Sources of Advocacy: When Does the Media Give Voice to Egyptian Advocacy NGOs?

Andrew Heiss
Duke University
andrew.heiss@duke.edu

Ken Rogerson
Duke University
rogerson@duke.edu

We use quantitative text analysis methods to examine how a subset of advocacy NGOs were used as sources by Egyptian English-language media publications in the months following the 2011 uprising. Using a combination of manual classification and fully automated computational topic modeling, we reveal descriptive insights into patterns of sourcing by both state-owned and independent Egyptian media outlets. We find that NGOs are used as sources sparingly in comparison to overall journalistic output, but that the reporting that referenced these NGOs was not necessarily superficial or shallow. There is evidence that state-run media tends to be more superficial in its use of NGOs as sources while independent media are more likely to identify specific individuals and use direct quotes and paraphrases.

On February 20, 2011, Jeffrey Ghannam claimed that social media proved to be a "far bigger threat" to Egyptian president Hosni Mubarak than other forms of protest. Recognizing that social media cannot change politics by themselves, he noted that the media were "chronicling and amplifying the revolution that is happening on the streets."[1] The potential impact of media—both traditional and social—on the tumultuous political changes in Egypt and the Arab Spring writ large have grabbed the attention of social science researchers with good reason.

The relationship between the media and revolution, social movements, and political protests has been an attractive topic for generations of scholars: from Billington's

---

[1] Jeffrey Ghannam, "Freedom, Beyond 140 Characters," *The Washington Post* (February 20, 2011): B3.

Fire in the Minds of Men which explored the relationship between media and revolutions,[2] to Deutsch's famous work on nationalism, social communication, and the uses of the media for smaller scale political protest during the Cold War in Eastern Europe.[3] Some research is more descriptive and country specific[4] while other researchers explore the nuances of different types of media in different situations. At times, these situations are social movements that seem revolutionary. At other times, they take place on a smaller scale, through organized and/or spontaneous political protest or lobbying the existing political establishment.

In the past few years, the Middle East has risen to the forefront of this discussion through the Arab Spring: a series of political changes, some successful and others not. The role of media in the Arab Spring—especially social and independent media—has been at the core of broader global debates. Cheerleaders praise the value of independent media in political change[5] while others caution that the media's impact may be less important than originally thought—or even nonexistent in some cases.[6]

The relationship between the media and NGOs

The expectation for the media's role will be different in democratic and autocratic societies, but in both the relationship between social movements (which we operationalize as NGOs) and the media is always unequal. Social movements suffer from "asymmetrical dependency"[7]—movements need media but media do not necessarily need to cover social movements to fulfill their mission of disseminating information. To gain media support for their causes, movements change and reframe their message and mission to be more palatable for media audiences. Further, under autocratic regimes specifically, opposition organizations and independent media must navigate enhanced government scrutiny in order to spread their messages.

[2] James Billington, *Fire in the Minds of Men: Origins of the Revolutionary Faith* (New York: Basic Books, 1980).

[3] Karl Deutsch, *Nationalism and Social Communication* (Cambridge, Massachusetts: The MIT Press, 1958).

[4] See William A. Rugh, *The Arab Press* (Syracuse, New York: Syracuse University Press, 1987); Festus Eribo and William Jong-Ebot, *Press Freedom and Communication in Africa* (Trenton, New Jersey: Africa World Press, 1997).

[5] Xose Soengas, "The Rold of the Internet and Social Networks in the Arab Uprisings - an Alternative to Official Press Censorship," *Comunicar* 21, no. 4 (2013): 147–155.

[6] Alexa Robertson, "Connecting 'Old' and 'New' Media and the Arab Spring," *The International Journal of Press/Politics* 18, no. 3 (2013): 325–341.

[7] William A. Gamson and Gadi Wolfsfeld, "Movements and Media as Interacting Systems," *Annals of the American Academy of Political and Social Science* 528 (July 1993): 125.

One reason for the asymmetry is the media's reliance on news values,[8] or criteria by which news organizations decide what is worth covering. Many social groups understand that conflict is often the most prominent news value. For example, groups that engage in violent activity often do so to attract media attention. While media are not dependent on this type coverage to survive, groups who engage in violent activities, such as terrorists, learn to exploit the media's desire for sensationalism and the coverage of vicious, disruptive actions. Given this relationship, even though it is not explicitly agreed to, violent organizations are able to showcase their actions to global audiences, disseminate their messages to interested parties, and gain recognition and legitimacy.[9]

Violent groups are not the only organizations that seek out media attention. Non-governmental organizations (NGOs) face stiff competition for resources and attention both domestically and internationally. Human rights causes that are adopted by the global advocacy community are not always those with the most compelling normative claims. Rather, organizations can successfully champion their causes only if they navigate the harsh Darwinian marketplace for attention. Clifford Bob has argued that "tacitly and at times openly, needy groups vie with one another for the world's sympathy, elevating themselves above their competitors and differentiating themselves from similar causes."[10] Access to the media is a critical element of NGO strategy.

Online technologies have had an impact on media activity, though, for this project, media include both print and online. The internet is a form of media that produces, consolidates and shares information with the "masses." Some researchers argue that its growth as a news source serves as a challenge to traditional media, since many of the previous constraints and controls by corporations or the government are more difficult to apply online.[11] Gance et al. argue that the internet provides a space for those to speak up who have typically been silenced by the military or government-controlled media.[12] Schuler and Day discuss how the internet can expand the public sphere to allow for democratization on a larger and larger scale.[13] It is clear from this literature that both independent media outlets and non-governmental organizations seek to spread distinct messages.

[8]Leonard Downie and Robert Kaiser, *The News About the News* (New York: Vintage Press, 2003), chap. 8.

[9]Brigitte L. Nacos, *Mass-Mediated Terrorism: The Central Role of the Media in Terrorism and Counterterrorism* (Lanham, Maryland: Rowman & Littlefield, 2007), 91–108.

[10]Clifford Bob, *The Marketing of Rebellion: Insurgents, Media, and International Activism* (Cambridge, MA: Cambridge University Press, 2005).

[11]Annabelle Sreberny and Gholam Khiabany, *Blogistan* (London: I.B. Tauris & Co. Ltd., 2010).

[12]Stephen Gance, An De Vaney, and Yan Ma, *Technology and Resistance* (New York: Peter Lang Publishing, 2000).

[13]Douglas Schuler and Peter Day, eds., *Shaping the Network Society* (Cambridge, Massachusetts: The MIT Press, 2004).

Reporters and news sources

In the ongoing Arab Spring, civil society and advocacy NGOs have been touted as crucial sources of information and commentary on political and social movements. Reciprocally, organizations have attempted to use the media to broadcast their advocacy agendas to a wider audience. However, given the asymmetry between media and social movements, when do journalists turn to these NGOs for information, interpretation, or commentary?

There are two theoretical points to consider: (1) where do NGOs fit in the scheme of news sources, and (2) when they are used, how are they used? Though a comprehensive typology of news sources does not yet exist, scholars have recently made a distinction between political/institutional and civil society sources.[14] Reporters are more likely to pay attention to elite, public sources, such as government officials or the police.[15] In 1978, Stuart Hall et. al. proposed a "pyramid of access" that describes those with more or less connection to journalists. Following the thinking of Stuart Hall, et al.,[16] civil society organizations like NGOs find themselves at the bottom of the pyramid of journalistic access, and are thus at disadvantage for both gaining access to the media and coming across as credible, helpful news sources.

