

# LINKING AUDIENCES TO NEWS II

2012 NETWORK ANALYSIS OF CHICAGO WEBSITES

An **Advancing Chicago's**  
**Information Ecosystem** Report

Funded by The Chicago Community Trust and  
the John D. and Catherine T. MacArthur Foundation

# Table of Contents

Executive Summary .....	1
Background and Context .....	3
How the Research Was Conducted .....	5
Introduction to Network Analysis.....	10
Structure of the Chicago News Ecosystem.....	11
Network Roles .....	14
Web Analytics Research .....	24
Exploring the Ecosystem With the Interactive Dashboard .....	28
Looking to the Future.....	30
About the Authors.....	32

## Executive Summary

The research project described in this report is the second examination of Chicago's "new news ecosystem" undertaken by the authors on behalf of The Chicago Community Trust, using sophisticated network analytics. The first study, based on data gathered in August 2010, provided insights into the most important news and information websites in Chicago, based on an analysis of the hyperlinks connecting these sites. It may have been the first network analysis focusing specifically on links among news sites in a local U.S. market.

This new study, with data from April 2012, updates and extends the findings of the first study and adds a novel analysis of web-analytics data for almost 100 websites in the region.

The key findings from our links research are:

- 301 websites form what we consider to be the core of the Chicago news and information ecosystem – sites that are connected to each other by one or more links. This network is 13 percent larger than the network of 268 connected sites identified in the 2010 study.
- Chicago's 2012 network of sites is more tightly linked than in 2010. Almost 3 percent of all possible site-to-site links exist – almost twice what we found in 2010.
- Twenty-six sites rose to the top of the Chicago news ecosystem, based on a variety of different measures of what we call "centrality," the extent to which they play central roles in the network of sites. The top sites are a mix of legacy and legacy-affiliated sites, micro-publishers and organizational and institutional sites.
- Half of the top 26 are enduring leaders – ones that also topped our site lists in 2010, suggesting that they are not only important in the network but have remained so over time. These sites are: [chicagoartistsresource.org](http://chicagoartistsresource.org), [chicagoist.com](http://chicagoist.com), [chicagoreader.com](http://chicagoreader.com), [chicagotribune.com](http://chicagotribune.com), [cityofchicago.org](http://cityofchicago.org), [communitymediaworkshop.org](http://communitymediaworkshop.org), [gapersblock.com](http://gapersblock.com), [metrarail.com](http://metrarail.com), [msichicago.org](http://msichicago.org), [nytimes.com](http://nytimes.com), [suntimes.com](http://suntimes.com), [transitchicago.com](http://transitchicago.com) and the now-defunct [windycitizen.com](http://windycitizen.com).
- The other half are new to our top sites list, including [chicagoargus.blogspot.com](http://chicagoargus.blogspot.com), [chicago.everyblock.com](http://chicago.everyblock.com), [chicagonow.com](http://chicagonow.com), [chicagoradioandmedia.com](http://chicagoradioandmedia.com), [examiner.com](http://examiner.com), [huffingtonpost.com](http://huffingtonpost.com), [illinois.gov](http://illinois.gov), [progressillinois.com](http://progressillinois.com), [proyectolatina.org](http://proyectolatina.org), [siskelfilmcenter.org](http://siskelfilmcenter.org), [steppenwolf.org](http://steppenwolf.org), [thesixthward.blogspot.com](http://thesixthward.blogspot.com) and [wbez.org](http://wbez.org).
- Three national sites with a local presence now play very important roles in the ecosystem because they have actively embraced linking to other sites: [patch.com](http://patch.com) (AOL's hyperlocal network), [huffingtonpost.com](http://huffingtonpost.com) (which has an active Chicago subsite) and [chicago.everyblock.com](http://chicago.everyblock.com) (which, under the ownership of MSNBC, has expanded its influence in the network by enabling and promoting conversations among residents of Chicago neighborhoods).

By obtaining and analyzing Web analytics for 98 of the 301 sites, we were able to take a closer look at how users move around the ecosystem by clicking on links from one site to another. This research, new to the current study, revealed several other insights:

- Smaller websites in the region rely more on traffic from within the ecosystem than larger sites. The share of traffic that smaller sites receive from other sites in the Chicago area is more than 11 times as great as the share of inside-the-region traffic that large sites received.
- Social media, especially Facebook, are critically important to driving traffic – especially for small sites. Facebook and Twitter drove more than half of the referred visits for small sites – almost three times the percentage of Facebook/Twitter visits referred to the largest sites.
- Sites that are most successful in driving traffic to other sites seem to have embraced and prioritized external links – doing it frequently and prominently. The best example we found in our study was [gapersblock.com](http://gapersblock.com).
- There is a demonstrable statistical relationship between the number of links a site provides and the amount of traffic it drives – and that relationship is especially strong for linking by the smallest sites.

Our report concludes with ideas for using, applying and extending this research in Chicago and elsewhere.

We'd like to add one final note about [windycitizen.com](http://windycitizen.com), which rated as one of the most influential sites in both of our studies. Its creator, Brad Flora, shut down the site in June 2012. Flora had described the site as a “local Digg,” where users could submit and vote up Chicago stories they thought were interesting. While its audience was never huge, our research showed that it had a key position in the network because of the way it could bring news to the attention of its audience. In announcing the shutdown of the site, Flora said he had not managed to build a thriving revenue base – and noted that he was devoting more and more time to his online-advertising business, NowSpots. The demise of [windycitizen.com](http://windycitizen.com) demonstrates that even sites that are important in the “new news ecosystem” are not guaranteed to become successful businesses. We hope our research can be useful as entrepreneurs, journalists and civic organizations look for ways to sustain news coverage in local markets.

We are grateful to The Chicago Community Trust for the opportunity to do this work and learn more about the local news ecosystem. Thanks to the Trust and the John D. and Catherine T. MacArthur Foundation for funding this research.

## Background and Context

The modern news environment in a metropolitan community like Chicago incorporates hundreds of small websites, not just a handful of major media companies. This diversification has the potential to serve audiences with a greater breadth and depth of news than ever before. But this “new news ecosystem” can only serve the community if the stories they cover reach the people who should know about them. On the Web, audiences discover content through distribution mechanisms based on networks: hyperlinks, search engine algorithms and social sharing via services such as “e-mail to a friend,” social networks and blogs. To understand Chicago’s “new news ecosystem,” it is desirable to study the network of sites and links that serve our community.

This is the second network study we have conducted of the Chicago “new news ecosystem” on behalf of The Chicago Community Trust. The first study<sup>1</sup>, based on data gathered in August 2010, is to our knowledge the first network analysis focusing specifically on links among news sites in a locally defined market. Its goal was to help the Trust – and others interested in improving the availability of news and information in Chicago – take action to help improve the “information health” of the ecosystem.

The first study found:

- There were more than 400 websites that provide news and information relevant to Chicagoland residents.
- Of these sites, almost eight in 10 received few if any links FROM other sites in the network. And more than four in 10 sites did not link TO other sites in the network.
- The sites that did give and receive links were clustered into content-oriented communities – for instance, sports, arts and restaurants – that linked heavily among themselves but didn’t get links from sites focusing on other content domains.
- Some sites were more important to the ecosystem than others. They linked out more to other sites, were more likely to receive links, or both.
- Organizations and institutions – which in the mass media era relied heavily on newspapers and TV to get their messages to local residents – now can get their stories out via their own websites. They were among the most widely linked-to sites in the region.
- Websites operated by traditional media tended not to link out to other sites. The sites most likely to link out were organizations and institutions, as well as online-only publications.

---

<sup>1</sup> <http://www.cct.org/research/research/linking-audiences-to-news>

The first study was enlightening, but like all such research, it was a snapshot at a single point in time. It also yielded ideas for additional research. The new study, focusing on the Chicago news/information ecosystem as of April 2012, extends the initial findings by:

- Providing a second picture of the network that confirms many of the findings of the original research;
- Focusing more closely on the sites that are linked together, rather than the ones that are isolated;
- Identifying a small number of sites that turned out to be important in the ecosystem in both studies;
- Providing the results of the links research in an online dashboard (<http://public.syndiosocial.com/ChiLinks>) that users can explore themselves;
- Assembling and analyzing data from Web analytics for 98 different websites, allowing an exploration of the relationship between links and audience usage.

