

Extreme Networking: Finding Connections in the Hall Hoag Collection

Project Background

Before getting into the visualization process I think it is important to shed some light on the background of the collection.

The Gordon Hall and Grace Hoag Collection of Dissenting and Extremist Printed Propaganda contains printed organizational literature (largely pamphlets and leaflets), with smaller numbers of photos, audio-visual items, manuscripts, and monographs published by fringe and extreme groups from the right and the left. The Hall Hoag Collection spans the political spectrum and constitutes the country's largest research collection of right and left wing U.S. extremist groups in the 20th century.

The collection was purchased by Brown University in the Mid-1980s and Gordon Hall began organizing the material and shipping it to Brown. Prior to his death in 2001 Mr. Hall organized 168,000 items emanating from over 5,000 organizations, which constitutes the Hall-Hoag Collection Part I. The material that remained constitutes Hall-Hoag Part II. Part II contains 700,000 items (~1800 linear feet) from ~35,000 different organizations.

In 2013, Brown received a 3-year CLIR Grant (Council on Library and Information Resources) to process the *Hall-Hoag Collection Part II*. Hall Hoag Part II is one of the largest collections at Brown University and prior to the start of this grant it was mostly unprocessed. Hall Hoag Part I is one of the most widely used collections at Brown and processing Part II will further increase the interest in the Hall Hoag collection and will eliminate a good deal of the back log of unprocessed material at Brown.

This grant project will be completed over three phases:

Phase 1. Processing (Years 1-2)

- Hire and train staff (Project Manager and student workers)
- Complete an inventory of materials in Hall-Hoag Part II

Phase 2. Collation and classification (Years 1-3)

- Collate all materials produced by the same organization
- Arrange topical headings in a meaningful hierarchical structure

Phase 3. Product and sustainability (Years 2-3)

- Create a collection-level Encoded Archival Description (EAD) finding aid and
- Organizational-level Encoded Archival Context (EAC-CPF) records

- Create a search interface built on the faceted elements contained in the EAC record
- Collect feedback and refine the interface and search functionality

The funding from CLIR gives Brown the resources and opportunity to spend more time organizing and describing these materials than would normally be available and gives Brown the opportunity to explore a new way of processing and providing access to an archival collection through creating EAC-CPF records. The data being collected provides the perfect platform for visualizing the information.

The Collection

Gordon Hall saw extremism, both left wing and right wing as a threat to American democracy and good government, and he made it his life's work to expose them. Hall saw himself as a member of what he called the "mainstream middle" defined as those who have faith in the American system, accept gradual change, work to bridge the gap between rich and poor and seek unity of all Americans. Hall began to infiltrate and investigate these groups and actively collected their printed propaganda. Hall researched these groups by collecting their materials and attending their meetings and gave lectures on the threat of these radical groups. Gordon Hall focused his efforts on extreme left and extreme right organizations. He categorizes them as follows:



- **Extreme Left:** The radical left in the United States views reform as a vehicle for the perpetuation of a despised system. In vivid contrast to the liberal emphasis on peaceful reform and orderly change, the

radical left seeks total change – disruptive, disorderly and revolutionary.

- **Extreme Right:** The radical right has scrapped parliamentary procedures and democratic processes without which any organization departs the conservative tradition in favor of authoritarianism, totalitarianism and dictatorship. The organizations on the radical right are led almost without exception by a lone individual and there is no machinery for change or the registering of disapproval.¹

Although the original focus was on extremist groups, Hall also collected materials from "dissenting"

TYPES OF ORGANIZATIONS

- Anti-Abortion organizations
- Anti-Integrationist organizations
- Anti-Semitic and Racist political parties
- Christian Identity organizations
- Communist organizations
- Communist political parties
- Cults
- Ku Klux Klan organizations
- LaRouche organizations
- Marxist-Leninist organizations
- Militant Anti-Communist organizations
- Militant Populist organizations
- Neo-Nazi organizations
- Pacifist organizations
- Racial and Ethnic Consciousness organizations
- Right-Wing Christian religious organizations

¹ Hall, Gordon *Left Right Center: Volume 1, Number 1 (January 1968)*

groups or groups that were not yet extremist, but might at some point turn in that direction.

By the late 1960's he had recruited a circle of like-minded volunteers to help in his collection efforts. Grace Hoag, an alumna of Smith College and mother of five, was among those helpers and collaborators. After attending one of Hall's lectures in the late 1960s, Hoag signed up to volunteer. She and others attended meetings and rallies of various groups and brought back their literature. Some volunteers took photographs for Hall to add to his research files. Hoag also assisted Hall in sorting and organizing the materials retrieved by the helpers from the meetings they attended.²

An additional note should also be given on the items in the collection. When Gordon Hall passed away in 2001 all of his belongings were shipped to Brown. This means that there are documents that are neither dissenting nor extremist currently in the collection. These are being continuously weeded out of the collection. They include Time Magazine, The Boston Globe and other mainstream publications.



Processing The Material

The Hall Hoag Part II was shipped to Brown in 800 large cartons that each contained material from 250 to 300 organizations without folders and in no order. Starting in 2006, over the summer student workers removed the contents of the boxes, sorted the material by organization, type of material, or title of the serial or book. Materials were grouped by organization and placed in acid-free folders. Once the folders were placed into acid free boxes, the folder titles (name of organization, type of material, or title of serial or book) were inventoried Excel.

By March 2013 all 800 shipping cartons were emptied and the result was 1600 record center boxes of foldered material and an Excel inventory containing over ~160,000 lines. However, there was no arrangement to the collection and materials from the same organizations were scattered throughout the

boxes. For example, there are 400 folders of material from the American Nazi Party, in 260 different boxes. If a researcher wanted all of the material on the American Nazi

² <http://blogs.brown.edu/hallhoag/gordon-hall-and-grace-hoag/>

Party, 260 boxes would have to be pulled. We are in the process of collating or bringing together of like-organizations. We are in the midst of this process now should be able to complete this work over two summers with 8 students working 35 hours a week. By the end of next summer the collection should be organized in alphabetically. The collection is so large that it is hard to make an impact on the physical arrangement without a large number of people working on it. Therefore we have been limited to working on this during the summer, when students are available for full time work.

Data Collection Policy

Using the collection inventory work on establishing intellectual control over the collection has begun. This work includes two parts:

1. EAD Finding Aid for the entire collection
2. EAC-CPF records for each organization in the collection

The EAD finding aid cannot be completed until the collection is organized and the inventory is updated with the proper locations, but the work to create the EAC-CPF records has begun.

Encoded Archival Context – Corporate bodies, Persons, and Families (EAC-CPF) is an XML schema that primarily addresses the description of individuals, families and corporate bodies that are associated with archival records that supports the linking of information about one agent to other agents to show/discover the relationships amongst record-creating entities, and the linking to descriptions of records and other contextual entities. It is a standard language that can be used to establish the context of archival record creators.³ We are creating EAC-CPF records to find connections between the organizations in this collection based on names, dates, geographic locations, members, publications, subjects, related archival collections.

As mandated by the grant we must create an EAC-CPF record for each of the ~35,000 organizations that have material in the collection. To save the time in creating an XML file for each of the organizations I have created a FileMaker Pro database, which can store information on each organization and has been customized to export this data in valid EAC-CPF XML. More importantly for the work here the creation of a customized FileMaker Pro database allowed me to store data that can be easily exported in a variety of formats including Excel, TSV, CSV and XML.

The type of information gathered was determined by two factors.