Given that NGOs are less likely to be used as sources, under what circumstances do journalists turn to them for information? One possible explanation is the journalistic notion of objectivity and the practice of telling all sides of a story, following Lacy et al. who conclude that "sources provide information and interpretation about news stories and they add credibility to truth claims. Source selection is crucial to story balance, a norm for American journalism."[17] It may be that NGOs in authoritarian regimes are also called on to provide this balance to official—often state-dominated—sources.

From this existing body of research, we can hypothesize that (1) media organizations will cover a story if there is conflict present, (2) the sources for that story will be more likely to include elite or political sources rather than civil society sources, and (3) when more social sources sources are used—such as NGOs—they will be used sparingly with superficial context and shallow nuance.

In this paper, we explore the third hypothesis—how are NGOs been used as news sources. To do so, we use quantitative text analysis methods to examine how a subset

[14]Rodney Tiffin et al., "Sources in the News: A Comparative Study," *Journalism Studies* 15, no. 4 (2014): 374–91; see also Paul Manning, *News and News Sources: A Critical Introduction* (London: SAGE Publications, 2001).

[15]See the indexing hypothesis in Lance Bennett, "Toward a Theory of Press/State Relations," *Journal of Communication* 40, no. 2 (1990): 103–25.

[16]Stuart Hall et al., *Policing the Crisis: Mugging, the State and Law and Order* (London: Palgrave Macmillan, 1978).

[17]Stephen Lacy et al., "How Radio News Uses Sources to Cover Local Government News and Factors Affecting Source Use," *Journalism & Mass Communication Quarterly* 90, no. 3 (September 2013): 460.

of advocacy NGOs were used as sources by Egyptian English-language media publications in the months following the 2011 uprising. Political and media texts provide prime corpora for data analysis, but robust and reproducible computational and statistical methods have only recently been developed and have been underapplied in the political science and communications literature. Using a combination of manual classification and fully automated computational topic modeling, we reveal descriptive insights into patterns of sourcing by both state-owned and independent Egyptian media outlets.

Data

English-language Egyptian news organizations

Despite the authoritarian nature of former president Hosni Mubarak's regime, Egypt has long had an active journalism sector, with dozens of widely circulated Arabic-language newspapers—Al-Ahram, the official news outlet for the state, has published continuously since 1875. In the late 2000s journalists and private sector investors opened English-language publications either as entirely new ventures or as spin-offs of already existing Arabic newspapers. We focus on three of the most prominent publications: al-Ahram English, Daily News Egypt, and Egypt Independent.

In the late 1990s, Al-Ahram Arabic began publishing a separate English newspaper, and following a move to online distribution in the mid 2000s, al-Ahram English has become the most read English-language publication in the Middle East.[18] A group of independent journalists formed Daily News Egypt in 2005 and quickly built up a wide audience, securing distribution rights with the International Herald Tribune. In May–June 2012, the ownership of the paper changed hands, but it has continued to publish daily both in print and online and continues to claim to be "the only independent English-language printed daily in the country."[19] In the late 2000s, the independently owned Arabic daily al-Masry al-Youm began including a weekly English supplement, which evolved into Egypt Independent, published in print and online beginning in November 2011 under the direction of an independent editorial board. However, al-Masry al-Youm's owners unexpectedly shuttered the publication in April 2013 following alleged government pressure. Egypt Independent has since resumed publication, but as a weekly English-language supplement under the editorial control of al-Masry al-Youm Arabic. In mid-2013, former Egypt Independent staff members

---

[18]"Al-Ahram tops online newspapers in Arab world: Forbes," *al-Ahram English*, December 27, 2012, http://english.ahram.org.eg/NewsContent/1/64/61353/Egypt/Politics-/AlAhram-tops-online-newspapers-in-Arab-world-Forbe.aspx.

[19]"Daily News Egypt: Final Words," *Daily News Egypt*, April 22, 2012, http://thedailynewsegypt.com/2012/04/22/daily-news-egypt-final-words/

formed Mada Masr, a new online-only independent media outlet that slowly grown in prominence and audience.[20]

Egyptian NGOs

Egypt also has a long history of associational life and civil society. Islamic religious endowments (*awqāf*) have provided substantial non-governmental social services since medieval times, and Western-style charitable, civic, and advocacy associations emerged in the 1800s.[21] While many of these organizations were closed or marginalized following the rise of Gamal Abdel Nasser in 1954, in the 1980s Hosni Mubarak greatly expanded the purview of civil society and permitted thousands of NGOs and other organizations to apply for official incorporation. However, the government reserved the right to shutter and ban any organization that engaged in political activity (or any activity against the state), in addition to a host of other stifling regulations.[22] The emerging civil society sector took advantage of this newfound openness, but at the cost of subjecting itself to state authority. Law 84 of 2002, or the Law on Non-Governmental Societies and Organizations, granted formal civil society organizations the legal permission to operate, but within fairly restrictive boundaries. Law 84/2002 established a set of "red lines" that organizations were forbidden to cross—organizations focused on social reform, political liberalization, or civil rights advocacy were not legally allowed to pursue any agenda contrary to the regime. Despite these restrictions, however, dozens of advocacy and civic organizations have been active in Egypt (often backed by foreign interests) and have carefully navigated or circumvented governmental regulations to do so.[23] There are now an estimated 20,000–30,000 registered NGOs in the country, including labor unions, charities, advocacy organizations, human rights groups, policy and legal think tanks, and religious organizations, thus comprising perhaps the largest NGO sector in the region.

Data collection and NGO identification

To understand how the English-language Egyptian media cover the activities of advocacy NGOs, we collected all articles published online by state-owned Al-Ahram En-

---

[20]See http://www.madamasr.com/.

[21]Amani Kandil, Stefan Toepler, and Lester M. Salamon, "Egypt," in *Global Civil Society: Dimensions of the Nonprofit Sector*, ed. Lester M. Salamon and S. Wojciech Sokolowski, vol. 2 (Bloomfield, CT: Kumarian Press, 2004), 217–26.

[22]Amaney Jamal, *Barriers to Democracy: The Other Side of Social Capital in Palestine and the Arab World* (Princeton: Princeton University Press, 2007), 124.

[23]Sarah Yerkes, "Capture, Co-Optation, and Pluralism: Navigating the Civil Society Arena in the Arab World" (PhD thesis, Georgetown University, 2012).

glish[24] and independently published Daily News Egypt[25] and Egypt Independent[26] between November 24, 2011 and April 25, 2013, the first and final days of Egypt Independent's regular weekly coverage under an independent editorial board. We limit the coverage of all three publications to Egypt Independent's lifetime to analyze the maximum number of English-language articles published simultaneously by the three outlets (i.e. to avoid looking at events that Egypt Independent could not have covered). The corpora for these three publications include both the text of each published article and metadata such as publication date, author name, subject tags, URL, and word count.