This research project can help people better understand the structure of the hyperlink network of news/information websites in the region. We hope the report and dashboard will be useful to people interested in improving the state of news and information in the Chicago area.

Our research was funded by The Chicago Community Trust and the John D. and Catherine T. MacArthur Foundation. It is part of the Community News Matters program (<http://www.communitynewsmatters.org>), a funders collaborative created by The Chicago Community Trust in 2009 in response to the Knight Foundation's Community Information Challenge (<http://www.informationneeds.org>), a five-year, \$24 million initiative to inspire community foundations to find creative ways to use new media technology to keep communities informed. The goals of Community News Matters are to increase the flow of truthful, accurate and insightful news and information in the Chicago region and spur development of new business models for news.

## How the Research Was Conducted

For this project, we assembled two data sets: a network of links built by crawling Chicago-area news and information websites, and Web analytics data provided by publishers of almost 100 of these sites.

- The links study used the same general approach as in the first research project, but with some important changes in methodology. The methodological changes mean that we need to be cautious in comparing the results of the two research projects – but we will point out similarities and differences in our findings when we believe them to be relevant.
- The analytics initiative was entirely new – and, to the best of our knowledge, is the first effort to assemble Web analytics data for a significant sampling of the websites in a local news market.

### The Link Network

To identify the link structure among websites, researchers use webcrawler software, or spiders, that start with a list of “seed sites” and then follow links to other sites. We used VOSON (<http://www.uberlink.com/>). This was a different tool than the one used for the first Chicago links research. Though using a different webcrawler limits our ability to compare findings between the two studies, we chose VOSON because it gave us a great deal of control over configuring the rules for the crawl process. We analyzed the link data using analytics developed at Syndio Social.

We chose our seed sites using the following process:

- We began with the list of 438 sites generated through the original study;
- We removed any sites that had not been updated in 3 months or were isolates (not linked to or from other sites) in the first study;
- We then added in (1) the Chicago-area Patch.com sites, which didn’t exist when we did the last study; and (2) sites from the Community Media Workshop’s database that we did not already have in our list of seed sites (after removing those sites that hadn’t been updated recently).

This resulted in a list of 380 seed URLs.

VOSON is a Web-based software tool that maps and analyzes online networks. It was developed at Australian National University by Robert Ackland. The crawl worked like this:

- The list of 380 seed URLs was entered into VOSON. Each seed URL led to one page of a website, usually the homepage (Depth One).

- VOSON searched each page for links to other pages within the site. Accessing these front-page links gave a new set of pages located one level into the site (Depth Two). Then the process was repeated for pages three levels into the site (Depth Three).
- VOSON followed the external links, or hyperlinks, from these three sets of pages – the front pages, the pages one level in and the pages two levels in.
- Sites that received external links from two or more sites formed the new network to be crawled. The entire process was repeated two more times, using the sites found after each iteration as the new set of seed sites.
- We removed any isolates still remaining from the final data set.

Due to a glitch in the crawl process, seven sites<sup>2</sup> were erroneously excluded from the seed site database, and this was not discovered until it was too late to conduct the crawl again. The glitch could have resulted in the exclusion of a few sites from our network, but we don't believe it would have materially affected the general conclusions of the study.

The final network consisted of 301 connected sites, excluding isolates and nonlocal sites that provide services to publishers such as AddThis, Facebook and WordPress. We consider this to be the core of the Chicago news ecosystem. This is almost 13 percent more sites than the 268 connected sites we focused on in the previous study. We counted more than 8,000 links connecting the 301 sites.

## **The Analytics Research**

After the network data was gathered, we contacted publishers of Chicago area websites and asked them to provide data from their web analytics systems about “referring domains.” Web analytics systems, the most popular of which are Google Analytics and Omniture (now part of the Adobe Digital Marketing Suite), take advantage of the way the Web’s hypertext transfer protocol (http) works. When someone clicks on a link that takes them from website A to website B, the server software at website B captures and stores the web address of the page on website A. Web analytics systems also capture information about the linked-to page as well as other data: time of day, the type of browser and operating system and in many cases the country or even city where the user is online.

We asked publishers to provide data for a two-week period (March 26-April 8) that included the period when the webcrawler was visiting their sites. The idea was to find the time period when the most website traffic was likely to be driven by the links identified in our webcrawler work.

Publishers of 98 websites provided data, though there were inconsistencies in what they gave us. Here are the issues we encountered and how we attempted to adjust for them:

---

<sup>2</sup> The sites excluded were: art.newcity.com (added back in the crawl), beachwoodreporter.com, beatkitchen.com, chicagoistheworld.com, newssun.suntimes.com, schubas.com and so-laze.com.

- Some sites provided us with complete referral data for the two-week period; some provided the top 200 or 500 URL's. One site provided only the top 50 referrers. Where we had incomplete data, we estimated total referred traffic based on patterns to other sites in our data set.
- Some sites provided data only for inbound visits driven by links from other publishers; others included clicks from search engines (“search traffic”) and people who visited the site by using a bookmark or typing in the URL (“direct traffic”). We attempted to exclude search and direct traffic for all sites.
- Some analytics systems gave us referring URL's that included subdomains (e.g., evanston.patch.com) while others reported only primary domains (e.g., patch.com). To count referred traffic, therefore, we focused on primary domains.
- A few sites gave data for a slightly different time period than we requested. We made mathematical adjustments based on the data available.

Some publishers were reluctant to provide analytics data if it would be made available publicly, for fear it would be used by competitors or reveal details about their traffic that they did not want to share. Because understanding the structure of the overall network was more important to us than providing details about specific sites, in these cases we agreed not to release site-specific details from the analytics research without prior approval from those publishers.

When all the referrer data was assembled, we were able to analyze the traffic for 4.6 million website visits over the two-week period.

## Site Categorizations

To aid in our analysis of the network data, we coded sites by “category” and “scope.” All sites were grouped into one of the following categories:

- **Legacy:** A Web publication related to or corresponding with a mainstream or traditional media brand. The site carries the brand and/or content of this traditional media brand. (Examples: chicagotribune.com, chicagoreader.com.)
- **Legacy-affiliated:** A Web publication owned by a mainstream or traditional media brand but carrying a separate name. (Examples: chicagonow.com, vocalo.org.)
- **Micro-publisher:** A Web-first publication focused on a topic, geographic area or audience segment. (Examples: gapersblock.com, evanstonnow.com.)
- **Organization/institution:** A website providing news or information from an organization, company, institution, government agency or nonprofit. (Examples: fieldmuseum.org, cityofchicago.org.)

- **National brand:** A website with national scope that provides Chicago-area news and information, usually with a structure of geographically segmented subsites.<sup>3</sup> (Examples: [chicago.everyblock.com](http://chicago.everyblock.com), [huffingtonpost.com](http://huffingtonpost.com), [sbnation.com](http://sbnation.com).)
- **Web service:** A website that provides services to Web publishers. (Examples: [addthis.com](http://addthis.com), [wordpress.com](http://wordpress.com), [facebook.com](http://facebook.com), [flickr.com](http://flickr.com), [quantcast.com](http://quantcast.com), [twitter.com](http://twitter.com), [vimeo.com](http://vimeo.com).) Because our focus was on the local ecosystem, we excluded service sites from the final analysis.

The “scope” categorization was intended – for sites providing local news and information – to describe the type of information provided. We assigned a scope to all sites categorized as Legacy, Legacy-affiliated or Micro-publisher. We also assigned a scope to National Brand and Organization/Institution sites if they provided news or information about one or a few geographic areas within the region. Options for scope were:

- **Geo-publisher:** A Web publication focused on one or a few specific geographic areas (neighborhoods or municipalities) within the Chicago region. (Examples: [dailyherald.com](http://dailyherald.com), [evanston.patch.com](http://evanston.patch.com), [pioneerlocal.com](http://pioneerlocal.com).)
- **Niche publisher:** A Web publication focusing on a specific topic or audience segment. (Examples: [gapersblock.com](http://gapersblock.com), [timeoutchicago.com](http://timeoutchicago.com), [vocalo.org](http://vocalo.org).)
- **Mass media:** A website branded with a major mass media outlet. (Examples: [chicagotribune.com](http://chicagotribune.com), [nbc5chi.com](http://nbc5chi.com), [suntimes.com](http://suntimes.com).)

