April 27, 2017

2- Nonprofit Investigative Journalism Content Analysis Section Final Report

Matthew L Hale

Nonprofit Investigative Journalism:
A Snapshot of Content and Reach

An Exploratory Research Report sponsored by
The Jonathan Logan Family Foundation

By
Dr. Matthew Hale
Associate Professor and MPA Director
Seton Hall University, Center for Public Service

&

Joseph Riccardelli & Lisa DeLuca
Seton Hall University, Center for Public Service

with research assistance provided by
Ryan Stetz,
Seton Hall University

Content Analysis Results

Content Analysis Sample

This study examines a total 2,309 stories that were extracted from the 2016 archives of nine separate nonprofit investigative news organizations. These organizations represent three distinct types of nonprofit investigative news outlets; national, state or local and academic-centered organizations. The organizations are listed in Table One:

Table One: Source of Stories

Source	N	% of total
National	**1284**	**55.6%**
ProPublica	553	23.9%
Center for Investigative Reporting	441	19.1%
Center for Public Integrity	290	12.6%
State/Local	**864**	**37.4%**
Arizona Center for Investigative Reporting	13	0.6%
NJ Spotlight	693	30.0%
Inewsource.org	158	6.8%
Academic	**161**	**7.0%**
The Investigative Reporting Workshop at American University	59	2.6%
The Wisconsin Center for Investigative Journalism	81	3.5%
The Investigative Reporting Program (UC Berkeley)	21	.9%
Total	**2,309**	**100%**

Story Sampling Methodology

A few clarifications concerning the overall sample are important. Our goal was to capture all of the "stories" produced by these nine organizations in calendar year 2016. To do this we explored all of the story archive functions for each outlet. We took an expansive view of what constitutes a "story" by including regular news stories, podcast, video stories and even blog posts and social media notifications in our sample. Despite this effort to be expansive and inclusive in what we captured and analyzed there are three specific types of stories that are not included in the sample. Inclusion of these "outlier" stories would have skewed the sample. Descriptions of each type of outlier follow.

First, we excluded stories about the Panama Papers. This major international investigative collaboration revealed a maze of legal (and illegal) ways that elites around the

world are able to hide money from various government authorities. The Center for Public Integrity (https://www.publicintegrity.org/) was a significant partner in this effort and virtually all of the organizations in our sample covered it to some extent. However, because this was such a major event and garnered such significant mainstream media coverage, including these stories would very likely artificially skew the results. We have therefore excluded stories concerning the Panama Papers from the sample.

Second, in an effort to include a wide range of nonprofit news organization in our sample we elected to include NJ Spotlight (http://www.njspotlight.com/) as representing a state/local nonprofit investigative news organization. As we started conducting our analysis of NJ Spotlight stories it became apparent that while the NJ Spotlight produces a significant number of investigative and explainer pieces, it also produces far more straight news and op-ed interviews that the other organizations. Overall, we coded more than 1800 stories for NJ Spotlight which if we were to include would constituted around 50% of all the stories analyzed. As a result, including them all would clearly skew the results. We, therefore, decided to exclude the straight news and op-ed stories produced by NJ Spotlight from the overall sample. Even after excluding these, the story universe still contains over 600 stories from NJ Spotlight. However, we felt this was a reasonable compromise and preferable to excluding NJ Spotlight entirely.

Third, ProPublica (https://www.propublica.org/) recently added a locally focused site in Illinois and has plans to increase their presence in the state and local markets. The ProPublica stories we extracted were all from the archive function on their main site.

Methodology

The content analysis instrument was developed by the principal investigator based on previous work conducted at the USC Annenberg's Norman Lear Center (see https://learcenter.org/project/news/pubs/). A full copy of the instrument is available upon request. The principle investigator and a senior graduate student extracted stories from the various archives. The principal investigator and two graduate students then conducted the primary content analysis. The two graduate students received training in use of the coding instrument and four separate tests for inter-coder reliability tests were conducted on a sample

of stories. The standard statistics for measuring inter-coder agreement are Scott's Pi and Chronbach's alpha. These both measure improvement in agreement that coders have over what they would have based on random chance. In this case, inter-coder agreement on both measures exceeded .80, which is the standard acceptable measure.

After reaching an acceptable level of inter-coder reliability, the two graduate students each coded approximately ½ of the stories. During the coding process, each coder received a sample of the same stories to check for coder "drift." Thankfully, none occurred. The students were able to ask the principle investigator questions on any story and if students remained unsure on how to code a specific variable, the principle investigator made the final decision.