We then selected a subset of 618 articles that mention a group of self-identified advocacy NGOs with missions focused on civil liberties and human rights. On May 30, 2013, these NGOs signed an open letter originally penned by the Egyptian Initiative for Personal Rights (EIPR), expressing deep concern over new draft legislation to severely limit the power of the civil society sector, prohibiting foreign donations and funding, requiring more onerous state registration, imposing heavy fines for minor infractions, and other restrictions.[27] While this list of signatory organizations is not representative of the full Egyptian civil society sector, or even all Egyptian human rights organizations,[28] the fact that these organizations signed an English-language open letter, published online and publicized by English-speaking Twitter users, bloggers, and reporters, with the clear goal of reaching an English speaking audience, is indicative of a self-selecting preference toward English-language media. We feel that this admittedly non-comprehensive subset is a good sample of Egyptian advocacy NGOs that regularly engage with the English-speaking Egyptian media.

Methods

Close textual analysis of every article in our corpus is infeasible and impractical. However, recent developments in computational linguistics have led to new methods for analyzing large quantities of text. Natural language algorithms can identify and extract parts of speech, parse sentences to determine which actors do which actions (the "who did what to whom" that forms the foundation of event data), calculate relative positive or negative emotional sentiment, and uncover latent topical structures

---

[24]See http://english.ahram.org.eg.

[25]See http://www.dailynewsegypt.com.

[26]See http://www.egyptindependent.com.

[27]Egyptian Initiative for Personal Rights, "The Muslim Brotherhood Lays the Foundations for a New Police State by Exceeding the Mubarak Regime's Mechanisms to Suppress Civil Society," May 30, 2013, http://www.eipr.org/en/pressrelease/2013/05/30/1720.

[28]Finding a comprehensive list of human rights-oriented NGOs is unfortunately impossible as the Ministry of Social Affairs does not provide a public list of registered organizations.

within a text. Non-computational analysis large corpora of text is generally difficult and costly to implement—manual content analysis often requires multiple coders to closely read each document in a corpus, creating issues of consistency and repro-ducibility.

The promise of robust and accurate computational text analysis is alluring, but algorithmic methods are still limited by the complexities of language. For example, computational context-free grammars can successfully parse a sentence such as "The NGO issued a report yesterday," identifying the sentence's subject, verb, direct object, and time frame. However, algorithms cannot parse sentences where words are used metaphorically, such as in the joke "Time flies like an arrow; fruit flies like a banana,"[29] or in cases where words play syntactically challenging roles, such as the classic sen-tence "The horse raced past the barn fell." Even less contrived examples present dif-ficulties in parsing, such as "Mohammed el-Masri, the spokesman for an NGO that participated in the human rights conference last year, said that the new NGO law is 'difficult and restrictive and I disapprove of it.' "—nested dependent clauses make it difficult for an algorithm to determine who or what is being quoted and could easily lead to false positives in event data (i.e. Did the spokesman participate in the con-ference? Is 'NGO' a proper name? Does the spokesman or proper-name-'NGO' dis-approve of the law?). Similarly, algorithms cannot pick up on extra-textual details such as specific word connotations or context—sarcastic texts are almost likely to be misclassified.

Despite these limitations, computational text analysis can still be usefully applied to large bodies of political texts as long as its deficiencies are recognized and accounted for. To ensure that our methods are appropriate and robust, we apply Grimmer and Stewart's four-fold framework for automated political text analysis in order to to prop-erly blend computational and human analytical methods. We describe our methods according to the following principles:

1. All quantitative models of language are wrong—but some are useful.
2. Quantitative methods for text amplify resources and augment humans.
3. There is no globally best method for automated text analysis.
4. Validate, validate, validate.[30]

All quantitative models of language are wrong—but some are useful

As noted above, natural language is notoriously difficult to parse algorithmically—no model will perfectly capture the nuanced meanings of a text, let alone its topics, parts

---

[29] Justin Grimmer and Brandon M. Stewart, "Text as Data: The Promise and Pitfalls of Automatic Content Analysis for Political Texts," *Political Analysis* 21, no. 3 (Summer 2013): 270.

[30] Ibid.

of speech, or overall sentiment. In quantitative prediction and causal inference, a critical first step in analysis is to model the data-generating process for the data at hand. Researchers can include additional covariates, complex or nested functional forms, or nonparametric fitting to precisely model underlying processes for many types of data. Such an approach is impossible with language—"the complexity of language implies that all methods *necessarily* fail to provide an accurate account of the data-generating process used to produce texts."[31]

However, automated quantitative language models can still be useful for specific analytical tasks. We use two types of computational text models to uncover important insight related to the Egyptian media's use of NGOs as sources: (1) part-of-speech tagging for source identification and (2) topic modeling for frame identification. These models, described below, should be evaluated not based on whether they perfectly describe the corpora, but whether they can successfully provide insights into the media-civil society relationship.

The quantitative results from these models are not particularly sophisticated—often returning basic frequencies or counts—and accordingly, we do not build methodologically complex models. Instead, we primarily use simple tables and graphs, following the criticism of Christopher Achen: "None of the important empirical generalizations in the discipline has emerged from high-powered methodological research. Instead, almost without exception, they were found with graphs and cross-tabulations."[32] While we do not necessarily agree with Achen's polemic against complicated statistical methods, we believe that our more simple analysis—rooted in imperfect language models—does provide important empirical insight into our main research questions and hypotheses.

Part-of-speech tagging   To identify when NGOs are used as sources in news articles, we use part-of-speech tagging—one of the simplest forms of automated text analysis—to parse text and classify it into different lexical categories such as nouns, verbs, and adjectives. We identified sentence boundaries with a tokenizer algorithm that had been pre-trained for English text[33] and extracted every sentence from the three corpora that mentioned one of the 40 advocacy NGOs. We then used a separate algorithm[34] to assign a lexical category to each word and extracted a list of all the verbs. In theory, sentences with verbs that mark attribution, such as "said" or "reported" os-

---

[31] Ibid.

[32] Christopher H. Achen, "Toward a New Political Methodology: Microfoundations and Art," *Annual Review of Political Science* 5 (2002): 442.

[33] Specifically we used `nltk.tokenize.punkt()` in the NLTK Python library: http://www.nltk.org/_modules/nltk/tokenize/punkt.html.

[34] Specifically we used `nltk.pos_tag()` in NLTK: http://www.nltk.org/api/nltk.tag.html, which is pre-trained on the Penn Treebank project.

tensibly indicate that the NGO was used as a source. Part-of-speech tagging is useful for showing that NGOs are commonly used as sources, but the utility of the method is limited without more sophisticated analysis.

Topic modeling with LDA   We then use topic modeling to identify the underlying topics or frames that each news article covers. Topic modeling is more technically complicated than categorizing data based on a list of verbs and is a useful application of distant reading, or using visual and quantitative tools to generate abstract "graphs, maps, and trees"[35] to analyze—or "read"—large corpora of text.[36] These models allow scholars to "step back from individual documents and look at larger patterns among [the entirety of an archive]."[37]
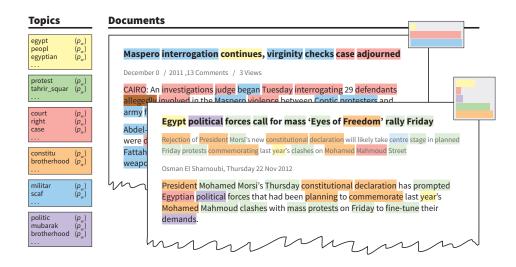

Figure 1: Diagram of LDA process

One topic modeling algorithm—latent Dirichlet allocation (LDA)—uses a multilevel generative Bayesian probabilistic process to model the mixture of unknown latent topics in a corpus of documents.[38] Simply put, LDA assumes that every document in the corpus is composed of a random distribution of corpus-wide topics and

---

[35] Franco Moretti, *Graphs, Maps, Trees: Abstract Models for Literary Theory* (London: Verso, 2005).