1. What would help create robust EAC-CPF records
2. What could be easily gathered considering the scope of the collection

This led to the decision to try to collect the following data at a maximum:

1. An authoritative version of the organization name

³ <http://eac.staatsbibliothek-berlin.de/about/ts-eac-cpf.html>

- a. It was important to start to standardize the names of organizations to help weed out the duplicate entries for the same group. Where possible I used the Virtual International Authority File and used their version of organization names. If it could not be found there I manually changed organization names to remove "The" from the beginning and any notes that may have been included in the inventory. Additionally if we want to eventually connect this data with other archival collections the authoritative information will be crucial.
2. Broad subject categories: I looked through the types of organizations that had been researched and I chose categories based on that. They include (I provide some definitions for groups that might not be apparent:
 - a. Animal Rights/Environment
 - b. Corporations: This mostly contains organizations that are include but not extreme. GM would be an example
 - c. Educational Groups: Many of these groups focus on home schooling, and community education. This group also contains university publications.
 - d. Government: National and local government groups and other groups related to the US government.
 - e. Hate Groups: Groups that are specifically racist or against certain groups of people.
 - f. Human Rights/International Focus: This was a catch all for groups with an international focus typically on a 3rd world country.
 - g. Left Wing Political Groups
 - h. Media/Publishers/Bookstores
 - i. Medicine/Health
 - j. Minority / Women / LGBT
 - k. Religion
 - l. Right Wing Political Groups
3. Narrow subject categories: Many groups needed further analysis and eventually will need more than one category. Currently they only have one narrow category to help define their scope. These include:

<ul style="list-style-type: none"> • African American • Anarchism • Animals Rights • Anti-Communist • Anti-Government • Atheism • Christianity • Communism • Community Organizations • Conservatism • Conservative 	<ul style="list-style-type: none"> • Corporations • Counter Culture • Cults • Economics • Education • Environmentalism • Feminism • Free Enterprise • Gun Rights • Health • Homophobia • Human Rights
---	---

- Immigration
- International Scope
- Judaism
- LGBT
- LaRouche
- Labor Unions
- Left Wing
- Libertarianism
- Local
- Local Government
- Media
- Militias
- Nationalism
- Native American
- Newspapers and Periodicals
- Objectivists
- Pacifism. & Nonviolence.
- Prisoners Rights
- Pro Choice
- Pro Life
- Progressive
- Protest Groups
- Publishers
- Quakerism
- Racism
- Right Wing

4. Organizational histories: Currently this field only contains a link to a biography page for organizations. Typically this is a Wikipedia link.
5. Start dates and end dates: This data was collected in the format of YYYY-MM-DD. I then used a FileMaker Pro database to calculate these fields by decade. It seemed much easier to search for and group organizations by data. The exact data or year was not lost, but it seemed useful to have both. The decade ranges are useful for filtering, but the exact date will be important for researchers.
6. Geographic locations: Here I collected city, state and country. There were far too many organizations to try to find exact locations.
7. Prominent members: Members are crucial for finding connections between organizations and creating a network. Here I collected the following information:
 - a. Name
 - b. Authoritative version of name and authority ID (VIAF)
 - c. Years involved
 - d. Position/Role
8. Related archival collections: This is another crucial field for networking the Hall Hoag collection to the broader archival community. Hyperlinks were gathered to offer related resources on the organizations in this collection.
 - a. Institution Name
 - b. Collection Name
 - c. Collection ID
 - d. URL to collection finding aid.

This is then combined with the information about the groups in connection with the collection

1. Where the items are located
2. How many total items are contained in the collection

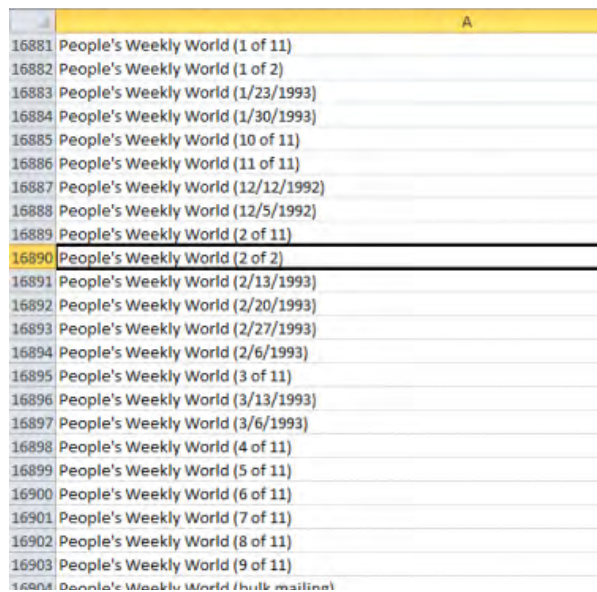
The amount of data and the type of data makes this project very different than other archival projects. Normally historical information would be gathered for the creator of a set of archival records. In this case we are gathering historical information on each organization within the collection and have used some techniques that overlap with the digital humanities to research, gather data and curate the data.

This also limits the scope of the data. We are collecting data about groups inside of a specific collection, not just on extremism as a whole. At the same time, the purpose of collecting the data is to provide further access to the material, not just to visualize the material to necessarily tell a story about extremism. As we have started going through the process of cleaning and gathering data, my hope is that I will eventually have data that can tell a story on its own even though its basis is an archival collection. But it seems important to note that there are inherent limitations to the data and when gathered the data is really telling us something about the Hall Hoag collection first, and maybe something broader second.

Cleaning/Storing The Data

The first step in this process was determining what organizations were in the collection. To create the organizations list, we started with the inventory that contained ~160,000 lines in an Excel spread sheet and then had to weed out the duplicates. Rather than going through each line in Excel and deleting the duplicates, I used a few tricks.

Starting with Excel, the “delete all duplicates” and “text to columns” functions were vital. The “delete all duplicates” function cut the number from 160,000 down to about 75,000. However, this only deleted the lines that were exactly the same. Many duplicates still existed due to misspellings or variations on the organization name. Excel could not recognize them as the same organization and therefore could not automatically delete them.

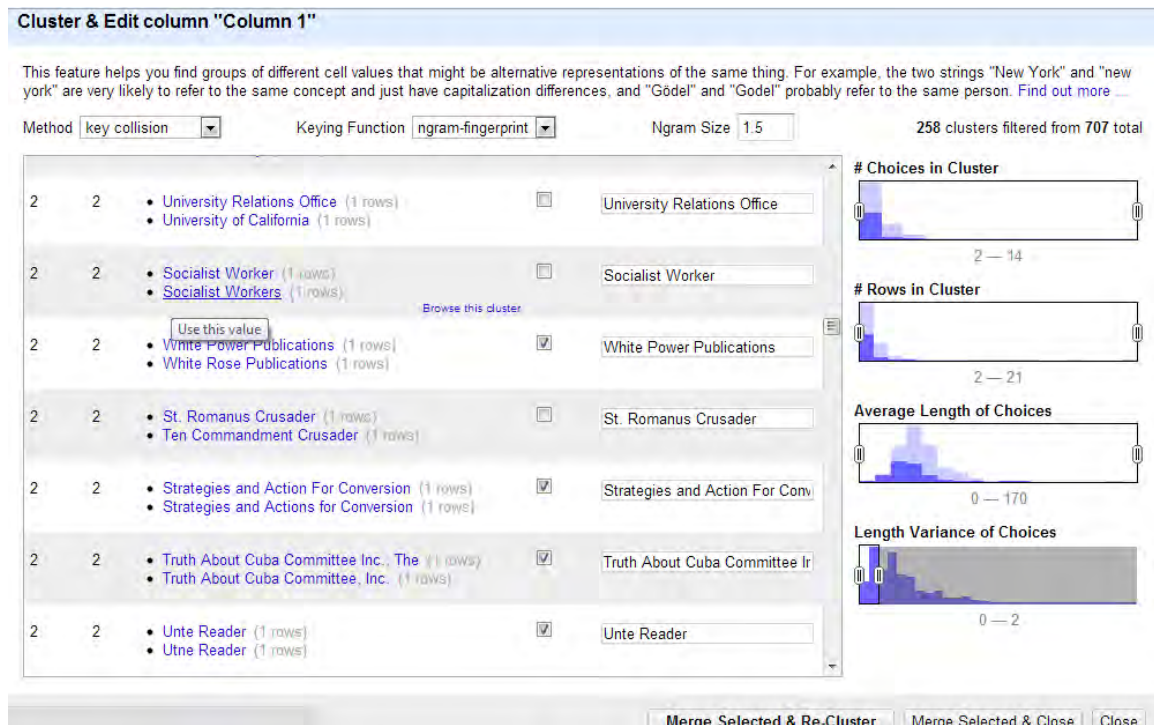


	A
16881	People's Weekly World (1 of 11)
16882	People's Weekly World (1 of 2)
16883	People's Weekly World (1/23/1993)
16884	People's Weekly World (1/30/1993)
16885	People's Weekly World (10 of 11)
16886	People's Weekly World (11 of 11)
16887	People's Weekly World (12/12/1992)
16888	People's Weekly World (12/5/1992)
16889	People's Weekly World (2 of 11)
16890	People's Weekly World (2 of 2)
16891	People's Weekly World (2/13/1993)
16892	People's Weekly World (2/20/1993)
16893	People's Weekly World (2/27/1993)
16894	People's Weekly World (2/6/1993)
16895	People's Weekly World (3 of 11)
16896	People's Weekly World (3/13/1993)
16897	People's Weekly World (3/6/1993)
16898	People's Weekly World (4 of 11)
16899	People's Weekly World (5 of 11)
16900	People's Weekly World (6 of 11)
16901	People's Weekly World (7 of 11)
16902	People's Weekly World (8 of 11)
16903	People's Weekly World (9 of 11)
16904	People's Weekly World (bulk mailing)

This image shows the different ways that “People’s Weekly World” was entered. Using the “Text to Column” function and delimitating the text by the “(“ symbol the data was separated by the “(“

symbol. I then ran the delete all duplicates function again and was left with 50,000 organization names.