---

<sup>3</sup> In describing the findings of the links research, we will refer to National Brand sites in this report by using the URL that best characterizes the primary Internet domain for the links captured by our webcrawler. This differs depending on whether the seed site was a subdomain (e.g., [chicago.everyblock.com](http://chicago.everyblock.com) or [evanston.patch.com](http://evanston.patch.com)) or a URL containing a subdirectory (e.g., [huffingtonpost.com/chicago](http://huffingtonpost.com/chicago) or [examiner.com/chicago](http://examiner.com/chicago)). For seed sites with subdomains, we will use the full home-page URL ([chicago.everyblock.com](http://chicago.everyblock.com)). For seed sites with subdirectories, we will use the primary domain ([huffingtonpost.com](http://huffingtonpost.com) or [examiner.com](http://examiner.com)). This is consistent with the way the webcrawler captured the links in our database.

Here is a breakdown of the sites we identified:

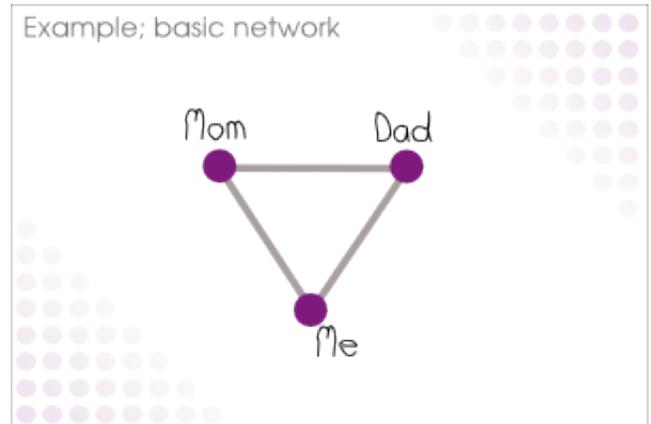
	SCOPE				
CATEGORY	Geo-publisher	Niche publisher	Mass media	None	TOTAL
Legacy	<b>34</b>	<b>56</b>	<b>14</b>		<b>104</b>
Legacy-affiliated		<b>6</b>			<b>6</b>
Micro-publisher	<b>15</b>	<b>36</b>			<b>51</b>
National brand	<b>58</b>			<b>6</b>	<b>64</b>
Organization/institution	<b>6</b>			<b>70</b>	<b>76</b>
<b>TOTAL</b>	<b>113</b>	<b>98</b>	<b>14</b>	<b>76</b>	<b>301</b>

In addition to these categorizations, for purposes of interpreting the Web analytics data, we grouped sites into three size categories based on the number of referring visits from other websites during the March 26-April 8 period:

- Large sites are those with more than 500,000 referring visits over two weeks. There are three in our database: [chicagotribune.com](http://chicagotribune.com), [suntimes.com](http://suntimes.com) and [dailyherald.com](http://dailyherald.com).
- Medium sites had between 50,000 and 500,000 referring visits. This group of 15 sites is dominated by sites affiliated with traditional media companies, such as [chicagomag.com](http://chicagomag.com) and [chicagonow.com](http://chicagonow.com) (both owned by the Tribune Co.), [chicagoreader.com](http://chicagoreader.com), [wbez.org](http://wbez.org) and the Sun-Times' community newspaper sites ([pioneerlocal.com](http://pioneerlocal.com), [beaconnews.suntimes.com](http://beaconnews.suntimes.com), [couriernews.suntimes.com](http://couriernews.suntimes.com), [heraldnews.suntimes.com](http://heraldnews.suntimes.com), [napervillesun.suntimes.com](http://napervillesun.suntimes.com) and [southtownstar.suntimes.com](http://southtownstar.suntimes.com)).
- Small sites (there are 80 in our database) had fewer than 50,000 referring visits during our timeframe. Most were categorized in our links research as "micro-publishers." These include hyperlocal sites covering a town or neighborhood as well as niche sites focusing on a targeted topic or audience.

## Introduction to Network Analysis

On a fundamental level, the world is full of “things” – such as people, organizations, communities and societies – that have relationships with one another. For example, people have friends, family and colleagues, while organizations have affiliates, competitors and stakeholders. These relationships are often complex; their structure creates a web or network. Although people have many one-on-one interactions, in most cases, people have relationships with multiple people. These people then, in turn, have relationships among themselves. For example, a person would have a relationship with her mom and dad while the two of them also have a relationship; hence the three people form a basic network, as shown in Fig. 1. In network terms, each individual in an interpersonal network is a **node**.

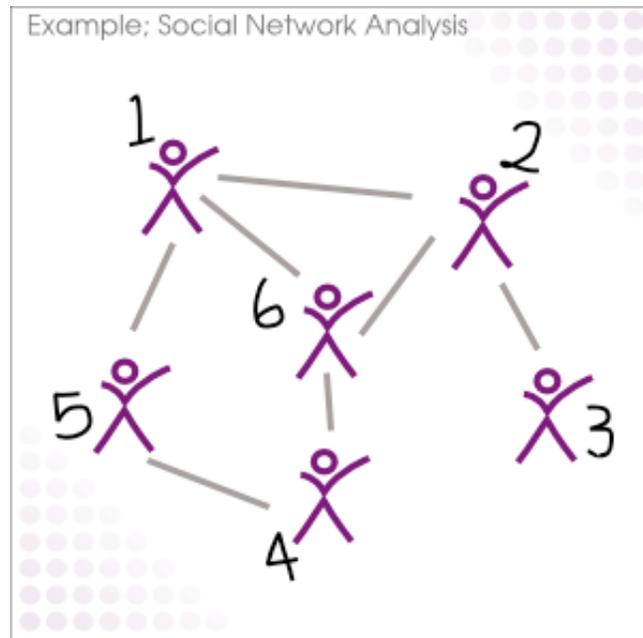


**Fig. 1: A three-node network**  
In network terms, each individual in an interpersonal network is a **node**.

### How Networks Are Analyzed

To study networks, we represent them as a set of nodes connected to one another by a set of edges. Nodes can be people, organizations, websites, or resources. Edges (or ties, links or connections) are the relationships that connect the nodes.

For example, consider Fig. 2, a basic network of six people (the nodes). Some nodes are connected, but others are not. Node 1 is connected to three other nodes directly (nodes 2, 5 & 6), but node 3 is only connected to one other node (node 2). It takes node 1 just one step to reach nodes 2, 5 and 6 and two steps to reach the remaining two nodes (3 and 4). All of the nodes are connected to one another via no more than two steps.



**Fig. 2: A basic six-node network**

From this simple setup, analysts can use a set of network metrics to quantify and compare how certain nodes and relationships contribute to the flow of information or resources through a network.

## Structure of the Chicago News Ecosystem

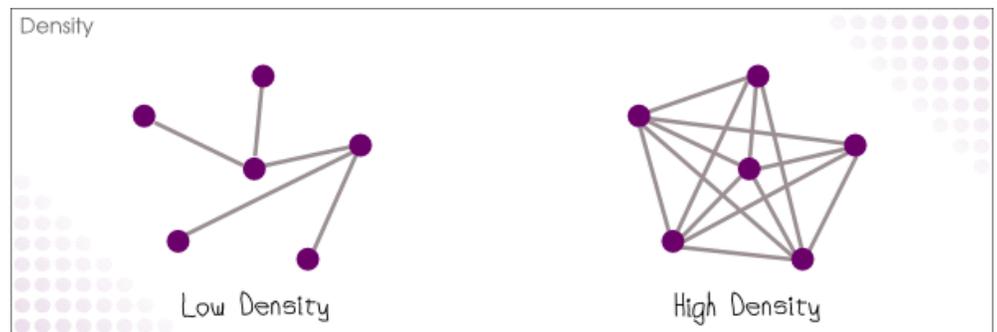
Network diagrams are visual representations of how a network is arranged structurally. We will use different types of network diagrams to show how sites in the hyperlink network are interrelated, as well as to point out key insights.