[36] Matthew G. Kirschenbaum, "The Remaking of Reading: Data Mining and the Digital Humanities," October 2007, http://www.csee.umbc.edu/~hillol/NGDM07/abstracts/talks/MKirschenbaum.pdf.

[37] Robert K. Nelson, *Mining the* Dispatch, Interactive topic modeling project (University of Richmond Digital Scholarship Lab, 2011), http://dsl.richmond.edu/dispatch/.

[38] David M. Blei, Andrew Y. Ng, and Michael I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research* 3 (2003): 993–1022.

that each topic contains words that are more probable that others. Figure 1, adapted from Blei,[39] demonstrates the intuition behind the LDA process. In the figure, two documents are shown with different latent topic distributions (visible in the miniature histograms), indicating that each document should be proportionally composed of words belonging to those topics. The right of the figure shows excerpts from the different topic distributions, where each topic contains each word in the corpus, but occurring with varying probabilities. The documents themselves represent random draws of words from these topic distributions according to the within-document distribution of topic probabilities. By changing and refining the parameters of the model over hundreds of iterations, we can approximate the probabilities of the words in each topics, as well as the composition of topics in each document.

LDA has become a popular method of distant reading because of its simplicity and intuitiveness. In contrast with other textual analysis methods that rely on natural language and nuanced meaning, LDA explicitly does not account for syntax, grammar, word order, or context. Because it converts words to tokens, LDA is language agnostic and can be applied to any unstructured text-based corpus. Political communications scholars have used it to determine underlying political agendas in US Senate press releases[40] and to measure how anti-Muslim fringe organizations dominate US media discourse,[41] among other applications.

Quantitative methods for text amplify resources and augment humans

Given that quantitative models are imperfect, it is inadvisable to rely solely computational methods for analysis—human inputs, such as close reading of the text for nuance and detailed domain knowledge, are still essential elements of text analysis. Both of our linguistic modeling approaches combine quantitative automation with qualitative nuance.

Manual classification of source types   As can be seen in Table 1, many of the most common verbs indicate that NGOs were possibly used as sources, including "said / says / saying," "according," "filed," and "published." These initial findings augment what human analysts could find, showing that NGOs are indeed often used as sources, but unless we make the naive assumption that every use of the word "said" indicates sourcing, the computational results require human input. Other more complicated

---

[39]David M. Blei, "Probabilistic Topic Models," *Communications of the ACM* 55, no. 4 (April 2012): 77–84.

[40]Justin Grimmer, "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases," *Political Analysis* 18, no. 1 (Winter 2010): 1–35.

[41]Christopher A. Bail, "The Fringe Effect: Civil Society Organizations and the Evolution of Media Discourse About Islam Since the September 11th Attacks," *American Sociological Review* 77, no. 6 (December 2012): 855–79.

computational methods could potentially parse the grammar of the extracted sentences, determining if the NGO is the object of the source-indicative verb. However, automated classification methods have their own model assumptions and can misclassify syntactical structure in more complex sentences.

Verb	Count	Verb	Count	Verb	Count
said	196	be	53	being	23
is	165	says	44	issued	23
was	95	according	43	released	23
are	75	filed	42	published	21
were	67	been	40	held	20
have	60	had	32	signed	19
has	58	including	24	saying	18

Table 1: Most common verbs in sentences that mention NGOs

Rather than develop a comprehensive grammar algorithm, we classified each NGO mention by hand, determining if the sentence used an NGO in one of the following ways:

- Direct quote: an NGO is cited with quotation marks[42]
- Paraphrase: an NGO's statement or quote is summarized with no quotation marks[43]
- Statement: the publication cites a portion of an NGO's report, official statement, or press release, with no indication that they talked to an individual[44]

We also determined whether the article cited an individual within the organization or if it attributed the quotation to the organization in general. This categorization

---

[42] For example, " 'The conditions were very tough surrounding civil society,' said Magdy Abdel Hamid, head of the Egyptian Association for Community Participation Enhancement (EACPE)": Nadine Marroushi, "Election monitoring hindered by restrictions," *Egypt Independent*, May 28, 2012, http://www.egyptindependent.com/news/election-monitoring-hindered-restrictions.

[43] For example, "Fifteen people are still being treated at Qasr Al-Ainy Hospital, according to Heba Abou Saeda from the Egyptian Organization for Human Rights.": Safaa Abdoun, "Health ministry raises death toll to 12 as Cairo clashes continue," *Daily News Egypt*, December 19, 2011, http://www.dailynewsegypt.com/2011/12/19/health-ministry-raises-death-toll-to-12-as-cairo-clashes-continue/.

[44] For example, "The Arab Network for Human Rights Information (ANHRI) released a statement on Monday evening saying it was 'deeply alarmed' by the decision to sentence a man to 80 lashes for being drunk.": Luiz Sanchez, "ANHRI calls for amendment of constitution," *Daily News Egypt*, April 23, 2013, http://www.dailynewsegypt.com/2013/04/23/anhri-calls-for-amendment-of-constitution/.

represents the operationalization of our hypothesis. If we find (1) that the quantity of NGO sourcing is small, and (2) that a publication tends toward statements rather than direct quotes (or even paraphrases), then the sourcing can be classified as superficial and shallow.

This manual categorization allowed us to pick up on nuances that computers would have been unable to capture, such as the distinction between a statement and a direct quote—both of which would appear as a direct quote to an algorithm. Automated analysis simplifies source categorization, but does not remove the need for human validation.

LDA output and domain knowledge Though topic models return far more data than simple part-of-speech tagging, it still requires human input and subjective decisions. After cleaning and preparing our corpora,[45] we generated a topic model with 20 topics, summarized in Table 2.[46] The first column shows the Dirichlet α parameter for each topic and is approximately equal to the proportion that topic appears in the corpus. The second column lists the top ten most likely words for each topic and is the primary output of the LDA process—these words represent the probabilistic clusters of tokenized words in the corpus, categorized by the underlying latent topics.

	Dirichlet α	Top ten words	Short name
1	0.927	egypt revolut peopl egyptian year work time ad countri day	Post-revolutionary Egypt (catch-all)
2	0.394	law state articl govern inform public right author draft legisl	Legislation and governance
3	0.334	case court investig charg trial prosecut mubarak law verdict defend	Trials
4	0.252	polic tortur report offic kill protest death violenc investig interior_ministri	Police torture
5	0.225	protest clash street march tahrir_squar demonstr squar report day cairo	Protests and clashes
6	0.201	report media human_right journalist state anhri newspap minist arab_network statement	Media and censorship

[45] Following the pattern laid out by Soriano, Au, and Banks (Jacopo Soriano, Timothy Au, and David Banks, "Text Mining in Computational Advertising," *Statistical Analysis and Data Mining* 6, no. 4 (August 2013): 273–85), we (1) filtered out common stop words, (2) reduced each word to its grammatical stem (e.g. "continue", "continuing", and "continued" became "continu"), and (3) created tokens of commonly appearing bigrams, (e.g. "human right," "tahrir squar," "arm forc," and "civil societi").