To weed out the misspellings and more complicated duplicates I used [Open Refine](http://openrefine.org/)⁴, which is a tool that will find similar pieces of data and clean it. I uploaded the list of 50,000 into Open Refine and it found the items that were similar and allowed me to change them all to one common name.



The last step was looking at each remaining line of data manually and deleting the found duplicates. In all, the process took about 60 hours to complete and we ended up with ~35,000 unique organization names.

The list of 35,000 names was imported into FileMaker Pro. I chose to build a FileMaker Pro database to tracking data about each organization because Brown already had licenses for the software and I already knew how to create a database in FileMaker from previous experience. Also, Brown has not worked with EAC-CPF before so there was no existing procedure to create these records and I was left to my own devices.

⁴ <http://openrefine.org/>

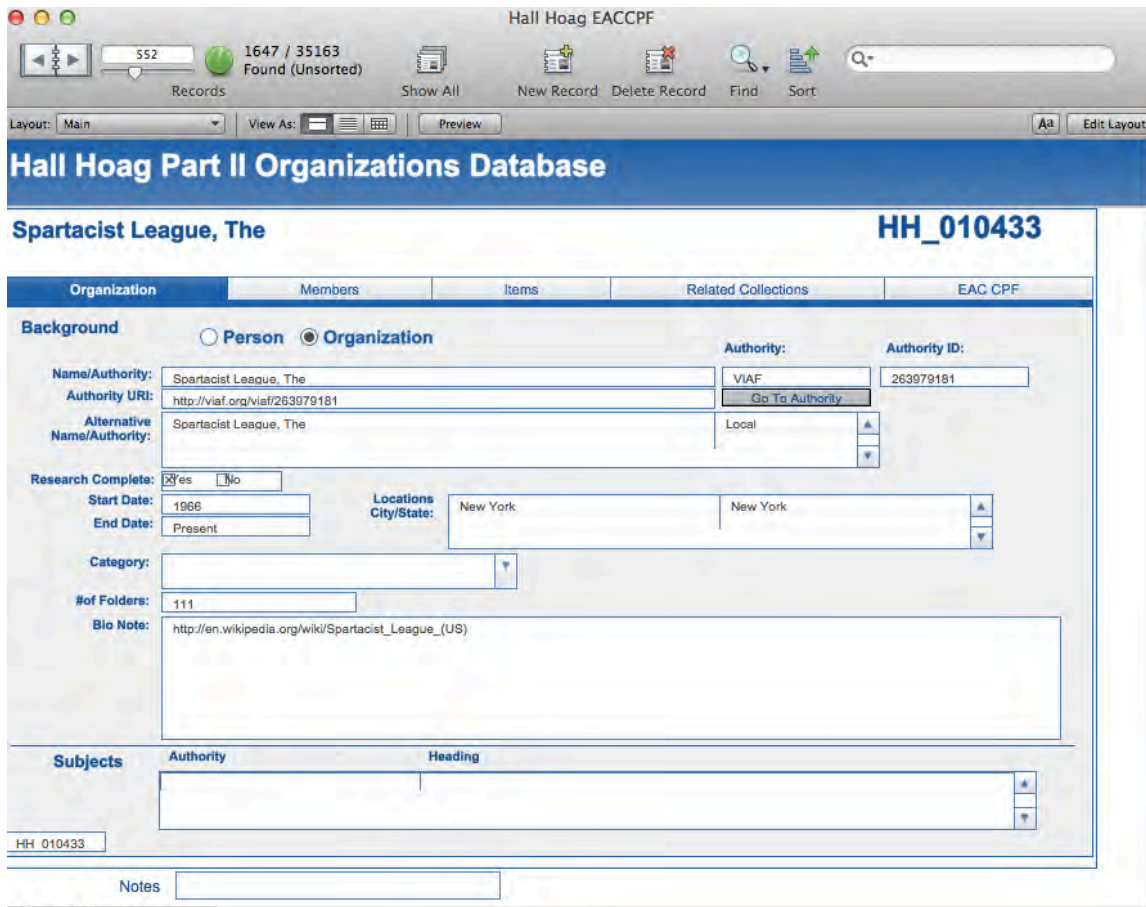


Figure 1: FileMaker Pro Screen Shot

That is not to say there are no other advantages to using FileMaker to create the EAC-CPF records. EAC-CPF is a specialized XML schema that requires training and experience to write. I will be the only person with any EAC-CPF experience working on the grant. It simply is not feasible for me to write the code for 35,000 individual records. It takes very minimal training to enter data into FileMaker, and even more importantly it is very easy to import data into FileMaker from various data sources including, Excel, csv, tsv and XML.

To collect information on 35,000 organizations we have tried to automate as much work as possible. A search and retrieval script was run through Open Refine on the Virtual International Authority File for each organization.

VIAF is an online database that combines the name authority files of participating institutions (Library of Congress, the Deutsche Nationalbibliothek, and the Bibliothèque nationale de France and more) into a single name authority service.⁵ For our purposes we thought VIAF would be the authority file to use because it has such a large base of authority files. Many of the organizations in the collection are virtually unknown and starting from a large database seemed like the way to go.

This script ran a search for every organization name in the collection returning an XML result that contained an authority file for each. We had ~12,000 positive hits and we are currently working through a way to pull data out of the XML results including a standardized form of their names, ids/urls that will link the names to other authority files, and other items in the MARC record notes, such as names, dates, publications, subject headings and locations.

```
<count="1"><ns2:text>Annual report : promoting the protection of all
animals</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Animals in a research library,
1961.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Animals in a research
laboratory.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Animal
sheltering.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Animal sheltering
magazine</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Animal activist
alerts</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Anim. act.
alert</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>AllAnimals</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:da
ta count="1"><ns2:text>All
animals</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>52 simple things you can do to help end pet overpopulation : ways you can help our
animal companions- and have fun doing
it!</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data></ns2:titles><ns2:coauthors><ns2:da
ta count="5" tag="951"><ns2:text>Humane Society
International.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="4"
tag="950"><ns2:text>Mold,
Chris.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="4"
tag="950"><ns2:text>Salem, Deborah
J.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="4"
tag="950"><ns2:text>Rowan, Andrew
N.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="3"
tag="950"><ns2:text>Hodge, Guy
R.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="2"
tag="951"><ns2:text>Walter J. Klein
Company.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="2"
tag="950"><ns2:text>Lane, Nancy,
1960-</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="2"
tag="950"><ns2:text>Houk,
Randy.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="1"
tag="951"><ns2:text>Wildlife Network
(Calif.).</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="1"
tag="951"><ns2:text>Royal Society for the Prevention of Cruelty to
Animals.</ns2:text><ns2:sources><ns2:s>SUDOC</ns2:s></ns2:sources></ns2:data></ns2:coauthors><ns2:publis
hers><ns2:data count="3"><ns2:text>Humane Society of the United
States j</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="3"><ns2:text>The
Society's</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="3"><ns2:text>The Benefactor,
Inc.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data count="2"><ns2:text>The
Company.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="2"><ns2:text>Institute for the Study of Animal
Problems</ns2:text><ns2:sources><ns2:s>SUDOC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Time Inc. Custom
Pub.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>National Humane Education
Center.</ns2:text><ns2:sources><ns2:s>LC</ns2:s></ns2:sources></ns2:data><ns2:data
count="1"><ns2:text>Little,
```

Figure 2 VIAF Record Sample

We know that we will not be able to automate everything, so we chose a list of about 1,700 organizations that have the most material in the collection and had student research them using VIAF.org, WorldCat.org and Wikipedia (as well as Google searches). This work was done over a 2 week period with 8 students working 35 hours a day. The students filled in pieces of data in an Excel spread sheet, rather than attempting to write an organizational history. They found, locations, dates, member names, and related collections and this data was then imported into the FileMaker database.