Network analysts also use a set of network metrics unique to their field; these allow us to quantify and compare how certain nodes and relationships are vital to the flow of information or resources. For example, **centrality** is described by a series of measures that quantify the extent to which a node is (or isn't) at the center of its network, while **closeness** is the degree to which a node can reach other nodes in its network. Both are useful for measuring the ease with which information passes through a network.

### Network Density

*Density is an important measurement for understanding the structure of a network. In a nutshell, it can be defined as the ratio of the number of actual links in the network to the number of potential links.*

Density ranges from 0 to 1, where a 1 indicates that all of the potential links are present. In Fig. 3, the network on the left has a relatively low density, with five of 15 possible links. Hence, the density is .33. In the network to the right, the density is much higher. Fourteen out of the possible 15 links are present, giving this network a density of .93.



**Fig. 3: Examples of low- and high-density networks**

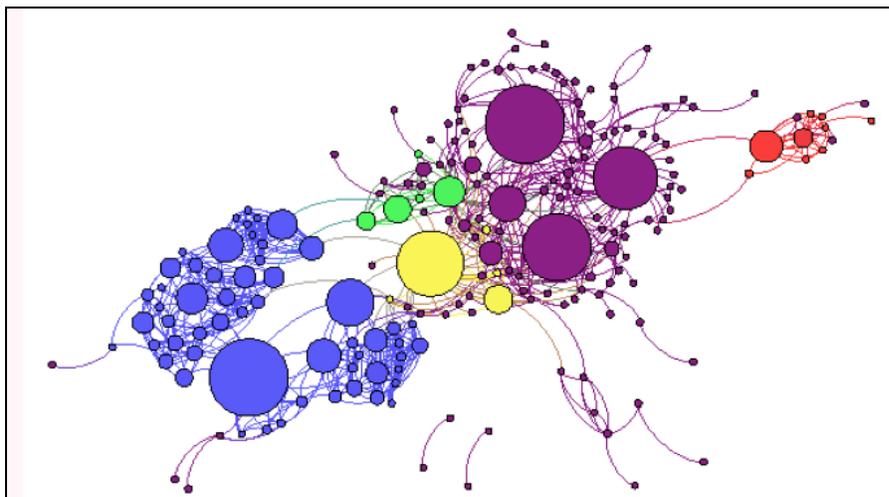
**Our findings:** The density of the Chicago information network found during the second crawl, excluding isolates, is 0.0295, meaning that 2.95% of all possible site-to-site links exist. While it is impossible to directly compare the current results to the previous ones since different web crawlers were used, we note that the first study found a network density of 0.161 (1.61%) for the 268 connected sites. So sites in the latest network study are more likely to link to one another than the sites studied in August 2010. Another indicator of greater linking is that the “average path length” – the average number of links needed to connect pairs of sites in the ecosystem – is significantly lower (from 2.9 to 2.1) than in the earlier study.

Here are the 15 sites with the most outbound links, and the number of sites they linked to:

<u>SITE</u>	<u>NUMBER OF LINKS</u>	<u>SITES LINKED TO</u>
1. mortongrove.patch.com	268	11
2. chicagobreakingnews.com	252	3
3. gapersblock.com	222	72
4. windycitizen.com (defunct)	219	32
5. progressillinois.com	192	27
6. lakeinthehills.patch.com	161	11
7. huffingtonpost.com	148	72
8. couriernews.suntimes.com	144	16
9. heraldnews.suntimes.com	140	14
10. napervillesun.suntimes.com	139	14
11. beaconnews.suntimes.com	137	13
11. chicago.everyblock.com	134	64
13. southtownstar.suntimes.com	131	22
14. newcitychicago.com	119	10
15. chicagoradioandmedia.com	118	46

## Clustering and Linking Patterns

Some clusters of websites in Chicago are tightly connected to one another. As one could imagine, it makes sense that sports sites link to other sports sites, while arts sites link to arts sites. To identify clusters, in Fig. 4 below, we viewed the network with continuous relationships (the number of links connecting the sites) as our edges. This allows us to impose a threshold so that a connection is displayed visually only if there were more than three links connecting the sites. The nodes are sized based on “betweenness centrality,” which you will read more about in the next section.



### Key

Red: Newcity sites  
 Green: Sun-Times sites  
 Yellow: Tribune sites  
 Blue: Patch sites  
 Purple: Other sites

### 5 largest “Other” sites, clockwise from top:

- \* huffingtonpost.com
- \* gapersblock.com
- \* windycitizen.com
- \* theexpiredmeter.com
- \* progressillinois.com

**Fig. 4: The Chicago website network, displaying only sites connected by at least three links**

In the diagram, it is easy to see that groups of sites operated by New City, the Sun-Times, Tribune Co. and Patch link heavily amongst themselves. There is even some clustering within these groups, such as with Patch, where it is split into two clusters for the northern suburbs and the western suburbs.

Looking at the patterns of links among the corporate website groups, we found that:

- 81.7% of links from sites owned by the Tribune Co. (chicago.metromix.com, chicagobreakingnews.com, chicagomag.com, chicagonow.com, chicagotribune.com, triblocal.com, vivelohoy.com, wgnradio.com and wgntv.com) went to other Tribune-owned sites;
- 80.4% of links from sites under the same ownership with the Sun-Times (beaconnews.suntimes.com, couriernews.suntimes.com, heraldnews.suntimes.com, naperville.sun.suntimes.com, oakpark.suntimes.com, pioneerlocal.com, southtownstar.suntimes.com and suntimes.com) went to other Sun-Times sites;
- 91.7% of links from Patch.com hyperlocal sites went to other Patch sites;
- 89.6% of links from sites operated by Newcity Communications, including the newcity.com website for its cultural weekly newspaper, went to other Newcity sites.

It's not unexpected to find that sites, especially from traditional media, link mostly to affiliated sites. A recent research paper<sup>4</sup> looked at links from different kinds of sites that covered three major news stories in 2010. Among the traditional-media news sites studied (from The New York Times, The Washington Post, The Wall Street Journal, CNN, ABC News and Time magazine), 91 percent of the links were internal – to the same site. Among journalistic blogs operated by the same companies, only 54 percent of links were internal. And among independent blogs, only 18 percent of links were internal.

While our webcrawling process did not capture information about internal links, our findings for these four corporate website groups are consistent with this research paper.

Also worth noting is the importance of several National Brand sites in bringing the Chicago ecosystem closer together through links: huffingtonpost.com, chicago.everyblock.com and patch.com. Among the five categories into which we grouped the Chicago websites, the National Brand sites were by far the most likely to link to other sites.

Even though many of the Patch links were to other Patch sites, we still counted 237 links from the 56 Patch sites in our database to other non-Patch sites in the ecosystem. This was

---

<sup>4</sup> Coddington, Mark, "Building Frames Link by Link: The Linking Practices of Blogs and News Sites" (International Journal of Communication 6 (2012), pp. 2007-2026.

considerably more than the number of external links from sites operated by the Tribune (90) and Sun-Times (176).

## Network Roles

When analyzing a network, one of the key approaches is to determine the different “roles” that each node plays based on its position in the network structure. The key metrics used for this type of analysis are different measures of **centrality**. As the term suggests, centrality refers to the extent to which a node is central to the network. Three of the most useful centrality measures are:

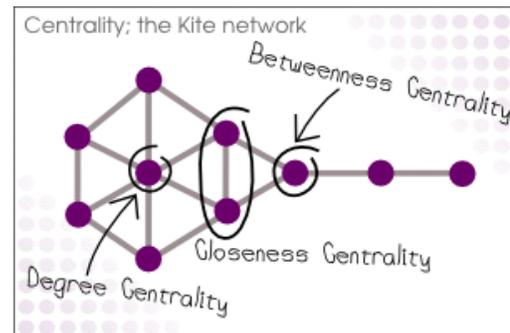
- Degree centrality;
- Betweenness centrality; and
- Closeness centrality.

We can use the "kite network," first developed by David Krackhardt, a leading researcher in social networks, to show the distinctions among the three measures.