[46] We used an open source implementation of LDA called MALLET (Andrew Kachites McCallum, "MALLET: A Machine Learning for Language Toolkit," 2002, http://mallet.cs.umass.edu; Shawn Graham and Ian Milligan, "Review of MALLET," *Journal of Digital Humanities* 2, no. 1 (Winter 2012)) to generate a model based on 1,000 sampling iterations.

	Dirichlet α	Top ten words	Short name
7	0.193	polit mubarak regim militari islamist egypt state power brotherhood elect	Regime politics
8	0.192	morsi constitut brotherhood presid parti member muslim_brotherhood declar constitu_assembl referendum	Muslim Brotherhood and constitution
9	0.147	ngos human_right organ civil_societi organis fund group egypt statement govern	Human rights and civil society
10	0.146	arrest children detain releas polic_station lawyer detaine moham detent prison	Police arrests
11	0.11	militari civilian scaf human_right militari_trial militari_court protest armi releas trial	Military trials
12	0.109	govern egypt price countri tax invest econom bank percent economi	Public economics
13	0.0819	worker union compani strike labor doctor syndic work privat state	Egyptian workers
14	0.0742	women woman women_right attack men femal sexual_harass group girl sexual_assault	Sexual violence
15	0.0639	vote elect poll_station voter candid campaign report judg violat shafiq	Elections
16	0.0622	egyptian morsi presid saudi_arabia saudi prison releas arrest insult_presid year	Morsi
17	0.0587	church christian copt villag muslim famili coptic home attack sectarian	Sectarian issues
18	0.0485	azhar religion islam religi baha saber channel constitut institut imam	Religious issues
19	0.0452	countri sinai univers egypt drug arab bahrain region border asset	Protestors and activism?
20	0.0441	resid land water area environment develop live plant farmer agricultur	Environmental issues

Table 2: Topic model summary

While the topic model does accurately output clusters of most likely words, human input is required to make sense of the latent topics—domain knowledge about the corpus is necessary to understand what those lists of words actually mean. In the third column, we used our understanding of post-revolutionary Egyptian politics and events to create labels based on the words generated by the topic model. The model's largest topic, labeled "post-revolutionary Egypt," appears in every document in the corpus because it absorbed generic terms related to news coverage of Egypt. Because

it is a catch-all category, we have omitted it from most of our analysis. Each of the other topics in the model represent key issues that advocacy NGOs were involved with between 2011–13, including issues like draft legislation, protests in Tahrir Square, the Muslim Brotherhood's actions in political power, extralegal military trials of civilian activists, human rights abuses by the military and police, and other related topics. Again, quantitative text models augment human ability, but do not replace it.

There is no globally best method for automated text analysis

A host of computational text methods have been developed in recent years, but each has a different limited scope and purpose. There is a danger that certain topics can become overused—counting the frequency and distribution of verbs in a corpus is quite easy, and LDA results are often described as "magic," which has increased these methods' popularity, for better or worse. Scholars in digital humanities, history, and political science repeatedly warn against applying these methods inappropriately, simply because they are easy or convenient.[47]

As discussed previously, there is no universal algorithm that can capture the full data-generating process behind a text—instead, methods must focus on slices of the corpora to be studied. Because our research design is explicitly focused on the topics that are reported when NGOs are used as sources, our use of parts-of-speech tagging and LDA are warranted and appropriate—we do not use them because they are easy or "magic."

Validate, validate, validate

Given the complications of natural language, any method—regardless of its ease or intuitiveness—will inevitably produce errors. Accordingly, we validate the output of our automated methods in multiple ways and find that the models perform satisfactorily.

Source type validation   Validating whether or not parts of speech were correctly categorized or whether the extracted verbs indicate that an NGO was used as a source is largely a nonissue. The output of the parsing models was used to inform the manual categorization of source types and was not used directly. When we came across errors during our manual coding (e.g. words like "Mohammed," "victims," and "freedom," as well as single apostrophes were often considered verbs), we were able to adjust our manual coding accordingly.

While we can circumvent issues of computational categorization, our manual classification still faces questions of human inconsistency in interpretation. We plan

---

[47] See Ben Schmidt, "When You Have a MALLET, Everything Looks Like a Nail," *Sapping Attention*, November 2, 2012; Grimmer and Stewart, "Text as Data."

on using multiple coders to classify the type of source based on the same natural language model output.

Topic model validation   We validate the output of our topic model in two ways. First, we verify that the discovered topics correlate with the stories reported in the subset of news articles. Figure 2 shows the topic composition of a randomly selected article from each publication (summarized in Table 3). In general, the model appears to perform well. Approximately half of the al-Ahram article is categorized under "human rights and civil society," which is unsurprising given that the article actually discusses how coalitions of NGOs protested against police raids of their offices. The most common topics in the Daily News Egypt article deal with media and censorship and police torture, which is to be expected since the Arab Network for Human Rights Information (ANHRI) deals with media freedom, often specifically tackling the abuse of and violence against journalists. Finally, the model performs moderately well on the Egypt Independent article. Though it does not draw out any overwhelmingly prominent topic, the article does deal with the internal politics behind business legislation that potentially grants regime-connected business owners the ability to arrest striking workers—all topics that the model correctly identifies.

Article ID	Publication	Article title
ahram_20090	Al-Ahram English	Human rights organisations condemn mass raids of NGOs in Egypt
dne_17390	Daily News Egypt	EU, Britain and ANHRI condemn escalation of violence
egypt_independent_5000	Egypt Independent	FJP mulls draft law granting private firms arrest powers

Table 3: Summary of randomly selected articles

Second, we verify that the topics in the articles that mention specific NGOs are related to those NGOs' missions. Since most of the advocacy NGOs we have selected deal with multiple policy and issue areas, we look at the distributions for two more niche topics that are likely to attract coverage of more specialized NGOs: (1) media and censorship and (2) sexual violence. Figure 3 shows the average proportion of these topics in articles that reference one of the top eight most mentioned NGOs. As expected, ANHRI is mentioned the most often in articles related to media and censorship (approximately 20%), while Nazra for Feminist Studies and The New Woman Foundation—the two most feminist-focused NGOs in our sample—are mentioned in the clear majority of articles related to sexual violence.
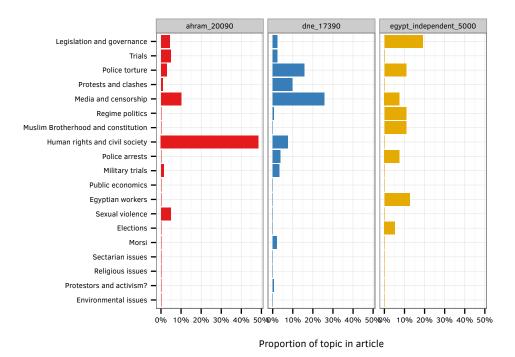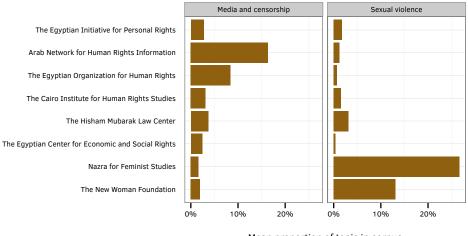
Figure 2: Topic composition of three random articles

Figure 3: Distribution of niche topics in articles that mention most common NGOs

Thus, despite the limitations of computational topic modeling, we feel that our LDA topic model satisfactorily identifies latent topics related to the advocacy NGOs that we have selected.