It is important to recognize a few points regarding the coding procedures. First, coders examined the stories "as if" they were normal news consumers. This means that they did not read each story multiple times or parse out all of the minute detail that may exist in each story. The rationale for this is that it best approximates how consumers engage with news material, while at the same time allowing for the recording of important details. This strategy has a number of implications.

First, coders centered on and coded for what was "primary" in each story. Many of the stories contained a wide variety of topics in them. For example, imagine a story centered on a particular criminal justice issue but at the end of the story, it mentions comments made by the 2016 presidential candidates about that issue. In this case, the story would likely be coded as focusing on crime or criminal justice issue and not the 2016 presidential election. It is possible that future research could provide a more specific and detailed analysis of all of the sub-topics covered in these stories.

The second implication of this consumer-simulation approach is that coders focused on coding "explicit" mentions for many of the variables. For example, stories are often written or co-produced with partner news organizations. Similarly, other stories are part of a multi-segment or series of stories. As coders were reading stories they knew to look for partnership references and references to the stories being in a series in the text, video or audio they were analyzing. However, they did not to scour the stories for references to partnerships or a series or search the entire story frame for links to other organizations. We believe that the coders

caught most of the partnerships and "series" stories but it is possible that we undercounted these variables to some extent if the references to these variables were not easily or immediately apparent.

The third and final coding instrument clarification is on the "story type" variable. The options on this variable are 1) a straight news story 2) "explainer" story 3) investigative journalism piece 4) data journalism 5) op-ed/Interview or 6) "other" format. The data journalism, op-ed/interview and the "other" formats are fairly straightforward and self-explanatory. However, the "line" between a "news story", an "explainer story", and an "investigative journalism story" is admittedly somewhat subjective. It is important to note that what matters for this research is that the coders were able to agree on the distinctions described below and code the stories accordingly.

One way of conceptualizing the difference between the three types of stories is on a two-dimensional (depth and time) continuum. A straight news story is the most immediate and current story. It describes who, what, where, when and how of a particular event. For example, a straight news story might report the results of a federal election commission fundraising filing and report how much money each candidate had raised over a given period.

An explainer story would be "next" on this depth and time continuum. The explainer story might focus more attention on a specific component of the original story and report on that facet in more explanatory detail. For example, an explainer story might examine candidate fundraising within a particular industry or from specific individuals and explain how the amount given to each candidate by that specific industry or individual helps better understand what the industry/individual wants from both candidates.

Finally, an investigative journalism story would take the same basic premise of this story but it would conduct original and much more in-depth research on the topic. So for example, an investigative journalism piece might examine the process known as "bundling" campaign donations and through secondary sources or by acquiring previously un-released government documents identify ways that specific industries or individuals may be circumventing public disclosure laws regarding campaign financing. In the development and training process, coders

received examples of each type of story and were always able to ask for assistance in coding this variable as they went along.

With those explanations, we now turn to the specific results.

Primary Story Topics

The most common story topic over the entire sample was government/non-elections stories, which accounted for just under one out of five stories (17.5%). The next most common story topic were those focusing on the 2016 Presidential election (10.7%) and stories about health or health care (10.7%). Stories about housing (2.4%) national security (2.2%) international news (1.9%) and natural disasters/catastrophes (2.4%) were the least common story topics.

Tables 2 through 5 contain the overall results and the results for each of the three types of outlets (national, state/local and academic) in terms of primary story topic.

Table 2: Overall Sample Top & Bottom Topics

Topic	N	% of total
Government/Non-election	403	17.5%
Presidential Election	248	10.7%
Health/HealthCare	247	10.7%
Business or Economy	173	7.5%
Story or Brand Promotion	173	7.5%
Legal Issues	165	7.1%
Housing	55	2.4%
National Security	50	2.2%
International News	44	1.9%
Natural Disaster/Catastrophes	34	1.5%

192 stories were coded as "other" these were stories that combine topics or stories focusing on other topics like animal cruelty or the Olympics. These stories are not included in the above table.