⁵ <http://www.oclc.org/research/activities/viaf.html>

Visualizing and Analyzing The Data

We are at an exciting time in the project, as we are just starting to gather information, but have already seen the value of this data. This data is already telling a story and as more and more information is added new insights can be drawn that would otherwise not be apparent even from looking at the physical materials.

For the visualization for the class project I chose to use Google Fusion Tables to visualize the data because it was a very good fit with the type of information I was collecting.

1. Mapping: There is a geocoding/mapping feature that works well with the location information I was collecting.
2. Card Display: There is a display feature to view the records as “cards” which fit well with the limited amount of faceted data I have. E.g. I do not have a lot of text in any of my data fields so I did not need a lot of space to display the data.
3. Filtering: The data I have collected are really facets or filters. States, Categories, and Dates are all items that can help group records together.
4. Summary Tables: This also works well with faceted data. You can “summarize” individual columns of data. This really means that Fusion Tables will give you a count of how many times a particular piece of data was used in a column. So if I summarize the “category” field I get a count of how many times each category was used. This makes the process of creating pie charts or bar graphs incredibly easy.

I imported the data from the ~1,700 organizations that were researched to use as a sampling of the data that will eventually be more robust. Use the link below to access this data:

<https://www.google.com/fusiontables/DataSource?docid=1pvw6qPkRXbYgDPatHVAnREstZLZfimA-SMZyrYM>

More than anything else, what this work has proven to me is a proof of concept. I can see that there is real value in this data and that more I can collect the better it will become. For our purposes here I want to look at the mapping, filtering and summary tables further.

Mapping

Using the city, state and country data I was able to plot these organization onto a map. Each location is represented by a red dot which when click provides additional information about the organization in that location. This allows someone using the map to click on an item and get a link to a related archival collection or a biography page. It also includes information right there on the screen about that organization. Click the map below to view this map:

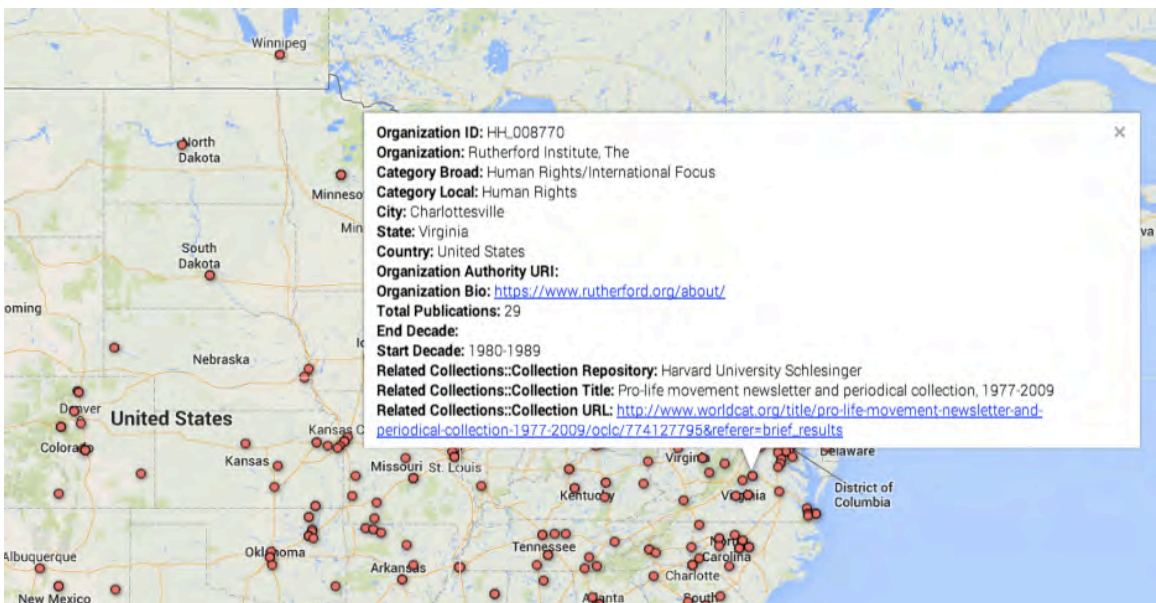
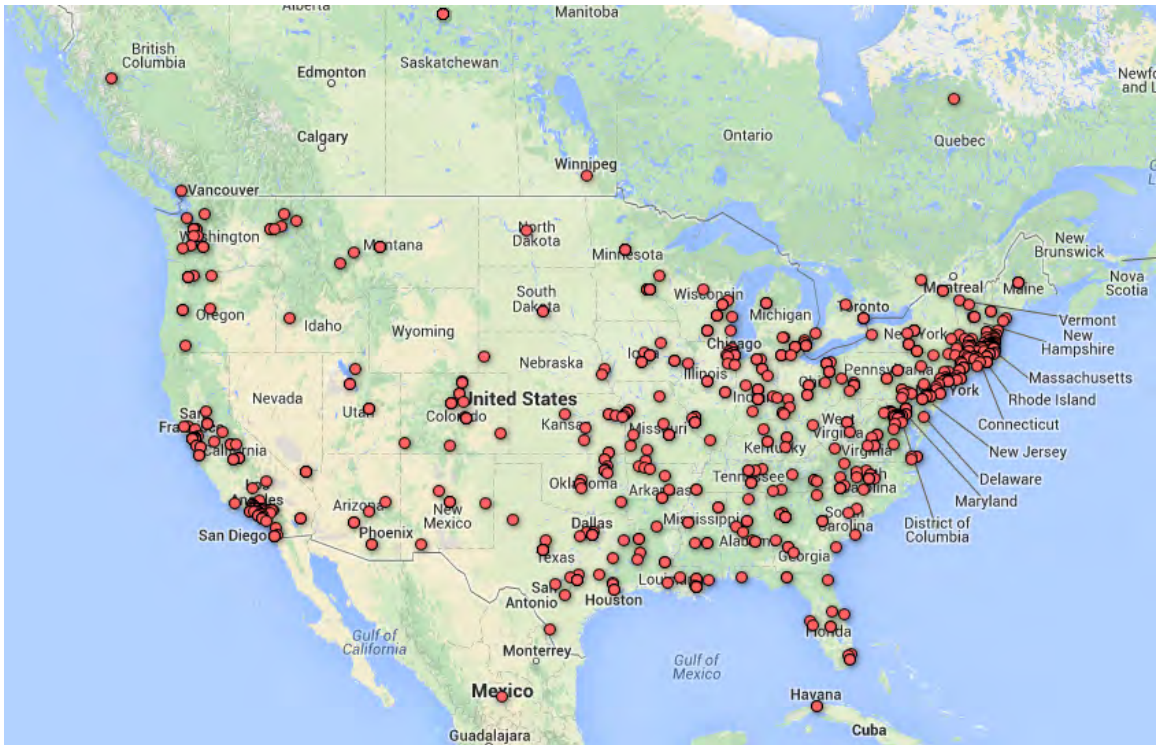


Figure 3 Additional Information Displayed on Map

Using Fusion Tables also limits the visualization in certain ways. Most importantly all organizations in a specific geographic location are mapped on top of one another. For example, this only allows for one organization in Los Angeles, California to be clicked on. As the project moves forward I will have to develop a tool that allows for many organizations in the sample location to be differentiated on a map. Something that perhaps can expand when click to show the scope of all organizations in a specific geographic area.