Results and discussion

Because we feel that we have built valid language models and appropriately blended computational and human methods, we can uncover key descriptive insights regarding the relationship between advocacy NGOs and the media. Specifically, we can use these models to describe (1) how NGOs fit in the scheme of news sources, and (2) how NGOs are used when the media turn to them as sources.

NGO coverage and sourcing

The choice to use an NGO as a source when discussing the organization follows a consistent pattern across the three publications. As seen in Table 4 78% of all NGO mentions use the organization as a source (i.e. rather than just mentioning the NGO in a list, the reporter quotes or paraphrases a member of the organization). However, there are clear differences in how often mentioning or sourcing occurs—our subset of 40 NGOs was used as a source far more often in independent publications than in state-owned media. Daily News Egypt and Egypt Independent used these NGOs as sources in around 2% of their articles on average, while al-Ahram English rarely consulted any NGO, with only 0.3% of their much more voluminous output using these organizations as sources.

Publication	Total articles	Articles with NGO mention	NGO used as source	NGO only mentioned	Source to mention ratio	Source to total ratio
Al-Ahram English	36,795	107	84	23	78.50%	0.23%
Daily News Egypt	10,327	276	215	61	77.90%	2.08%
Egypt Independent	8,822	235	183	52	77.87%	2.07%

Table 4: NGO mentions and sourced mentions by publication

This trend holds over time as well (see Figure 4). Both Daily News Egypt and Egypt Independent consistently use these organizations as sources in 2–3% of their articles each month (with the exception of Daily News Egypt in mid-2012, which stopped

publication for several weeks as it changed ownership), while al-Ahram never surpasses 0.5% in any month.
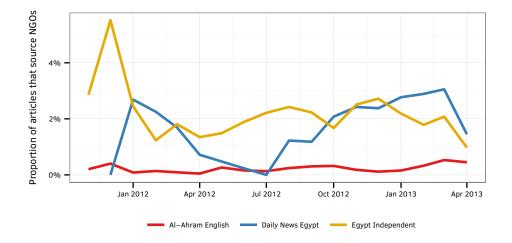

Figure 4: Proportion of articles that use NGOs as sources, November 2011–April 2013

The disparity in how often state-run and independent media outlets consult NGOs is indicative of the values these publications place on NGOs as sources. As discussed previously, previous empirical and theoretical work holds that reporters will turn to elite, government sources more often than civil society sources, thus placing NGOs near the bottom of the journalistic "pyramid of access." Both state-run and independent media use NGOs as sources relatively rarely, but it appears that the independently operated publications have closer connections to advocacy NGO and are more likely to turn to them for input into their coverage.

There are also notable and statistically significant differences in how the three publications use NGOs as sources (see Table 5 and Figure 5).[48] State-run media rarely quote NGOs directly—the two independent publications account for nearly 90% of all the direct NGO quotes in the corpus. Though Egypt Independent and Daily News Egypt account for the majority of NGO sources, the two publications differ in how they quote these organizations, likely due to differences in editorial style. Daily News Egypt is far more likely to cite a statement or a report issued by an NGO, accounting for more than half of all the statement sources in the corpus. Egypt Independent, on the other hand, comprises the majority of the paraphrases. Despite these differences in

---

[48]The cells of the contingency table are significantly different from the expected distribution ($\chi^2 = 27$, df = 6, p < 0.001), and are conditionally independent from each other (p = 0.004)

citation styles, independent media are far more likely to give direct voice to these advocacy NGOs, either through verbatim quotes or by disseminating information from NGO reports.

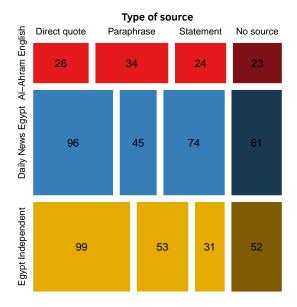| Publication	Direct quote	Paraphrase	Statement
Al-Ahram English	11.76%	25.76%	18.6%
Daily News Egypt	43.44%	34.09%	57.36%
Egypt Independent	44.8%	40.15%	24.03%
Sum	100%	100%	100%

Table 5: Percent of source type per publication

Figure 5: Source type by publication

There are also publication-level differences in whether the reporter references an individual within the NGO (i.e. the executive director, spokesperson, or lawyer) when citing the organization. As shown in Figure 6, individuals are far more likely to be cited when an organization uses a direct quote, while organizations are cited in reference to statements and reports—a logical outcome, given that reports are generally issued by an organization as a whole, while spokespeople are individuals that speak on behalf of the NGO. Interestingly, however, when al-Ahram paraphrases NGO statements, it is less likely to name their source directly, furthering the evidence that state-run media avoids using NGOs as direct sources (if they chose to cite them at all).

Finally, it is important to notes that reporters in these publications do not turn to all advocacy NGOs equally. A handful of advocacy organizations dominate discus-
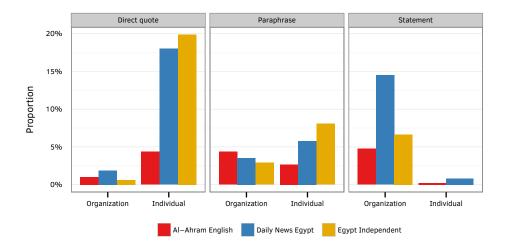
Figure 6: Source type by publication, grouped by individual mention

sion in the media and account for the majority of sourced mentions (see Table 6). Out of the 40 NGOs that signed EIPR's open letter, 18 of them were never used as sources by any publication and 13 of them were cited 5 times or fewer. Nine organizations account for more than 90% of all NGO citations, with EIPR alone representing 40%. This skewed distribution can be explained by several hypotheses. It could be that because EIPR penned the original letter, it sought out pseudo-affiliate organizations to sign it—each of the other organizations could simply be smaller nodes in an advocacy network dominated by EIPR. While this is possible, many of the organizations that did sign are not formally connected to EIPR, such as the Hisham Mubarak Law Center or ANHRI, each of which tend to pursue their own advocacy agendas. If we assume that these organizations are relatively independent from each other, the disproportionate access could be explained by organizational differences in approaches to the media. Some organizations might have better relations with the press, have more compelling missions, or have strategies that are specifically designed to take advantage of the media, thus allowing them to dominate in the market of NGO attention. Further research is needed regarding the structure of the Egyptian advocacy NGO network, the internal strategies and structures of individual NGOs, and their direct relationships to media outlets.

