienced the events of 9/11. Though the results don't definitively suggest that news media was the primary way people experienced the events of 9/11, they do surface that these subjects were included in descriptions provided. Close reading would be useful in further exploring these findings and investigating how media shaped these experiences.

Overall the results of the Text Analysis plugin directed my inquiry into the *Archive*. I was curious to explore the ways in which users described their experiences on 9/11. The features of this plugin reinforced that most of these experiences focused on the events and locations associated with the tragedy. These results also encouraged me to think about how news media (both print and television) may have shaped or contributed to the experience of that day.

### **Using a Corpus with the Ngram plugin**

The Ngram plugin offers another way to look across the contents the *September 11 Digital Archive*. Much like the Text Analysis plugin, users must first create and validate a corpus that draws on a particular Text Element and organizes results according to a search query, number sequence, and range. I used the corpus created from the *September 11: Bearing Witness to History: Stories Submitted Online* collection, which was created and validated in an earlier step.

The Ngram plugin utilizes text analysis to process the content of a large pool of texts. The text is evaluated and broken down into items, known as n-grams. Ngrams are items of speech, like words or phrases. For this plugin, the base item is a word, but in some ngram tools n-grams can focus on smaller units of language like syllables. The Ngram plugin evaluates the items in a corpus and produces information about them, like graphs, counts, and frequencies.

Users must select to create unigrams, bigrams, or trigrams. Unigrams are ngrams that organize and compare single items, like "car" and "flower." Bigrams are ngrams that evaluates pairs of

words, like “the car” or “pretty flower.” Trigrams organize corpus results by groups of three words, like “the car drives” or “a pretty flower.”

The process of creating ngrams may be time intensive and may only be initiated after a corpus has been validated. There are two options for exploring these results; ngram search and ngram frequencies.

### **Ngram Search**

The ngram search allows users to search the corpus for phrases according to a sequence. In this case, the sequence element was the date that contributors added to the collection, producing a graph of language frequency over time. With this plugin, users can choose to evaluate the frequency of a particular word or words in the entire corpus on a numeric scale.

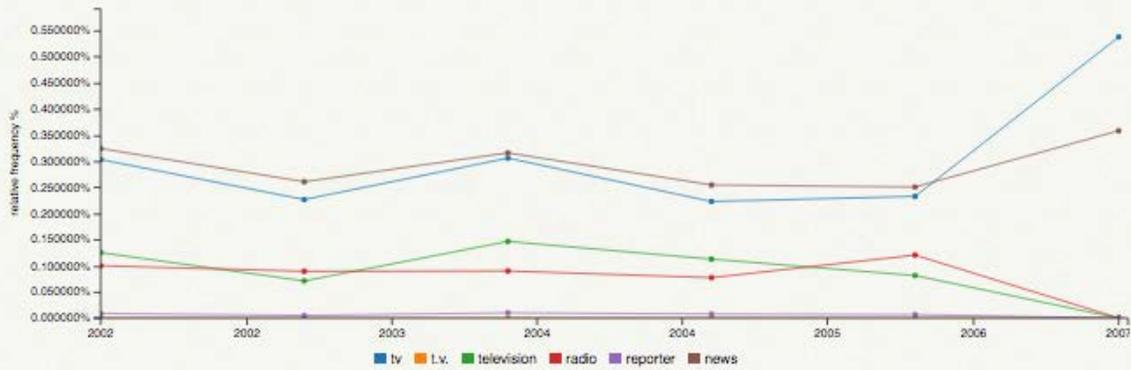
Passing a word or phrase for a sequence of time (within the span defined in the corpus) provides users with an opportunity to explore how language or ideas may have changed over time by comparing the frequency of their usage. These results are contingent on whether unigrams, bigrams, and trigrams have been generated for the corpus.

Utilizing the information gleaned from the Text Analysis process, I produced a unigram graph that explored the frequency with which contributions referenced media technologies in their experiences. Were there changes in the ways that contributors referred to media technologies in their descriptions of 9/11 experience?

## Search corpus "September 11: Bearing Witness to History: Stories Submitted Online"

Graph these comma-separated phrases:  between  and

### Sequence Graph



### Ngram Counts

Ngram	n	Count	Frequency %
news	1	3,867	0.303631%
tv	1	3,574	0.280625%
television	1	1,449	0.113773%
radio	1	1,232	0.096735%
reporter	1	101	0.007930%
t.v.	n/a	0	n/a