**Degree centrality** simply counts the number of direct links with other nodes.

In Fig. 5, the node marked “Degree Centrality” has the highest degree centrality – the most direct links (six) to other nodes in the network.

In a hyperlink network, links are directional; site A can link to site B, even if site B does not link back to site A. So a node has both in-degree centrality and out-degree centrality. In-degree centrality refers to the extent that a site receives links. Out-degree centrality refers to the extent that a site links to others.



**Fig. 5: Kite diagram**

**Betweenness centrality** refers to nodes that are central because they mediate between parts of the network.

In the kite diagram, the node marked “Betweenness Centrality” has just three direct connections – fewer than the average in the network – but has one of the best locations in the network. This node could be defined as playing a “broker” role; it has potentially great influence over what flows through the network because it connects nodes that would otherwise not be connected. Without this “high betweenness” node, the two nodes on the right would not be connected to the larger group of nodes on the left. Nodes with high betweenness often connect otherwise disconnected communities.

**Closeness centrality** is the extent to which a node is close (in number of “hops”) to other nodes in the network.

In the kite diagram, the two nodes circled and marked “Closeness Centrality” have the shortest

average paths to all other nodes in the network. These nodes can reach other nodes in the network in relatively few steps – two or fewer in most cases, and no more than three.

Based on these measures, we are able to identify sites that are:

- **Authorities** and **Hubs** (with high degree centrality);
- **Switchboards** (with high betweenness centrality);
- **Referrers** and **Resources** (with high closeness centrality).

**Authorities** are sites with high in-degree centrality: many other sites tend to link to them. **Hubs** are sites with high out-degree centrality: they link to many other sites. Fig. 6 illustrates the difference between Authorities and Hubs in a simple ‘star’ network.



**Fig. 6: Authorities and hubs**

In a hyperlink news network, Authorities are often first-source news sites; many other sites tend to link to them because they are a reputable or original source of knowledge. Typically in a hyperlink network, an 80/20 principle applies, meaning that 20 percent of the sites receive 80 percent of the links. Many of these links come from Hubs, which tend to link out to many sites either because they cover a wide variety of content, provide a huge volume of content, or keep many active links on their pages.

## Authorities

Fig. 7 on the next page shows the network with circles representing each site scaled by the number of in-links. Thus, the largest circles are the most important Authorities. We used a radial axis layout for this diagram, as well as those for Authorities and Resources. This layout ranks all the sites based on a given characteristic, and then orders them in a circle by that characteristic. When two nodes have the same score, they are placed in a 'spear' out from the circle. In this diagram, the nodes are sized by their Authority score, and arranged by degree.

## Top Authorities

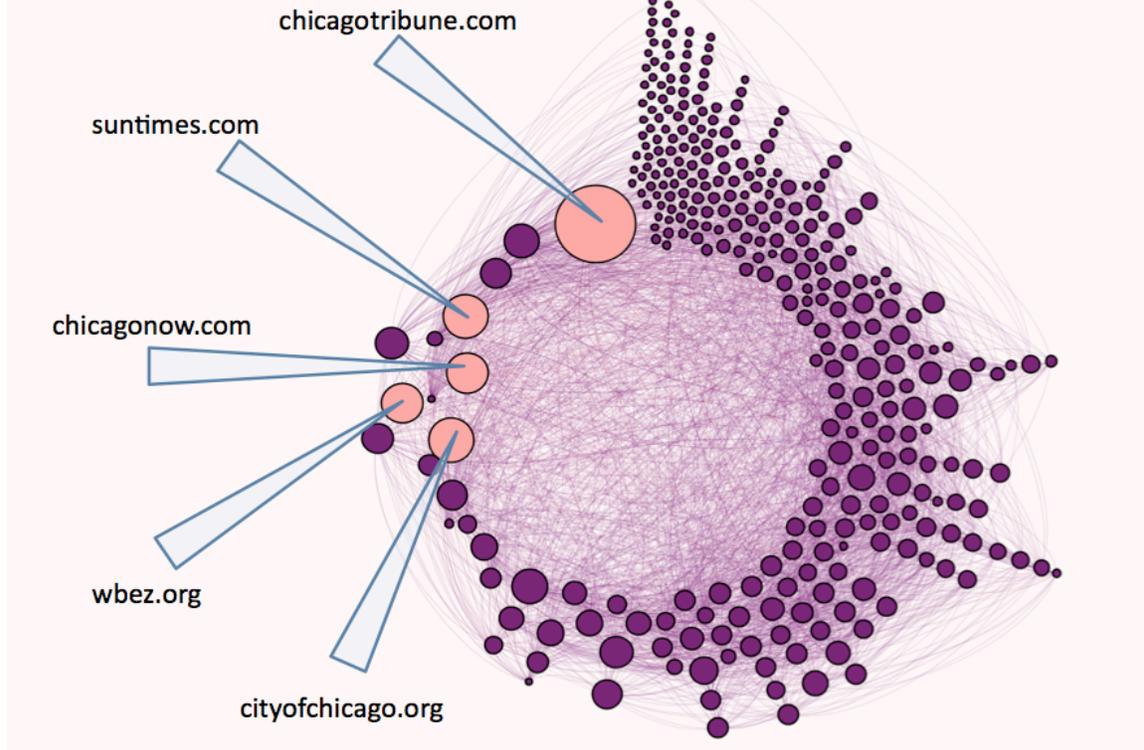


Fig. 7: Top authorities in the network, highlighting the top five

Top Authorities		
Ranking	Name	Category
1	chicagotribune.com	Legacy
2	cityofchicago.org	Org/inst
3	suntimes.com	Legacy
4	chicagonow.com	Legacy-affil
5	wbez.org	Legacy
6	transitchicago.com	Org/inst
7	huffingtonpost.com	Natbrand
8	chicagoreader.com	Legacy
9	nytimes.com	Legacy
10	chicagoist.com	Micro

Top Authorities include chicagotribune.com and suntimes.com (both leading sources of originally reported news in the region). Other prominent Authorities are: cityofchicago.org (Chicago’s city website), chicagonow.com (a blog hub operated by Tribune Co.) and wbez.org (the website for Chicago’s public radio station). Not surprisingly, Legacy and Legacy-affiliated sites (also including chicagoreader.com and nytimes.com) make up the majority of the top 10 Authorities. But as with the 2010 study, government-focused sites (cityofchicago.org and

transitchicago.com) also have a presence. And chicagoist.com, an independent news/blog hub, also makes the top 10.

## Hubs

Hubs are sites that guide users to relevant information on other sites by linking. Fig. 8 shows Hubs in the network, again using a radial axis arrangement ordered by degree:

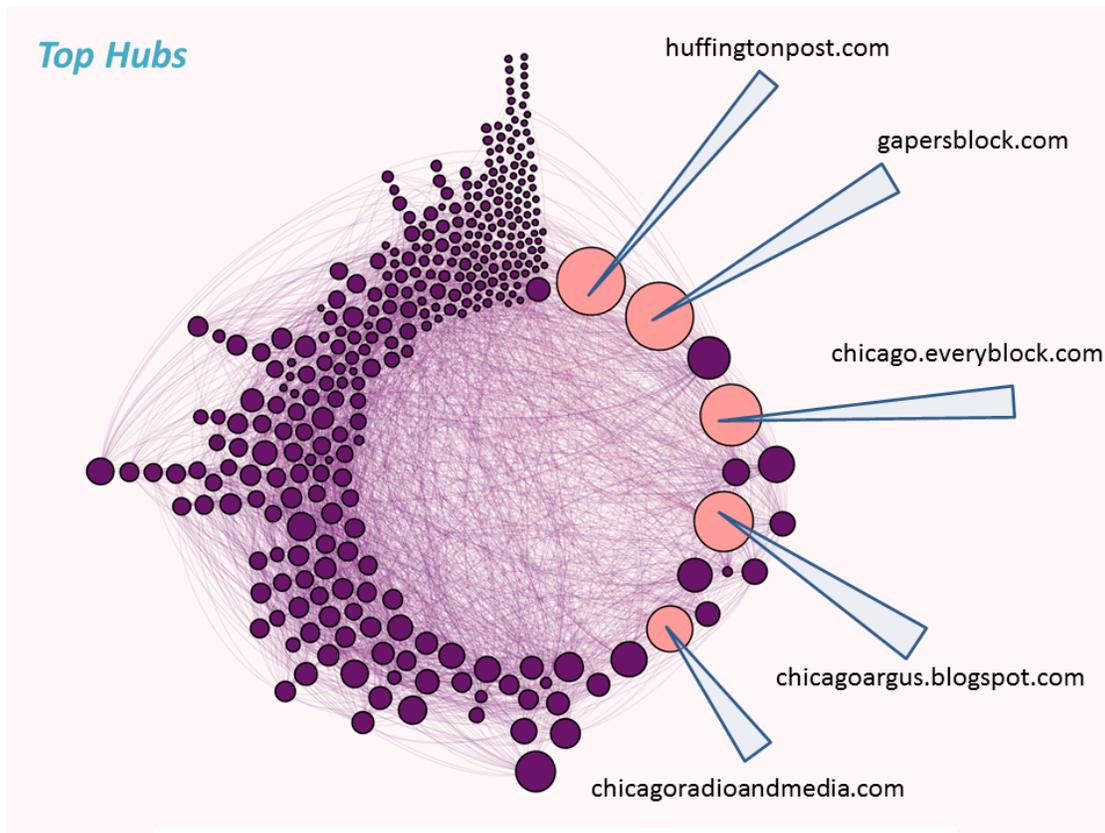


Fig. 8: Top hubs in the network, highlighting the top five

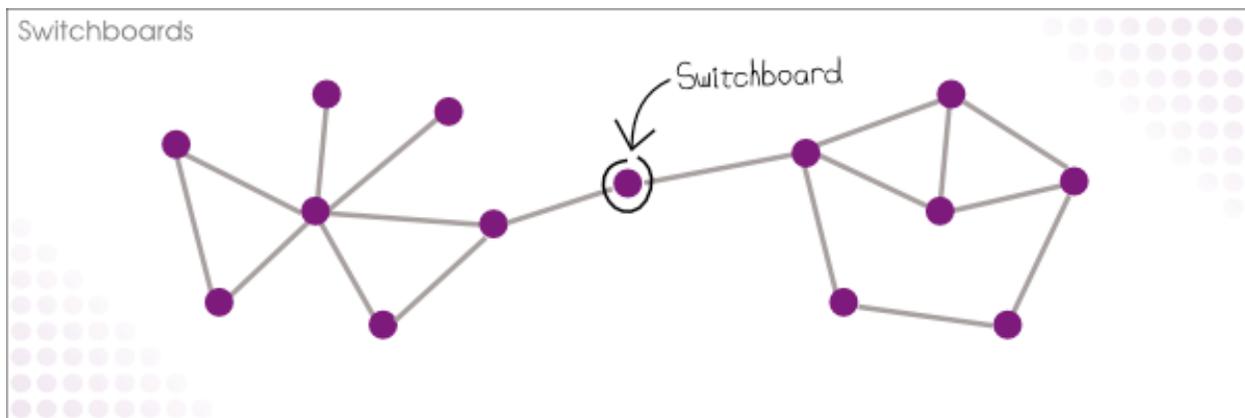
Top Hubs		
Ranking	Name	Category
1	<a href="http://huffingtonpost.com">huffingtonpost.com</a>	Natbrand
2	<a href="http://gapersblock.com">gapersblock.com</a>	Micro
3	<a href="http://chicago.everyblock.com">chicago.everyblock.com</a>	Natbrand
4	<a href="http://chicagoargus.blogspot.com">chicagoargus.blogspot.com</a>	Micro
5	<a href="http://chicagoradioandmedia.com">chicagoradioandmedia.com</a>	Micro
6	<a href="http://suntimes.com">suntimes.com</a>	Legacy
7	<a href="http://thesixthward.blogspot.com">thesixthward.blogspot.com</a>	Micro
8	<a href="http://chicagoreader.com">chicagoreader.com</a>	Legacy
9	<a href="http://examiner.com">examiner.com</a>	Natbrand
10	<a href="http://windycitizen.com">windycitizen.com</a>	Micro

Micro-publishers and national-brand sites (huffingtonpost.com, chicago.everyblock.com and examiner.com) make up eight of the top 10 Hubs in the Chicago news ecosystem, along with two Legacy sites (suntimes.com and chicagoreader.com). It is interesting to see that there are three sites – suntimes.com, chicagoreader.com and huffingtonpost.com – that appear on the top-10 list for both Authorities and Hubs. This illustrates the important position that these sites have in the news ecosystem; all of them link out and also receive links from other sites.

While our research indicates that the vast majority of the suntimes.com links go to other publications owned by the paper’s parent company, it is worth pointing out that these links still serve to guide the audience to potentially relevant content they might not otherwise have encountered.

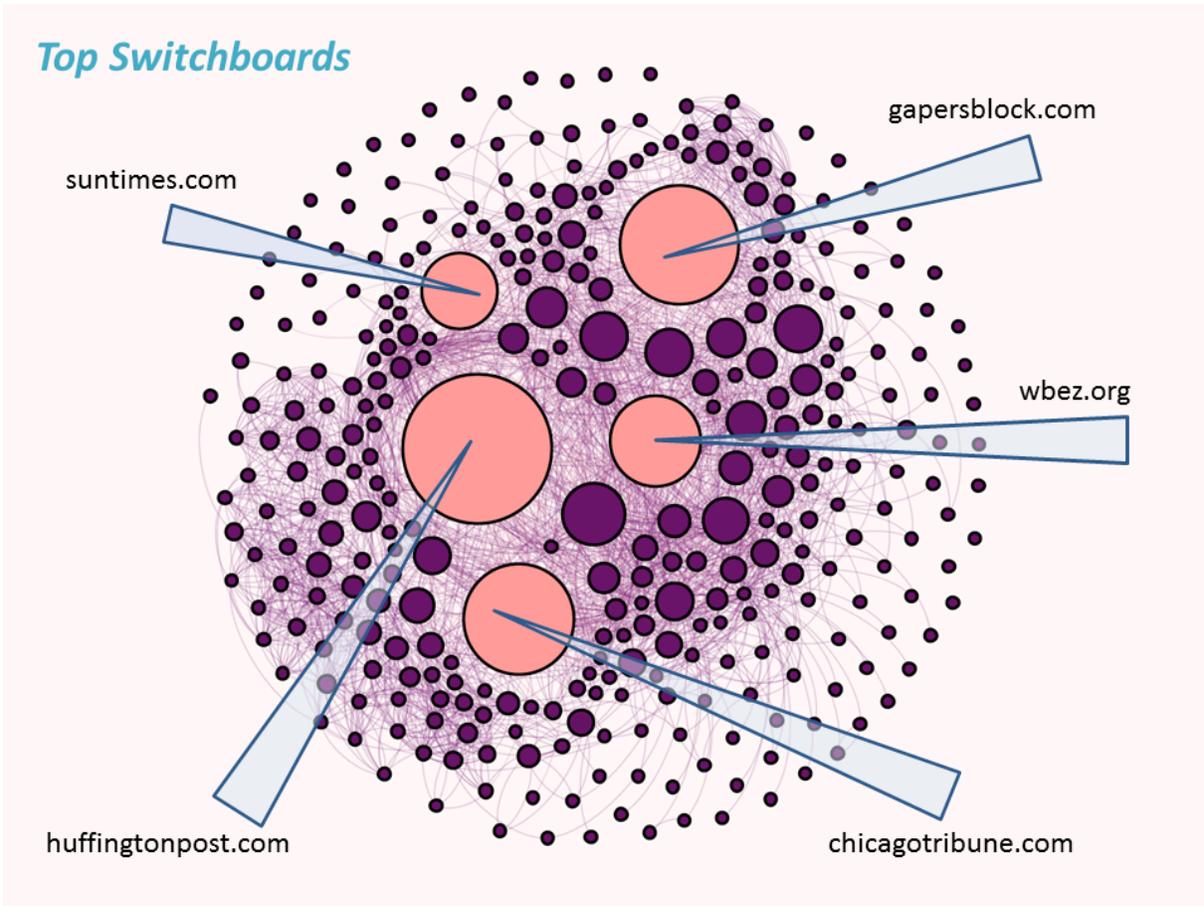
## Switchboards

Switchboards are sites with high “betweenness,” a measure of the number of *shortest paths* that pass through a particular node. A path can be defined as the steps – or “hops” – needed to get from any node in the network to another node. A site with high betweenness lies on many of the shortest paths – those with the fewest “hops” between different pairs of sites. In practice, this means Switchboards connect nodes that would otherwise not be connected at all – or, if removed from the network, would require users to make many more clicks to travel from one site to another site. See Fig. 9 for an example of a Switchboard in a simple network.