Filtering

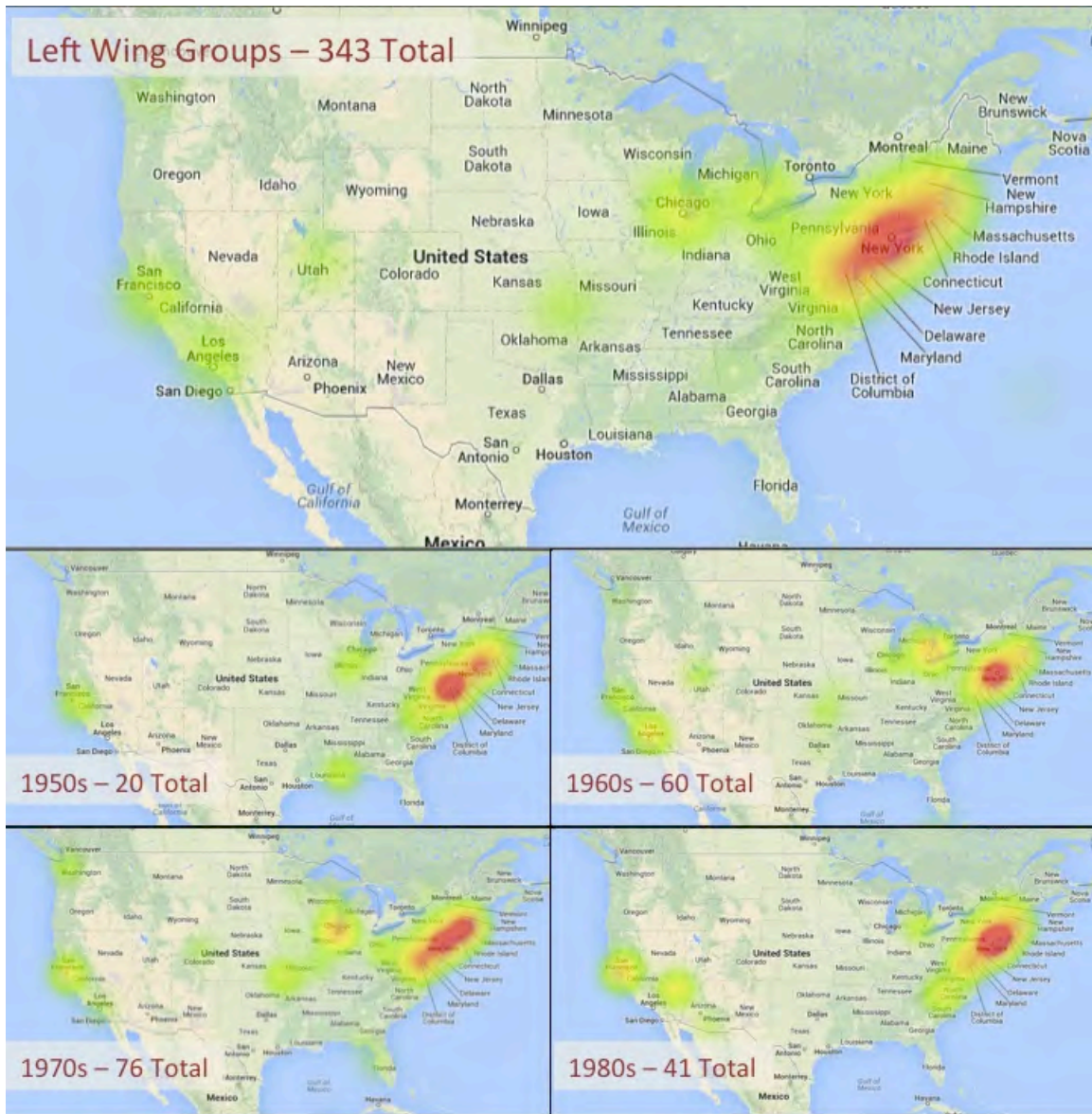
The filtering option can be used in conjunction with the map or the card display. In the image below I have filter the collection to groups in “Alabama” and “Right Wing.” This makes it possible for researchers who are only interested in specific groups based on filters to see what exists in the collection. Potential this would also tell them where these items are located in the Hall Hoag collection and could function as some form of a finding aid.

<p>Organization ID: HH_005690 Organization: Patriotic Education, Inc. Category Broad: Right Wing Political Groups Category Local: Right Wing City: Madison State: Alabama Country: United States Location (Combined): Madison, Alabama, United States Organization Authority ID: Organization Authority URI: Organization Bio: http://www.patriotic-education.org/ Total Publications: 11 End Decade: Start Decade: 1950-1959 Related Collections::Collection Repository: Related Collections::Collection Title: Related Collections::Collection URL:</p>	<p>Organization ID: HH_032012 Organization: Wallace, George C. Category Broad: Right Wing Political Groups Category Local: Right Wing City: Montgomery State: Alabama Country: United States Location (Combined): Montgomery, Alabama, United States Organization Authority ID: 65418751 Organization Authority URI: Organization Bio: http://en.wikipedia.org/wiki/George_Wallace Total Publications: End Decade: 1990-1999 Start Decade: 1910-1919 Related Collections::Collection Repository: Related Collections::Collection Title: Related Collections::Collection URL:</p>
<p>Organization ID: HH_017730 Organization: Ludwig Von Mises Institute, The Category Broad: Right Wing Political Groups Category Local: Libertarianism City: Auburn State: Alabama Country: United States Location (Combined): Auburn, Alabama, United States Organization Authority ID: 158360281 Organization Authority URI: Organization Bio: https://en.wikipedia.org/wiki/Ludwig_von_Mises_Institute Total Publications: 64 End Decade: Start Decade: 1980-1989 Related Collections::Collection Repository: Related Collections::Collection Title: Related Collections::Collection URL:</p>	<p>Organization ID: HH_024406 Organization: Centreville Press Category Broad: Right Wing Political Groups Category Local: Conservatism City: Centreville State: Alabama Country: United States Location (Combined): Centreville, Alabama, United States Organization Authority ID: Organization Authority URI: Organization Bio: Total Publications: 21 End Decade: Start Decade: 1880-1889 Related Collections::Collection Repository: Related Collections::Collection Title: Related Collections::Collection URL:</p>

Figure 4 Card Display

I have provided a sample visualization using a “heat map” and filtering the collection by Left Wing Groups and then further filtering by decade. Red represents the highest concentration and green represents the least. This allows you to see changes in the movement of left wing groups from the 1950s to the 1990s. Because Gordon Hall was based in the Boston there is a focus on the North East, but in different decades we can see other trends pop up. In the 60s and 70s we can see Michigan and Chicago gain more focus, which is probably in conjunction with the Students for a Democratic Society

forming in Michigan and the militant movements developing in Chicago around the 1968 Democratic National Convention and the Black Panther Party.



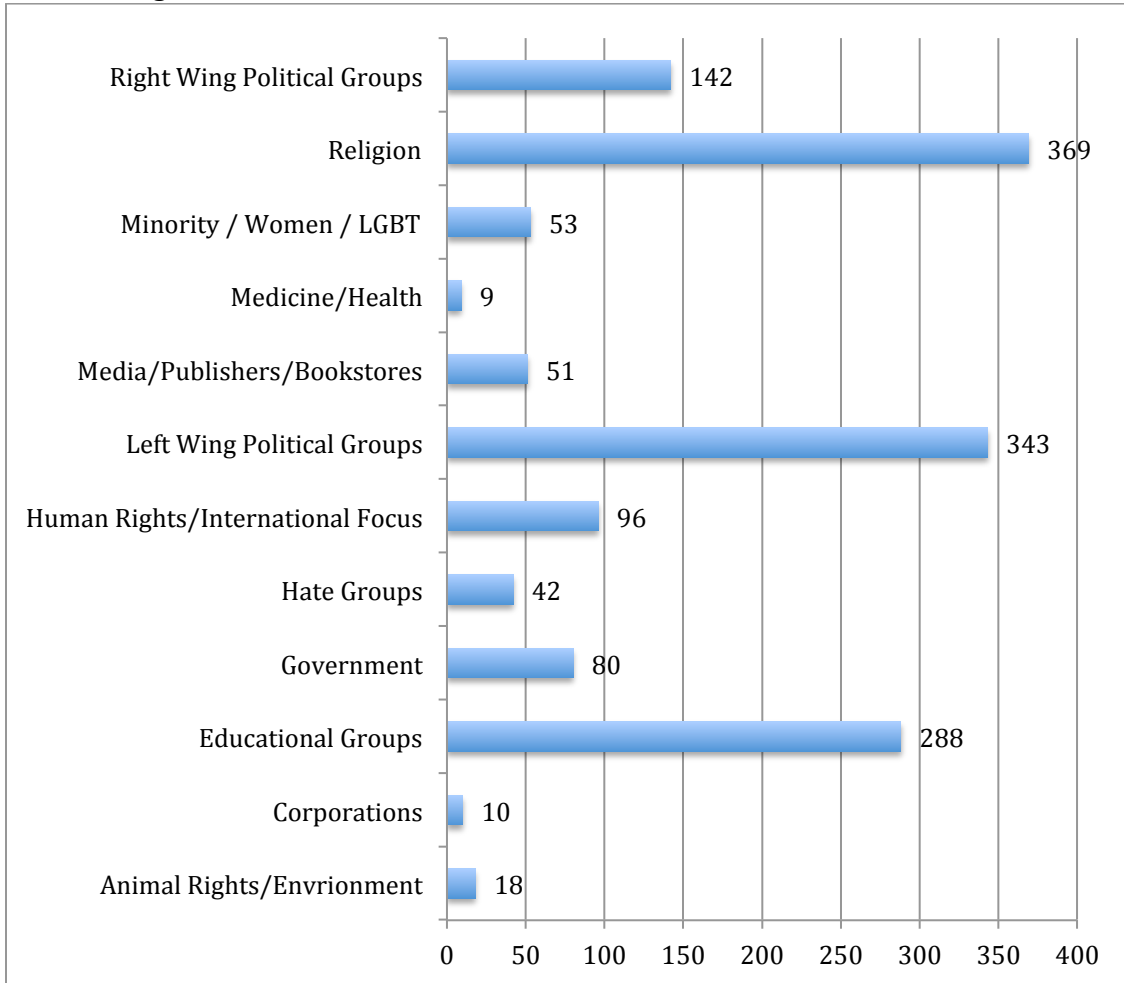
I also encourage you to look at the maps created in the fusion table and play around with the available filters. The most useful ones seem to be categories, locations and dates.