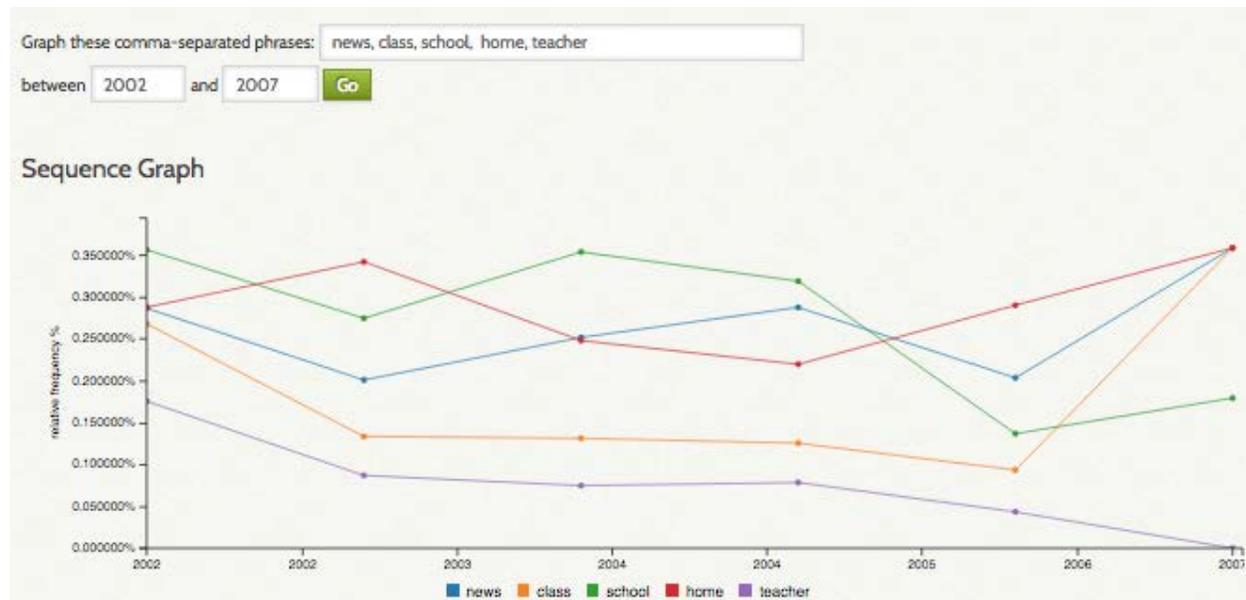
### Total Ngram Counts

n	Total Count	Total Unique Count
1	1,273,585	53,445
2	1,266,489	371,171
3	1,259,408	773,795

Passing the terms, “tv,” “t.v.,” “television,” “radio,” “reporter,” and “news,” produced a graph that confirmed that references to media technologies occurred consistently across the corpus, and with relative frequency. The results also suggested to me that while language usage across the period varied, “tv” was used more consistently over the corpus. References to television and radio were less frequent. Users across the corpus did not utilize the phrase “reporters” frequently across the corpus.

Using this insight, I generated a new corpus. Much like the initial effort, I configured the Ngram plugin to run on “NMAH Story:Story,” the appropriate text element for the *September 11: Bearing Witness to History: Stories Submitted Online* collection. Then, I refined the item pool by defining a search query (“tv”), sequence element (date entered), sequence type (date by year), and sequence range (2001-2007). After completing that process, my corpus of 7,420 items was reduced to 2,498 items that specifically discussed the experience of watching tv.

I validated my item pool and generated unigrams, bigrams, and trigrams. The resulting Ngram Search provided an opportunity to look across the language of contributors who discussed a mediated experience of 9/11.



### Ngram Counts

Ngram	n	Count	Frequency %
school	1	1,958	0.324980%
home	1	1,756	0.291453%
news	1	1,572	0.260914%
class	1	1,259	0.208963%
teacher	1	810	0.134440%

### Total Ngram Counts

n	Total Count	Total Unique Count
1	602,498	35,224
2	600,008	218,694
3	597,523	414,006

### Ngram Frequencies

The Ngram Frequencies are another way to engage with the content of the corpus. As with the Ngram Search, the results are contingent on whether unigrams, bigrams, and trigrams have been generated for the corpus. While the Ngram Search feature allows users to pass particular phrases (ngrams) and graph their results, the Ngram Frequencies identifies the total number and unique number of unigrams, bigrams, or trigrams, and provides this information in order of total count and frequency.

Corpus Viewer Q ...

◀ Back to corpus

### Ngrams in corpus "September 11: Bearing Witness to History: Stories Submitted Online"

Display this many   Unigrams  Bigrams  Trigrams

This corpus has 1,266,489 total bigrams with 371,171 unique bigrams.

Ngram	Total Count	Frequency %
i was	10,506	0.829537%
of the	7,463	0.589267%
it was	6,279	0.495780%
on the	6,127	0.483778%
in the	5,274	0.416427%
to the	3,991	0.315123%
and i	3,959	0.312596%
when i	3,441	0.271696%
the world	2,878	0.227242%
the news	2,805	0.221478%
at the	2,751	0.217215%
was in	2,708	0.213819%
i had	2,577	0.203476%

On the Ngram Frequencies page users can choose whether they would like to view the frequency results for unigrams, bigrams or trigrams, as well as the number of results they would like to display. The results page returns information about the corpus depending on the information passed. For instance, in order to get an overview of the corpus, I chose to display 100 bigrams. The return included, the number of total bigrams (1,266,489) in the corpus as well as the number of unique bigrams in the corpus (371,171).