**Fig. 9: Example of a Switchboard**

In Fig. 10 on the next page, we look at the Chicago website network based on betweenness. The size of the circles represents the betweenness score.



**Fig. 10: Top Switchboards in the network, highlighting the top five**

Top Switchboards		
Ranking	Name	Category
1	<a href="http://huffingtonpost.com">huffingtonpost.com</a>	Natbrand
2	<a href="http://gapersblock.com">gapersblock.com</a>	Micro
3	<a href="http://chicagotribune.com">chicagotribune.com</a>	Legacy
4	<a href="http://wbez.org">wbez.org</a>	Legacy
5	<a href="http://suntimes.com">suntimes.com</a>	Legacy
6	<a href="http://chicagonow.com">chicagonow.com</a>	Legacy-affil
7	<a href="http://chicagoreader.com">chicagoreader.com</a>	Legacy
8	<a href="http://chicagoist.com">chicagoist.com</a>	Micro
9	<a href="http://chicagoartistsresource.org">chicagoartistsresource.org</a>	Org/inst
10	<a href="http://chicago.everyblock.com">chicago.everyblock.com</a>	Natbrand

Nine of the top 10 Switchboards have also appeared among the top 10 Authorities or Hubs. (The one exception is [chicagoartistsresource.org](http://chicagoartistsresource.org), a foundation-funded website designed to promote Chicago’s artists.) This is appropriate. Hubs and Authorities are highly linked with other sites, so they are also likely to be positioned along many of the shortest site-to-site paths – the definition of a Switchboard. In particular, it makes sense that [huffingtonpost.com](http://huffingtonpost.com) is the top Switchboard in the network; it is known for being a news aggregator and linking to other sites.

We used a Fruchterman-Reingold layout for the Switchboard diagram. It is a physics-based simulation that imagines nodes as mass particles and edges as the springs between them. It attempts to minimize the energy in this system. With this algorithm, the most central nodes in the network move to the center of the visualization.

## Referrers and Resources

Some nodes, by virtue of their structural position in the network, are ideally placed to quickly send or receive information. Closeness centrality is the metric used to determine which sites fit these two roles.

The World Wide Web is a directed network, which means that one site can link to a second site without the necessity of a return link. This means there are two types of closeness centrality on the Web.

A site with high “**out-closeness**” has a position in the network that would enable a user following links to quickly reach other sites in the network. This means the site’s outbound links lead to sites with yet more outbound links to other sites. We call these sites **Referrers**.

A site with high “**in-closeness**” has a network position that would enable a user to reach it quickly by following links from other sites. This means the site’s inbound links come from other sites that are, in turn, linked from yet other sites in the network. We call these sites **Resources**.

In Fig. 11, the node with high in-closeness is a Resource while the node with high out-closeness is a Referrer.

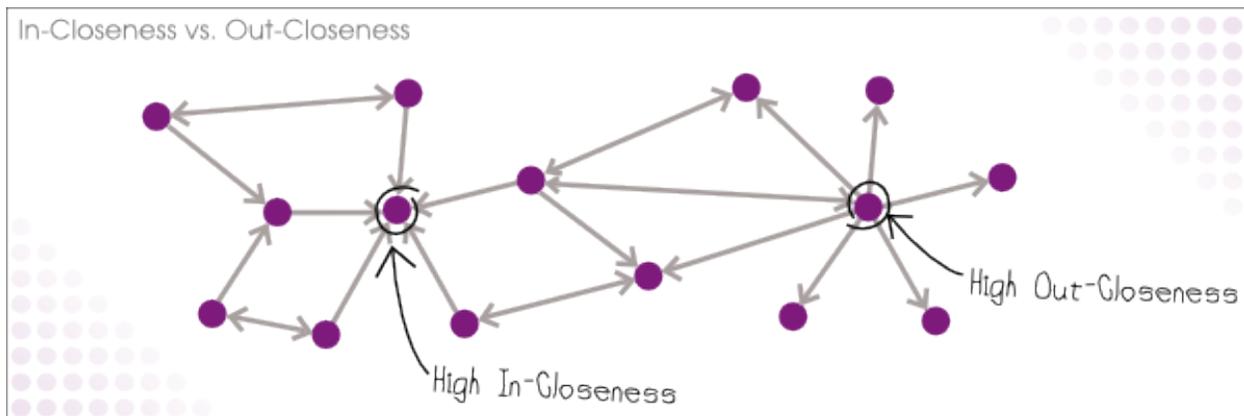


Fig. 11: In-closeness vs. out-closeness

## Referrers

Fig. 12 shows the Chicago network based on out-closeness. Once again, we used a Fruchterman-Reingold layout for this diagram.