<https://www.google.com/fusiontables/DataSource?docid=1pww6qPkRXbYgDPatHVAAnRstZLZfimA-SMZYrYM>

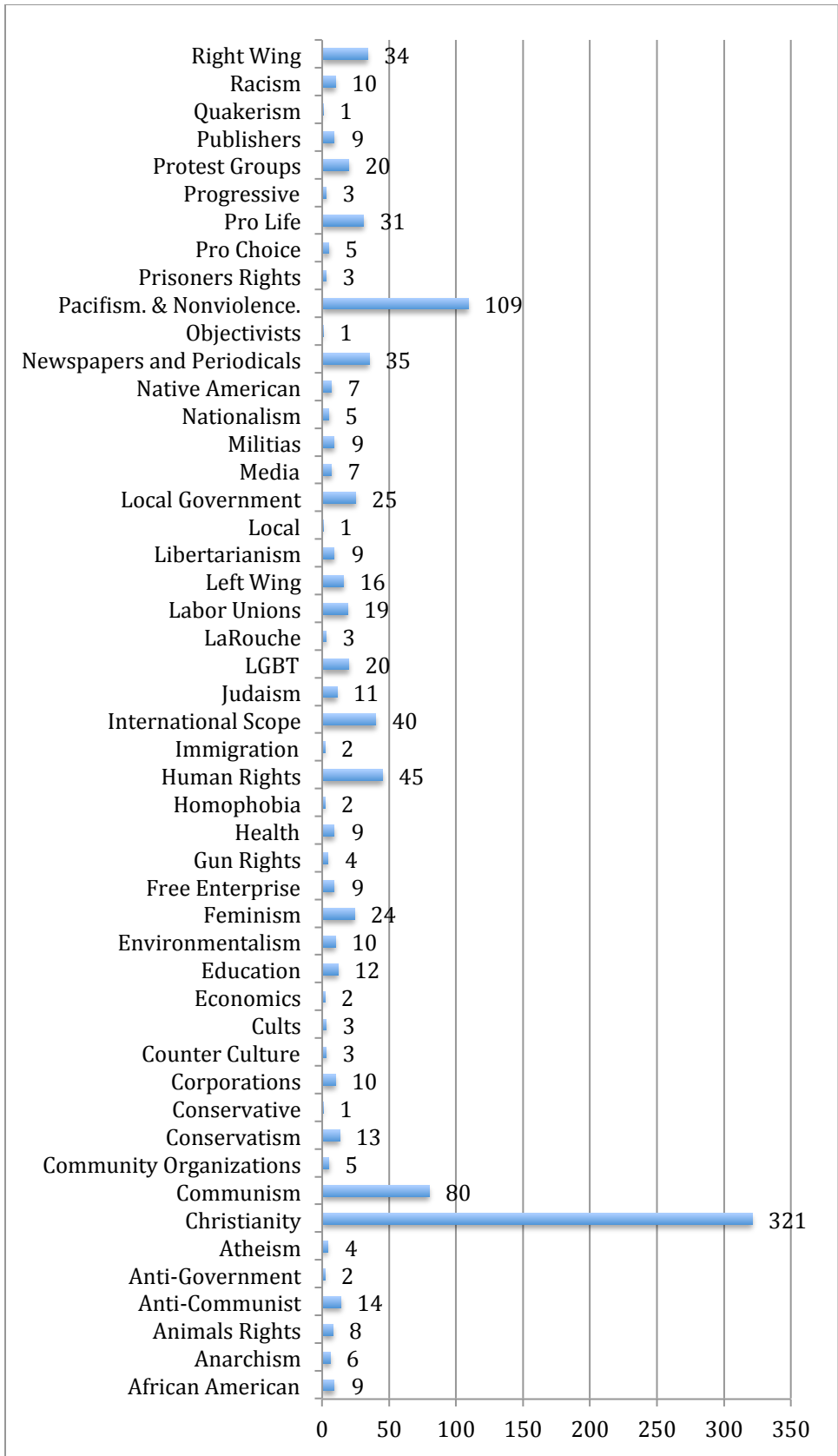
Summary Tables

I was able to summarize a lot of the data very quickly using fusion tables and create bar graphs based on the data. These provide a great overview of the data contained in the collection. I envision the graphs being used to show researchers the scope of the contents in a concise and easy to read way.