Organization	Al-Ahram English	Daily News Egypt	Egypt Independent	Total	Percent
The Egyptian Initiative for Personal Rights	38	84	71	193	40.04%
Arab Network for Human Rights Information	15	44	6	65	13.49%
The Egyptian Organization for Human Rights	5	17	34	56	11.62%
The Cairo Institute for Human Rights Studies	9	13	16	38	7.88%
The Egyptian Center for Economic and Social Rights	5	8	18	31	6.43%
The Hisham Mubarak Law Center	1	11	17	29	6.02%
Nazra for Feminist Studies	4	11	4	19	3.94%
The New Woman Foundation	2	4	3	9	1.87%
The Egyptian Association for Community Participation Enhancement	1	2	5	8	1.66%
Arab Organisation for Penal Reform	1	4	0	5	1.04%
Arab Program for Human Rights Activists	0	5	0	5	1.04%
The Land Center for Human Rights	0	0	5	5	1.04%
Baheya Ya Masr	0	3	0	3	0.62%
Center for Egyptian Womens Legal Assistance	0	2	1	3	0.62%
Habi Center for Environmental Rights	0	0	3	3	0.62%
Women and Memory Forum	1	2	0	3	0.62%
Andalus Institute for Tolerance and Anti-Violence Studies	0	2	0	2	0.41%
Appropriate Communications Techniques for Development	0	1	0	1	0.21%
Arab Foundation for Civil Society and Human Right Support	1	0	0	1	0.21%
The Egyptian Coalition for the Rights of the Child	0	1	0	1	0.21%

Organization	Al-Ahram English	Daily News Egypt	Egypt Independent	Total	Percent
The Egyptian Foundation for the Advancement of Childhood Conditions	1	0	0	1	0.21%
The International Center for Supporting Rights and Freedoms	0	1	0	1	0.21%
Total	84	215	183	482	

Table 6: Frequency of use as source in each publication

How are NGOs talked about when used as sources?

Our computational models also allow us to describe in which contexts NGOs are sourced, determining which topics or journalistic frames they are used to support. As discussed previously, turning to NGOs when covering specific topics can allow a publication to either create a sense of balance and credibility or push its own advocacy or editorial agenda. While we cannot determine if the reference to the NGO balances the story without looking at all other sources used in the article, we can comment on how the NGO is used as source and determine in which which frames the sourcing is more likely to occur.

First, we can compare the predominant topics in each of the three publications to see how these organizations differ in when using NGOs as sources. Figure 7 shows that each of the publications tend to give roughly the same amount of coverage to each topic, with several notable exceptions. Al-Ahram uses NGO sources more than either independent publication when reporting on police torture, which is surprising considering its state-sponsored editorial position. However, many of its articles regarding police torture are reprints from wire services—stories that are also often published by the independent press—and thus does not necessarily reflect an editorial stance against the practice. Al-Ahram also turns to NGOs more often than the independent press when discussing elections, President Morsi, and regime politics. These differences follow conventional wisdom—since the state tends to have greater interest in elite politics, its news apparatus would be expected to cover those topics more frequently.

Second, as expected, independent publications use NGOs as sources when reporting on issues that are more contentious and adversarial to the state, such as Mubarak's trial, military trials, media and censorship, human rights abuses, sectarian religious conflicts, and militant union action and worker relations. The model also reveals latent differences in which topics the two publications cover more. Daily News Egypt appears to be more responsive to NGOs concerned with human rights, censorship,
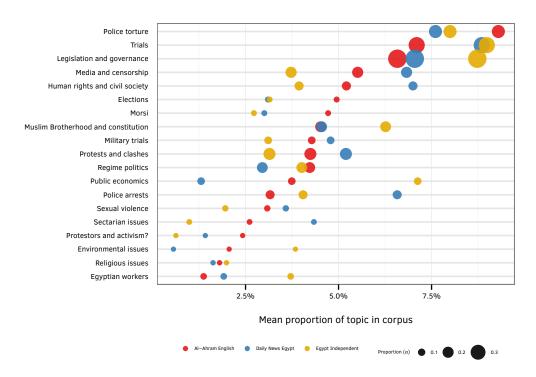
Figure 7: Average proportion of topics by publication, sorted by al-Ahram's highest to lowest proportion in the corpus, catch-all category suppressed

arrests, and sectarianism, while Egypt Independent uses NGOs more for coverage of regime politics, economic policy, labor policy, and environmental issues. Daily News Egypt gives more of a platform to anti-regime, pro-human rights advocacy organizations than Egypt Independent. We can see the same trend when looking at individual organizations. Figure 8 shows the distribution of topics by publication for the top four advocacy organizations. Egypt Independent turns to these organizations most often when covering higher level, elite politics such as legislation and trials, while Daily News Egypt (and even al-Ahram) turn to them for coverage on human rights issues.

Finally, we can see how these publications stylistically use NGOs as sources (see Figure 9). Similar to the division between Egypt Independent and Daily News Egypt, journalists generally tend to use NGO statements when discussing more elite topics, such as legislation, torture, or censorship, often citing a report to provide a counterpoint to elite sources and spokespeople. On the other hand, NGOs are paraphrased or directly quoted in more niche topics like sectarianism, the environment, or activist protests. This split does not imply that NGOs are never cited or quoted directly when the media covers elite topics—rather, it indicates that on average, articles written about
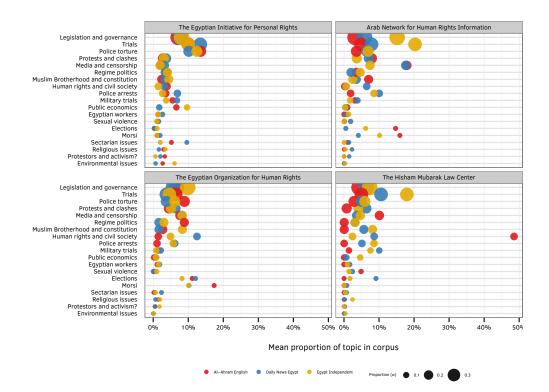
Figure 8: Average proportion of topics by publication for the top four NGOs, catch-all category suppressed

these elite topics will reference a report or statement (at times in addition to a direct quote).

Conclusion

Since the data outputs for this project are fairly recent and we are just beginning to dig through them, these preliminary observations contribute to the ongoing research in the area and also provide some space to analyze our hypothesis and propose further research questions.

Our principal contribution is that the expectation from our third hypothesis—that publications will turn to civil society sources (such as NGOs) rarely, with superficial context, and with shallow nuance—is only partially met. We found that NGOs were used as sources sparingly in comparison to overall journalistic output, but that the reporting that referenced these NGOs was not necessarily superficial or shallow. There is evidence that the state-run media tends to be more superficial in its use of
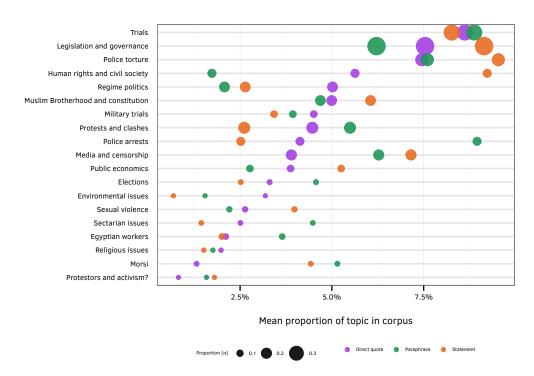
Figure 9: Average proportion of topics by source type, sorted by the highest to lowest proportions of direct quotes, catch-all category suppressed

NGOs as sources while independent media were more likely to identify specific individuals and use direct quotes and paraphrases. This distinction is important because it demonstrates different levels of effort on the part of the journalists in these news organizations. Statements, as we define them, are generally publicly accessible documents, received either through press conferences or on the internet. Directly quoting or paraphrasing NGO spokespeople implies that the journalist had to reach out to individuals, ask questions, and synthesize their responses.