Table 3 National Sources Top & Bottom Topics

Topic	N	% of total
Government/non-election	204	15.9%
Presidential election	189	14.7%
Story or Brand Promotion	125	9.7%
Legal Issues	114	8.9%
Health/Health care	98	7.6%
Crime	88	6.9%
Other elections	40	3.1%
Housing	40	3.1%
International News	38	3.0%
Natural Disaster/Catastrophes	25	1.9%

108 stories (8.4%) were coded as "other" or combination story topics. *These stories are not included in the above table.*

Table 4: State/Local Top & Bottom Topics

Topic	N	% of total
Government/Non-election	181	20.9%
Health/Health care	142	16.4%
Business or economy	92	10.6%
Education	85	9.8%
Other elections/Politics	76	8.8%
Environment	70	8.1%
Housing	15	1.7%
Story or Brand Promotion	11	1.3%
Catastrophes	6	.7%
National security	2	.2%

60 stories (6.9%) were coded as "other" or combination stories. *These stories are not included in the above table.*

Table 5: Academic Top & Bottom Topics

Topic	N	% of total
Story & Brand Promotion	37	23.0%
Government/Non-election	18	11.2%
Crime/Criminal Justice	15	9.3%
Environment	12	7.5%
Legal issues	10	6.2%
Health/Health care	6	6.1%
Business or economy	4	2.5%
Other elections/politics	4	2.5%
Catastrophes	3	1.9%
National security	3	1.9%

24 stories (14.9%) were coded as a combination or "other" category. *These stories are not included in the above table.*

The specific outlet results provide a number of interesting findings. These are discussed below.

Domestically Focused Organizations

Overall, the nine organizations generally ignored international news and news focused on national security or the military. The vast majority of stories across all outlets focused on U.S. domestic issues. This is quite consistent with previous research on for-profit mainstream media outlets in the United States (see https://learcenter.org/project/news/pubs/). In contrast, the nine organizations did not pay a great deal of attention to stories about natural disasters and catastrophes (fires, floods), which the same research shows are routinely found in large quantities on local television news and mainstream media.

Presidential Election Coverage Concentrated at National Level Organizations

While overall coverage of the 2016 presidential election was the second most common topic it was by far the most prevalent in the three national outlets. Overall 76.2%, of all of the stories focused on the presidential election appears on one of the three national outlets. The state/local outlets accounted for 20.2% of the Presidential election stories and the academic centers just 3.6%. This finding suggests that organizations particularly at the state/local level are carving specific areas in deciding what stories they will cover.

Story and Brand Promotion Content

All of the sites produce stories, blogs and other forms of content that are designed to either promote an individual story or some aspect of the organization's brand. For example, many of these stories described awards recently won by the outlet. Other promotion stories, particularly from the academic and smaller outlets point web site viewers to partner outlets where a co-produced story was actually released. Finally, some promotion stories provide information about live events held by the organizations. While the idea of "promotion" may sound at odds with news and investigative organizations these stories are clearly necessary for nonprofit organizations attempting to gain financial support, get their work noticed and even more importantly to engage their audience with the work of the organization.

Types of Stories

The most common type of story were "explainer" stories. As described above, these stories provide consumers with a more thorough explanation of a particular topic compared to a "straight news story" but do not reach the level of depth as an "investigative journalism story". Overall, explainer stories accounted for 41.4% of all stories. Straight news reporting was the second most common story type (27.6%), followed by traditional in-depth investigative reporting pieces (18.8%). While explainer stories were the most common in the entire sample, straight news reporting stories were actually the most common among the 3 national outlets (37.9%) and the 3 academic outlets (44.5%). Stories focused on the presentation of data or data journalism and op-ed/interview centered stories each accounted for about 4.5% of the total.

Tables 6 through 9 provide specific data on the type of stories each type of organization produced. When examining these tables perhaps the most interesting observation is how consistent the amount of stories coded as investigative journalism is across different type o outlet. Around 20% of all stories were investigative pieces across all three types of outlets.

Table 6: Type of Story Overall Sample

Type of Story	N	% of total
"Explainer" Story	957	41.4%
News reporting	638	27.6%
Investigative Journalism	433	18.8%
Op-ed/Interviews	109	4.7%
Data Journalism	104	4.5%
"Other"	68	2.9%

Table 7: Type of Story National Sources

Type of Story	N	% of total
News Reporting	487	37.9%
Explainer Story	350	27.3%
Investigative Journalism	256	19.9%
Op-ed/Interviews	87	6.8%
"Other"	57	4.4%
Data Journalism	47	3.7%
Total	1284	100%

Table 8 Type of Story State and Local Sources

Type of Story	N	% of total
Explainer Story	574	66.4%
Investigative Journalism	151	17.5%
News reporting*	80	9.3%
Data Journalism	55	6.4%
Op-ed/Interview*	3	0.3%
Other	1	0.1%

*NJ Spotlight stories in these categories are not included in the study.