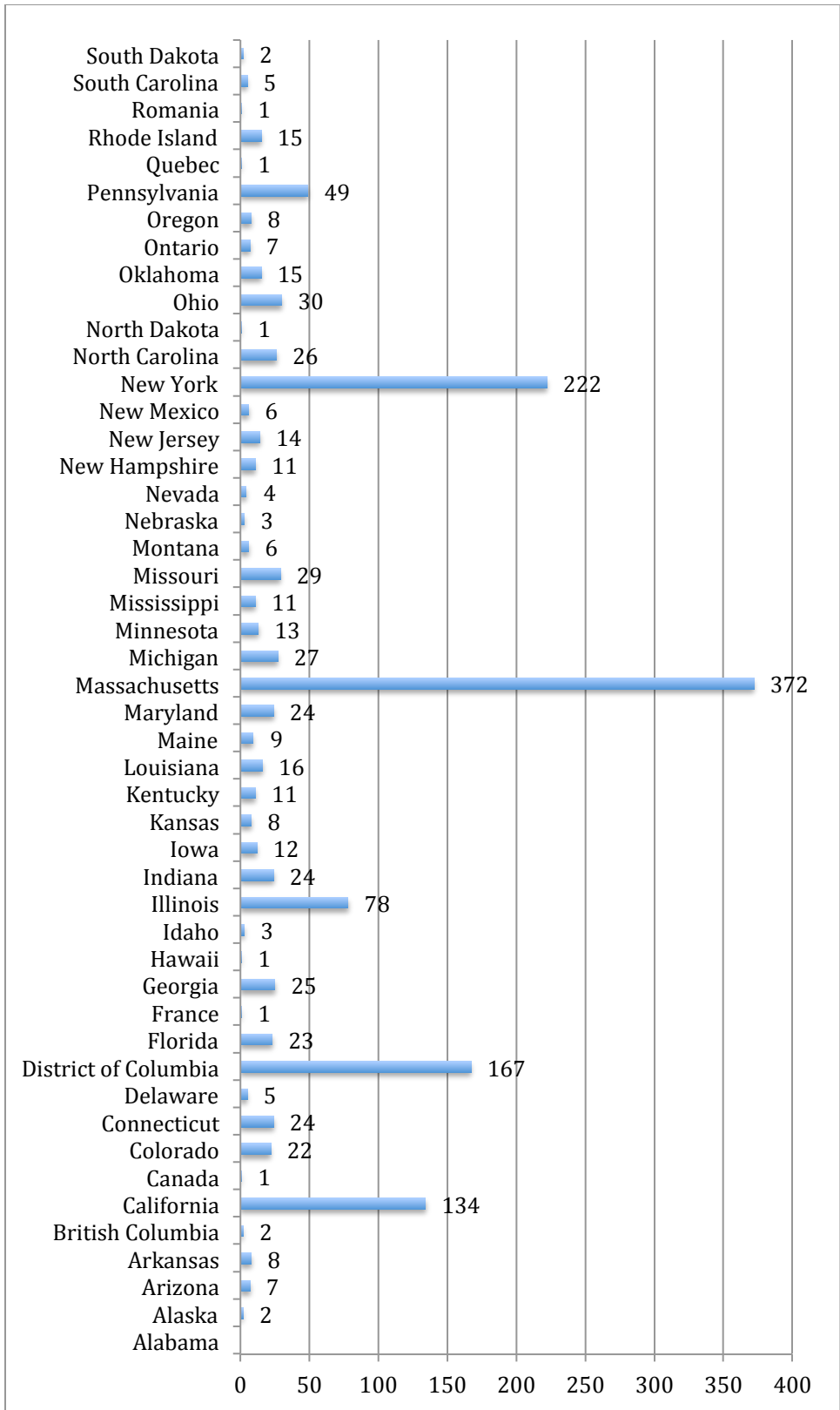
Broad Categories



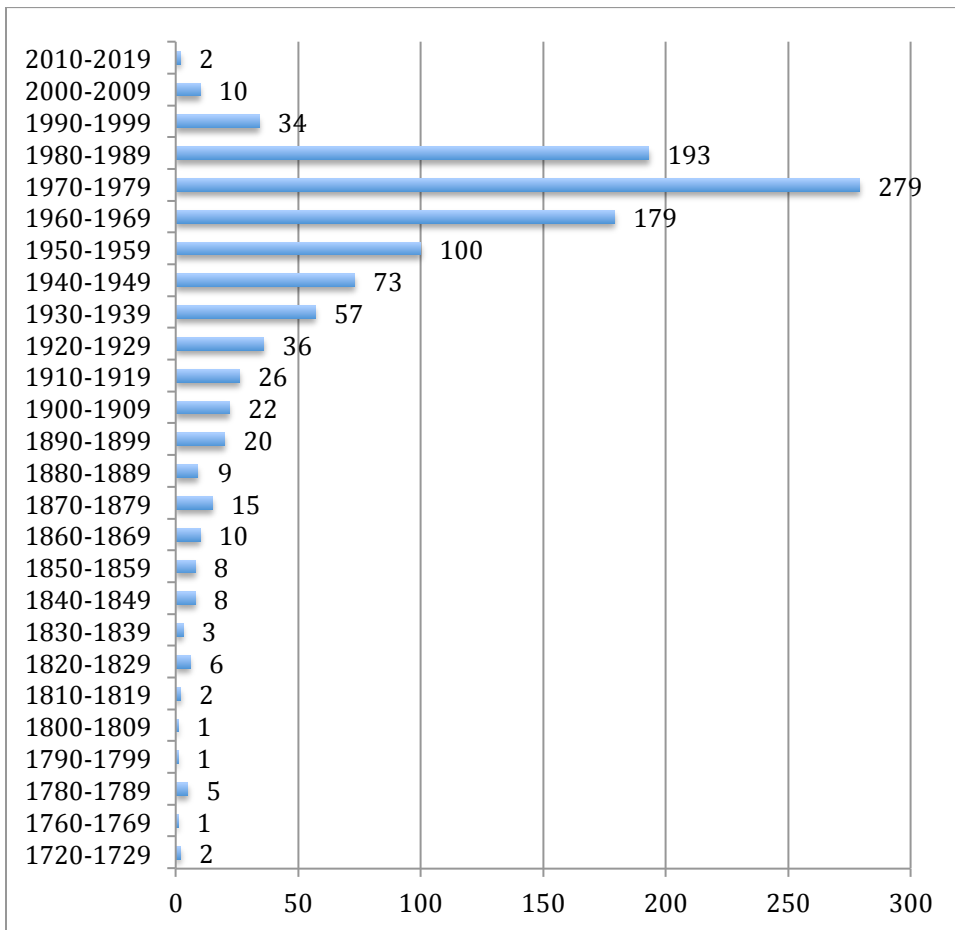
Narrow Categories:



States:



Decade Founded



Moving Forward

It will continue to be a challenge to find information on each organization, but with the funding from the grant we should be able to hire enough students to do at least some minimal research on the organizations we cannot find any information on through automation. Even if we are only able to find a limited amount of information for the organizations, we are hoping there will still be value in the volume of information collected. Simply having locations and start dates for 20,000 extremist groups would be a very valuable resource to researchers and can help facet search queries on the collection.

I think more avenues to automate this data will be possible moving forward as well. It is possible to use something called DBpedia, which is a database version of Wikipedia to query organization names and download full Wikipedia entries. If I am able to query the

whole collection and then pull this data into my database it will greatly improve the project.

The end goal of this project is to create an online interface that will allow users to manipulate this data and explore the collection in new ways by integrating the EAC and EAD records. The work I have done on this project has proven to me that this should be possible.

As end-users search the database, results will contain both organizational-level and item-level information, with links to the complete, corresponding EAC record as well as to the full EAD finding aid. The web interface will be a digital humanities resource in and of itself as a great place to find quick information about a myriad of extremist organizations that currently does not exist without having to visit the archives to view the materials. At the same time it would still provide access/location information of a traditional finding aid if a user still wanted to view the materials.

The Social Networks and Archival Context Project

(<http://socialarchive.iath.virginia.edu/>) offers a model for using EAC as a starting point for linking archival data between collections and creators. With SNAC user can start with a name, which takes you to a record that shows you all of the connections that name has to other entities, which are hyperlinked for further exploration. For example, George S. Patton is connected to 59 people. From his record we can see how he is connected to those people and go directly to their records, which will be connected to other people and other archival collections. It almost creates a six degrees of separation type of interface.

Our interface will establish a solid foundation for future efforts to link, share, enhance and visualize the EAC data through traditional methods as well as Web 2.0 collaborations with scholars and external repositories (An example would be geographically mapping out where all of these organizations are located). Since both EAD and EAC are based on XML, they are easily interoperable with other metadata schemes, such as MARCXML, MODS, and Dublin Core. This enables the collection to be linked to other related resources at Brown and beyond. Through the use of Web-based technologies such as XForms, the EAC data will be open to future enrichment by subject experts, scholars using the collection, and other crowd-sourced efforts to enhance this substantial resource.

```
<?xml version="1.0" encoding="UTF-8"?>
<eac-cpf xmlns="urn:isbn:1-931666-33-4"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:isbn:1-931666-33-4
http://eac.staatsbibliothek-berlin.de/schema/cpf.xsd">
  <control>
    <recordId>US-RPB-HH_002504</recordId>
    <maintenanceStatus>new</maintenanceStatus>
    <maintenanceAgency>
      <agencyCode>US-RPB</agencyCode>
      <agencyName>Brown University. Library. Manuscript Collection </agencyName>
    </maintenanceAgency>
  <languageDeclaration>
    <language languageCode="eng">English</language>
    <script scriptCode="Latn">Latin (Roman) </script>
  </languageDeclaration>
  <conventionDeclaration>
    <citation>AACR2</citation>
  </conventionDeclaration>
  <localControl localType="title">
    <term>EAC-CPF Record: United Presbyterian Church in the United States of
America </term>
  </localControl>
  <localControl localType="collection">
    <term>The Gordon Hall and Grace Hoag Collection of Extremist and
Dissenting Printed Propaganda </term>
  </localControl>
  <maintenanceHistory>
    <maintenanceEvent>
      <eventType>created</eventType>
      <eventDateTime standardDateTime="2014-02-05"/>
      <agentType>human</agentType>
      <agent>Johnson, Daniel </agent>
      <eventDescription>Created at Brown University using FileMaker
Pro.</eventDescription>
    </maintenanceEvent>
  </maintenanceHistory>
  <sources>
    <source xlink:actuate="onRequest" xlink:show="new" xlink:type="simple"
xlink:href="http://viaf.org/viaf/130187381" >
      <sourceEntry>Virtual International Authority File ID:
130187381</sourceEntry>
    </source>
    <source xlink:actuate="onRequest" xlink:show="new" xlink:type="simple"
```

```

xlink:href="http://library.brown.edu/collatoz/info.php?id=62">
  <sourceEntry>Hall-Hoag Collection of Dissenting and Extremist Printed
Propaganda</sourceEntry>
</source>
</sources>
</control>
<cpfDescription>
  <identity>
    <entityType>corporateBody</entityType>
    <nameEntry>
      <part>Presbyterian Church in the U.S.A. </part>
      <authorizedForm>VIAF</authorizedForm>
    </nameEntry>
    <nameEntry>
      <part>United Presbyterian Church in the United States of
America</part>
      <authorizedForm>Local</authorizedForm>
    </nameEntry>
  </identity>
  <description>
<existDates>
<dateRange>
  <fromDate standardDate="1958-05-28">5/28/1958</fromDate>
  <toDate standardDate="1983">1983</toDate>
</dateRange>
</existDates>
<places>
  <place>
    <placeRole>Headquarters</placeRole>
    <placeEntry>Pittsburgh, Pennsylvania </placeEntry>
  </place>
</places>

<biogHist><p>http://en.wikipedia.org/wiki/United_Presbyterian_Church_in_the_U
nited_States_of_America</p>
  </biogHist>
</description>
<relations>
  <cpfRelation cpfRelationType="associative" xlink:type="simple"
xlink:role="http://RDVocab.info/uri/schema/FRBEntitiesRDA/Person"
  xlink:href="http://viaf.org/viaf/18015989"
  xlink:arcrole="https://blogs.brown.edu/hallhoag/eac-cpf/role/">
    <relationEntry>Blake, Eugene Carson </relationEntry>
  </dateRange>
  <fromDate ></fromDate>