Users should note that stopwords are not stripped out of the corpus. In this case, this meant that much of the results included the most common words, like a, the, to, and, of.

## Ngrams in corpus "September 11: Bearing Witness to History: Stories Submitted Online"

Display this many   Unigrams  Bigrams  Trigrams

This corpus has 1,273,585 total unigrams with 53,445 unique unigrams.

Ngram	Total Count	Frequency %
the	74,817	5.874520%
i	55,827	4.383453%
and	39,887	3.131868%
to	34,958	2.744850%
was	34,504	2.709203%
of	23,650	1.856963%
in	23,171	1.819352%
a	22,597	1.774283%
that	21,791	1.710997%
my	21,767	1.709112%

The results of the Ngram Frequencies feature are useful for looking across the corpus text to answer questions. For instance, comparing the use of pronouns, these results suggest that contributors were much more likely to use first-person in their descriptions of their experiences of 9/11.

i	55,827	4.383453%
we	13,740	1.078844%

Users were nearly four times more likely to use "I" than "we." This finding is likely due to the collection method, which instructed contributors to reflect on their personal experiences. But this conveys information about the texts that would otherwise be gleaned by close reading. It confirms that users are describing their own experiences in the corpus.

Drawing on the result of the Text Analysis described above, language used to describe the context of experiences can also be located in the results of the Ngram Frequencies. Encouraged by the Text Analysis results described earlier, I also examined the frequency of language referring to media sources and education. As expected, these subjects were referred to frequently across the corpus.

school	5,002	0.392750%
news	3,867	0.303631%

Looking at trigrams, which offer context to the use of phrases to describe the experiences of 9/11, educational contexts are not among the trigrams with the highest frequencies.

Ngrams in corpus "September 11: Bearing Witness to History: Stories Submitted Online"

Display this many   Unigrams  Bigrams  Trigrams

This corpus has 1,259,408 total trigrams with 773,795 unique trigrams.

Ngram	Total Count	Frequency %
the world trade	2,194	0.174209%
i was in	2,065	0.163966%
world trade center	1,869	0.148403%
the twin towers	1,246	0.098935%
one of the	1,128	0.089566%
turned on the	1,117	0.088692%
it was a	1,094	0.086866%
was going on	912	0.072415%
what was going	906	0.071939%
on the tv	893	0.070906%

Media technologies, like news, tv, and radio, however were among the trigrams with the highest frequencies in the results.

The results of the Ngram plugin offered avenues for further exploration in the archive. The Ngram Viewer provides an overview of how language was used across the corpus, offering insights that help to streamline search process and locate items for close reading. Similarly, ngram frequencies offered new ways to explore the entire corpus answering questions about the types of descriptions submitted and revealing insights about contributors themselves.

## Conclusion

As Paul Arthur writes in his article ["Trauma Online: Public Exposure of Personal Grief and Suffering"](#), "These personal narratives serve to demonstrate that these two ways of remembering traumatic events - the collective and the individual - have traditionally been poles apart and often contradictory... In this case, one is a national narrative with the primary aim of mythologizing and nation building by generating and consolidating a unified public vision of events and people, whereas the other is the fragmentary, individualized, and messy form of autobiography." Moving between close and distant readings of the submissions of the *September 11 Digital Archive* is about moving between the individual and the collective experience of 9/11.

The Text Analysis and Ngram plugins made it possible to look across a collection of thousands of submissions and to gain insight on experience of 9/11. The results of this analysis offered a distant reading of the mediated experience of contributors, raising questions about context and

meaning, and allowing for a close reading of contributions like this one; “I witnessed it on tv. I was in school and more specifically [sic] in art class. After arriving [sic] in the room my teacher came in and turned on the television. As everyone sat in their seats the tv flickered, and everyone's [sic] eyes were firmly focused on the screen. The same image was repeating constantly. It showed the second plane fly right into the second tower with smoke still hovering in the air from the first attack.”

The results of these analyses guided a close reading of submissions, identifying a number that described watching the events of 9/11 unfold on tvs or in a classroom. They emphasized learning about 9/11 from newscasters at home, in the car, or at school. Overall findings suggest that many contributors did, as Paul Arthur suggested at the start of this study, experience the events of 9/11 through “their television screens.” It also suggests, as [Mary L. Dudziak wrote](#), “Students and teachers gathered around televisions, sharing the sense that ‘history’ was being made before their eyes... Across the nation and the world, people stopped in front of television screens.” (5)