Table 9: Type of Story Academic Sources

Type of Story	N	% of total
News Reporting	71	44.1%
Explainer stories	33	20.5%
Investigative Journalism	26	16.1%
Op-ed/Interview	19	11.8%
Other	10	6.2%
Data Journalism	2	1.2%

Topic and Type of Stories

When comparing the story topics with types of stories a somewhat different pattern emerged. Non-election stories about government remained the most common topic across explainer, straight news and investigative stories. Among stories coded as data journalism, the presidential election was the most common topic accounting for 27.9% of these stories. Interestingly among the op-ed/interview category, the most common story topic was crime/criminal justice, which accounted for 15.6% of all of these stories. Table 10 contains these results.

Table 10: Story Topic by Story Type

Top Topics in Explainer stories	Top Topics in Investigative Journalism Stories*	Top topics in News Stories	Top Topics in Data Journalism Stories	Top Topics in Op-Ed/ Interview Stories
Government/ Non-election (18.6%)	Government/ Non-election (15.9%)	Government/ Non-election (21.9%)	Presidential Election (27.9%)	Crime/Criminal Justice (15.6%)
Presidential Election (13.0%)	Health/Health care (13.6%)	Legal Issues (14.9%)	Health/Health care (14.4%)	Story/Brand Promotion (13.8%)
Health/Health care (12.1%)	"Other" Story Topics (10.4%)	Story/Brand Promotion (9.6%)	Non-Presidential Elections (11.5%)	Presidential Election (11.9%)
Business Economy (9.8%)	Crime/Criminal Justice (9.5%)	Health/Health Care (7.4%)	Business/Econ (10.6%)	Business/Econ (9.2%)
Non-presidential elections (7.7%)	Presidential Election (7.6%)	Presidential Election (7.2%)	Government/Non-election (6.7%)	"Other" story topic (9.2%)

*In the investigative Journalism Stories Education (7.4%) and Environment (7.2%) were about equal with presidential election stories

Story Delivery Method

The vast majority of stories (around seven out of 10 stories) across outlet type were delivered in a traditional written format, or just like a regular newspaper or magazine article. Table 11 provides information about story delivery methods for each type of outlet. The most surprising result is perhaps the apparent lack of video content and social media content within the sample.

Table 11: Story Delivery method by Outlet Type

Delivery Methods	Overall Sample	National Outlets	State/Local Outlets	Academic Outlets
Traditional written	73.5%	75.5%	70.1%	74.5%
Blog	11.4%	7.3%	15.2%	23.6%
Audio Focused (Podcast)	6.4%	9.3%	3.0%	0.6%
Data Visualizations	5.6%	3.3%	10.1%	0.0%
Video Focused (vblogs)	1.5%	1.5%	1.6%	1.2%
Social Media	1.4%	2.6%	0.0%	0.0%
Other Format	0.2%	0.4%	0.0%	0.0%

Partners

The coders looked for explicit references concerning media organization partnerships between in each story. A total of 99 different news organizations were explicitly mentioned as partnering with the 9 nonprofit organization in the sample. However, just 13 of these organizations (TIME, Huffington Post, Washington Post, Virginian Pilot, Texas Tribune, NBC News, NPR, WNYC, Al Jazeera America, New York Times, Mother Jones, New York Daily News, and PRI) accounted for 52.3% of all explicit partnership mentions. The remaining 47.3% of the partnership mentions were split across the other 86 organizations. Table 12 contains the percentage stories that mentioned a partnership. Table 13 indicates the number of partnership mentions for the top 12 organizations in the sample.

Table 12: Partnership mentions

Partnership Mentioned in the Story	Overall Sample	National Outlets	State/Local Outlets	Academic Outlets
Yes	16.1%	23.2%	2.7%	31.1%
No	83.9%	76.8%	97.3%	68.9%

Table 13: Who are the Partners?

Partner	Number of Mentions	% of all Mentions
Time	42	10.0%
Huffington Post	23	5.5%
Washington Post	21	5.0%
Virginian Pilot	19	4.5%
Texas Tribune	18	4.3%
NBC News	16	3.8%
NPR	16	3.8%
WNYC	15	3.6%
Al Jazeera America	14	3.3%
New York Times	14	3.3%
Mother Jones	11	2.6%
New York Daily News	11	2.6%
PRI	11	2.6%
Total	231	52.3%

86 other partner outlets accounted for the remaining 47.7% of all direct partner mentions