```

```

    <toDate ></toDate>
  </dateRange>
    <descriptiveNote><p>Clerk of the General
Assembly</p></descriptiveNote>
  </cpfRelation>
  <cpfRelation cpfRelationType="associative" xlink:type="simple"
xlink:role="http://RDVocab.info/uri/schema/FRBREntitiesRDA/Person"
xlink:href="http://viaf.org/viaf/100272137"
xlink:arcrole="https://blogs.brown.edu/hallhoag/eac-cpf/role/">
  <relationEntry>Brown, Robert McAfee </relationEntry>
  <descriptiveNote><p>Member/Theologian</p></descriptiveNote>
</cpfRelation>
  <cpfRelation cpfRelationType="associative" xlink:type="simple"
xlink:role="http://RDVocab.info/uri/schema/FRBREntitiesRDA/Person"
xlink:href="http://viaf.org/viaf/100176316"
xlink:arcrole="https://blogs.brown.edu/hallhoag/eac-cpf/role/memberOf">
  <relationEntry>Eisenhower, Dwight D. </relationEntry>
  <descriptiveNote><p>Member</p></descriptiveNote>
</cpfRelation>
  <resourceRelation resourceRelationType="creatorOf"
xlink:href=""
xlink:type="simple">
  <relationEntry>Records, ca. 1950-1972 </relationEntry>
<placeEntry>Prebyterian Historical Society </placeEntry>
  </resourceRelation>
</relations>
</cpfDescription>
</eac-cpf>

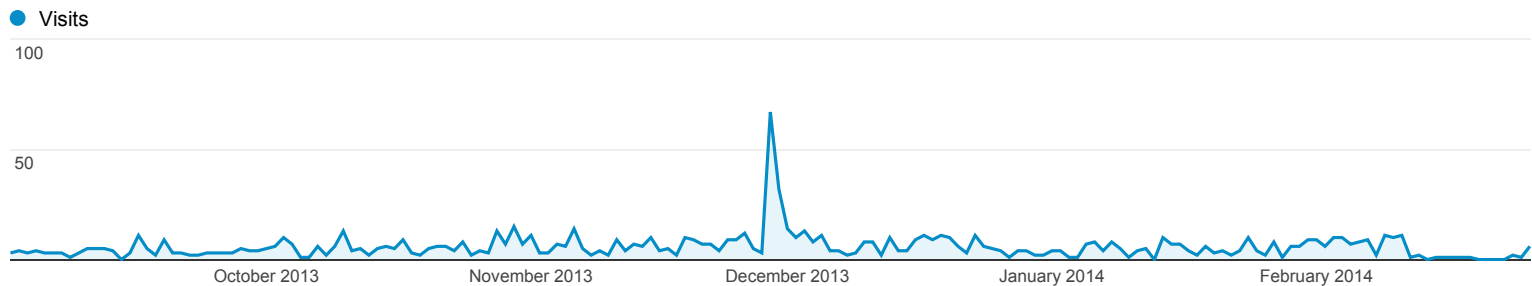
```

Sep 1, 2013 - Feb 26, 2014

Audience Overview

temp-HallHoag
1.46%

Overview

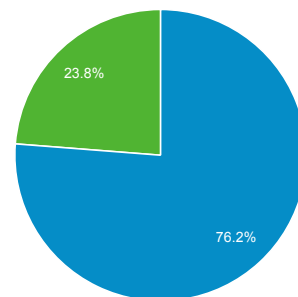


Visits
temp-HallHoag
1,027

Unique Visitors
temp-HallHoag
826

Pageviews
temp-HallHoag
2,178

■ New Visitor ■ Returning Visitor



Pages / Visit
temp-HallHoag
2.12

Avg. Visit Duration
temp-HallHoag
00:02:12

Bounce Rate
temp-HallHoag
66.70%

% New Visits
temp-HallHoag
76.24%

Language	Visits	% Visits
1. en-us	805	78.38%
2. ja-jp	79	7.69%
3. ja	41	3.99%
4. en-gb	19	1.85%
5. en	15	1.46%
6. fr	13	1.27%
7. pt-br	6	0.58%
8. de-de	5	0.49%
9. nb-no	5	0.49%
10. de	4	0.39%

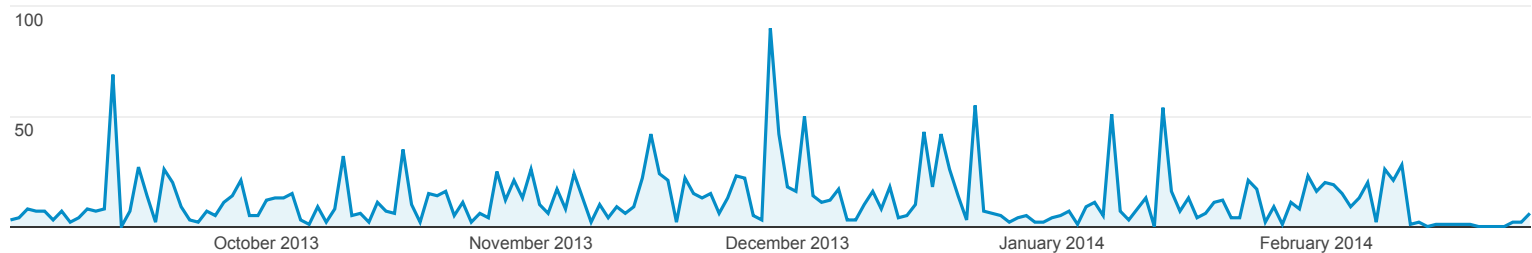
Sep 1, 2013 - Feb 26, 2014

Overview

temp-HallHoag
1.48%

Overview

Pageviews



Pageviews
temp-HallHoag
2,178



Unique Pageviews
temp-HallHoag
1,731



Avg. Time on Page
temp-HallHoag
00:01:58



Bounce Rate
temp-HallHoag
66.70%



% Exit
temp-HallHoag
47.15%



Page	Pageviews	% Pageviews
1. blogs.brown.edu/hallhoag/	392	18.00%
2. blogs.brown.edu/hallhoag/2013/06/25/black-panther-party/	206	9.46%
3. blogs.brown.edu/hallhoag/gordon-hall-and-grace-hoag/	87	3.99%
4. blogs.brown.edu/hallhoag/background/	72	3.31%
5. blogs.brown.edu/hallhoag/category/eac-cpf/	55	2.53%
6. blogs.brown.edu/hallhoag/2013/10/10/761/	51	2.34%
7. blogs.brown.edu/hallhoag/2013/04/24/jews-for-the-preservation-of-firearm-ownership/	48	2.20%
8. blogs.brown.edu/hallhoag/2013/01/28/john-birch-society/	37	1.70%
9. blogs.brown.edu/hallhoag/2013/11/14/national-white-americans-party/	35	1.61%
10. blogs.brown.edu/hallhoag/category/highlights/	33	1.52%