One of the more interesting implications of our research is that some NGOs seem to have better relations with media organizations than others. For example, why did all three publications use EIPR as a source so often while ignoring other NGOs, many with missions that are arguably more specifically focused on specific issue areas? Further research could answer this by analyzing the organizational structures of these advocacy NGOs and discovering whether (1) they rely on a dedicated media relations staff member, (2) they follow a clearly articulated media strategy, and (3) whether they have a personal or professional relations with these news outlets.

Finally, we believe that our innovative blend of human and computational methods can contribute to a more robust and reproducible approach to content analysis of large corpora of text—allowing us to further the elusive goal of understanding both text and context.

References

Achen, Christopher H. "Toward a New Political Methodology: Microfoundations and Art." *Annual Review of Political Science* 5 (2002): 423–50.

Bail, Christopher A. "The Fringe Effect: Civil Society Organizations and the Evolution of Media Discourse About Islam Since the September 11th Attacks." *American Sociological Review* 77, no. 6 (December 2012): 855–79.

Bennett, Lance. "Toward a Theory of Press/State Relations." *Journal of Communication* 40, no. 2 (1990): 103–25.

Billington, James. *Fire in the Minds of Men: Origins of the Revolutionary Faith.* New York: Basic Books, 1980.

Blei, David M. "Probabilistic Topic Models." *Communications of the ACM* 55, no. 4 (April 2012): 77–84.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (2003): 993–1022.

Bob, Clifford. *The Marketing of Rebellion: Insurgents, Media, and International Activism.* Cambridge, MA: Cambridge University Press, 2005.

Deutsch, Karl. *Nationalism and Social Communication.* Cambridge, Massachusetts: The MIT Press, 1958.

Downie, Leonard, and Robert Kaiser. *The News About the News.* New York: Vintage Press, 2003.

Egyptian Initiative for Personal Rights. "The Muslim Brotherhood Lays the Foundations for a New Police State by Exceeding the Mubarak Regime's Mechanisms to Suppress Civil Society," May 30, 2013. http://www.eipr.org/en/pressrelease/2013/05/30/1720.

Eribo, Festus, and William Jong-Ebot. *Press Freedom and Communication in Africa.* Trenton, New Jersey: Africa World Press, 1997.

Gamson, William A., and Gadi Wolfsfeld. "Movements and Media as Interacting Systems." *Annals of the American Academy of Political and Social Science* 528 (July 1993): 114–25.

Gance, Stephen, An De Vaney, and Yan Ma. *Technology and Resistance.* New York: Peter Lang Publishing, 2000.

Ghannam, Jeffrey. "Freedom, Beyond 140 Characters." *The Washington Post* (February 20, 2011): B3.

Graham, Shawn, and Ian Milligan. "Review of MALLET." *Journal of Digital Humanities* 2, no. 1 (Winter 2012).

Grimmer, Justin. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18, no. 1 (Winter 2010): 1–35.

Grimmer, Justin, and Brandon M. Stewart. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis for Political Texts." *Political Analysis* 21, no. 3 (Summer 2013): 267–97.

Hall, Stuart, Chas Critcher, Tony Jefferson, John Clarke, and Brian Roberts. *Policing the Crisis: Mugging, the State and Law and Order.* London: Palgrave Macmillan, 1978.

Jamal, Amaney. *Barriers to Democracy: The Other Side of Social Capital in Palestine and the Arab World.* Princeton: Princeton University Press, 2007.

Kandil, Amani, Stefan Toepler, and Lester M. Salamon. "Egypt." In *Global Civil Society: Dimensions of the Nonprofit Sector*, edited by Lester M. Salamon and S. Wojciech Sokolowski, 2:217–26. Bloomfield, CT: Kumarian Press, 2004.

Kirschenbaum, Matthew G. "The Remaking of Reading: Data Mining and the Digital Humanities," October 2007. http://www.csee.umbc.edu/~hillol/NGDM07/abstracts/talks/MKirschenbaum.pdf.

Lacy, Stephen, Steven S. Wildman, Frederick Fico, Daniel Bergan, Thomas Baldwin, and Paul Zube. "How Radio News Uses Sources to Cover Local Government News and Factors Affecting Source Use." *Journalism & Mass Communication Quarterly* 90, no. 3 (September 2013): 457–77.

Manning, Paul. *News and News Sources: A Critical Introduction.* London: SAGE Publications, 2001.

McCallum, Andrew Kachites. "MALLET: A Machine Learning for Language Toolkit," 2002. http://mallet.cs.umass.edu.

Moretti, Franco. *Graphs, Maps, Trees: Abstract Models for Literary Theory.* London: Verso, 2005.

Nacos, Brigitte L. *Mass-Mediated Terrorism: The Central Role of the Media in Terrorism and Counterterrorism.* Lanham, Maryland: Rowman & Littlefield, 2007.

Nelson, Robert K. *Mining the* Dispatch. Interactive topic modeling project. University of Richmond Digital Scholarship Lab, 2011. http://dsl.richmond.edu/dispatch/.

Robertson, Alexa. "Connecting 'Old' and 'New' Media and the Arab Spring." *The International Journal of Press/Politics* 18, no. 3 (2013): 325–341.

Rugh, William A. *The Arab Press.* Syracuse, New York: Syracuse University Press, 1987.

Schmidt, Ben. "When You Have a MALLET, Everything Looks Like a Nail." *Sapping Attention*, November 2, 2012.

Schuler, Douglas, and Peter Day, eds. *Shaping the Network Society*. Cambridge, Massachusetts: The MIT Press, 2004.

Soengas, Xose. "The Rold of the Internet and Social Networks in the Arab Uprisings - an Alternative to Official Press Censorship." *Comunicar* 21, no. 4 (2013): 147–155.

Soriano, Jacopo, Timothy Au, and David Banks. "Text Mining in Computational Advertising." *Statistical Analysis and Data Mining* 6, no. 4 (August 2013): 273–85.

Sreberny, Annabelle, and Gholam Khiabany. *Blogistan*. London: I.B. Tauris & Co. Ltd., 2010.

Tiffin, Rodney, Paul K. Jones, David Rowe, Toril Aalberg, Sharon Coen, James Curran, Kaori Hayashi, et al. "Sources in the News: A Comparative Study." *Journalism Studies* 15, no. 4 (2014): 374–91.

Yerkes, Sarah. "Capture, Co-Optation, and Pluralism: Navigating the Civil Society Arena in the Arab World." PhD thesis, Georgetown University, 2012.