## Top Referrers

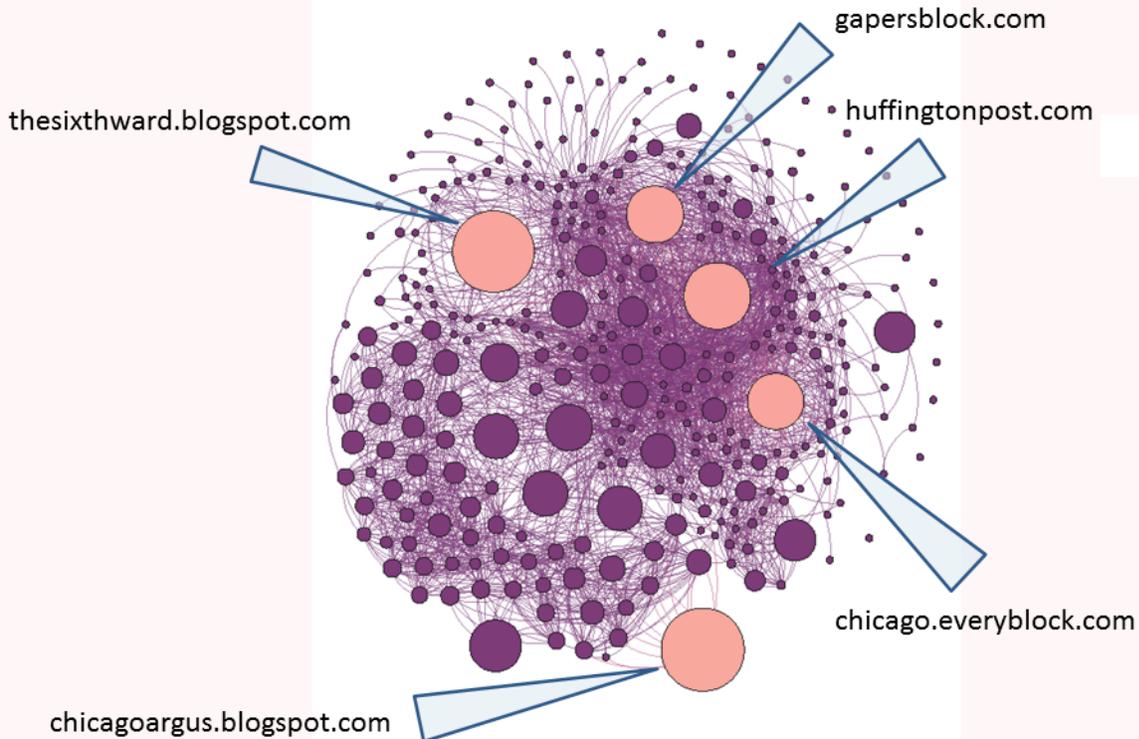


Fig. 12: Top Referrers in the network, highlighting the top five

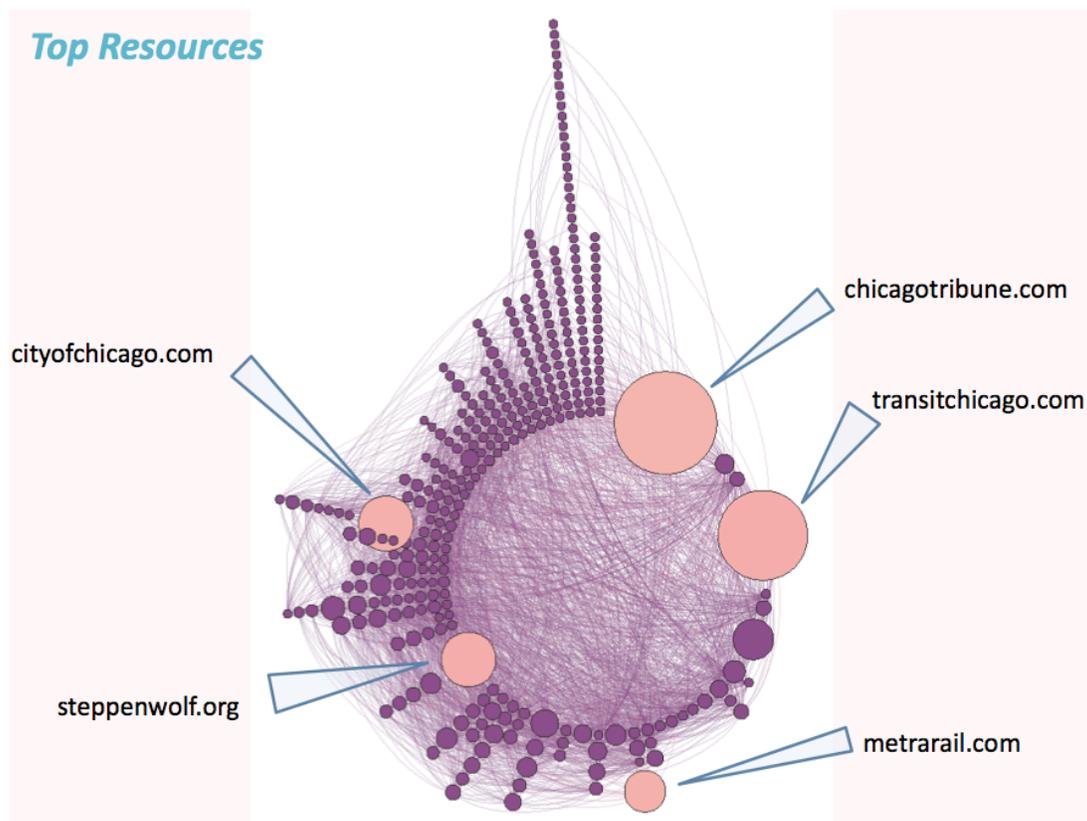
Top Referrers		
Ranking	Name	Category
1	chicagoargus.blogspot.com	Micro
2	thesixthward.blogspot.com	Micro
3	huffingtonpost.com	NatBrand
4	gapersblock.com	Micro
5	chicago.everyblock.com	NatBrand
6	chicagoradioandmedia.com	Micro
7	progressillinois.com	Micro
8	windycitizen.com	Micro
9	communitymediaworkshop.org	Org/Inst
10	proyectolatina.org	Micro

Not surprisingly, considering the way Referrers are measured, there is considerable overlap between Referrers and Hubs. Seven of the top Referrers also appear on the top-10 Hubs list. Three top Referrers don't show up on other top-10 lists: progressillinois.com (a union-financed news site covering state government and politics), communitymediaworkshop.org (a non-profit that helps community organizations communicate their stories to the media and the public) and

proyectolatina.org (a site that hosts multimedia stories about Latinas in Chicago). Among the top 10 Referrers, there are no Legacy sites; most are either micro-publishers or national brands with a major local presence (chicago.everyblock.com or huffingtonpost.com). What differentiates these sites is that they are targeted at a specific niche or geographic area.

## Resources

Fig. 13 shows the sites that are most easily reached by clicking from link to link in the Chicago news ecosystem. The visualization uses a radial axis with sites ordered by degree. This layout ranks all the sites based their in-closeness score, and then orders them in a circle by that characteristic. When two nodes have the same score, they are placed in a 'spear' out from the circle.



**Fig. 13: Top Resources in the network, highlighting the top five**

Top Resources		
Ranking	Name	Category
1	chicagotribune.com	Legacy
2	transitchicago.com	Org/inst
3	cityofchicago.org	Org/inst
4	steppenwolf.org	Org/inst
5	metrarail.com	Org/inst
6	illinois.gov	Org/inst
7	wbez.org	Legacy
8	chicagonow.com	Legacy-affil
9	msichicago.org	Org/Inst
10	siskelfilmcenter.org	Org/Inst

Five of the top 10 sites also appear among the top 10 Authorities: [chicagotribune.com](http://chicagotribune.com), [transitchicago.com](http://transitchicago.com), [cityofchicago.org](http://cityofchicago.org), [wbez.org](http://wbez.org) and [chicagonow.com](http://chicagonow.com). Five other sites don't appear on any of the other top-10 lists: [steppenwolf.org](http://steppenwolf.org) (theater company), [metrarail.com](http://metrarail.com) (commuter rail agency), [illinois.gov](http://illinois.gov) (state website), [msichicago.org](http://msichicago.org) (Museum of Science and Industry) and [siskelfilmcenter.org](http://siskelfilmcenter.org) (center for independent film affiliated with the School of the Art Institute of Chicago).

### Websites with High Centrality: Comparing Two Studies

Many of the most highly ranked websites, based on network measures, are different than the ones identified in the first study. This is not necessarily surprising, since linking patterns can change and since we modified our methodology. Still, there are several sites that appeared among the most influential websites in both studies:

- [chicagoartistsresource.org](http://chicagoartistsresource.org)
- [chicagoist.com](http://chicagoist.com)
- [chicagoreader.com](http://chicagoreader.com)
- [chicagotribune.com](http://chicagotribune.com)
- [cityofchicago.org](http://cityofchicago.org)
- [communitymediaworkshop.org](http://communitymediaworkshop.org)
- [gapersblock.com](http://gapersblock.com)
- [metrarail.com](http://metrarail.com)
- [msichicago.org](http://msichicago.org)
- [nytimes.com](http://nytimes.com)
- [suntimes.com](http://suntimes.com)
- [transitchicago.com](http://transitchicago.com)
- [windycitizen.com](http://windycitizen.com) (now defunct)

That these sites showed up prominently in similar studies two years apart suggests that they are not only important in the network but also have remained so over time.

## Web Analytics Research

The analysis of web-analytics data for 98 sites is a novel aspect of this study. We know of no prior comparable research examining the degree to which audiences actually follow links among sites in a local online ecosystem.

To put it simply, knowing whether or not sites link to one another is helpful in identifying the structure of the network – but knowing whether or not people click on those links is critical in understanding which links matter. Thanks to the way web servers work, every site has access to information on which inbound links are bringing traffic. But other than analytics providers themselves, no one else usually has access to information about multiple sites operated by different publishers in an ecosystem. We were able to gain that access for 98 sites.

Because of the inconsistencies in the data provided by publishers, there are limits to the conclusions we are able to draw at this time. But we have identified a number of interesting findings that demonstrate the value of this type of research. And we would welcome the opportunity to collaborate with other scholars, and with Chicago area websites, to collect additional data and extend this research in the future.

Here are some of our key findings:

### **1. Smaller sites rely more on traffic from within the ecosystem than larger sites.**

By matching referring URL's to the network of sites we assembled through the links research, we were able to attribute 390,000 visits as coming from sites in the Chicago ecosystem. (This understates the true amount of inside-the-ecosystem traffic, since we have analytics for only 98 of the 301 sites.)

Visits from other Chicago-ecosystem sites made up 8.5 percent of the 4.6 million visits in our Web-analytics database. But this varies widely depending on the size of the site:

- For large sites (chicagotribune.com, suntimes.com and dailyherald.com), 2.2 percent of visits come from within the ecosystem.
- For small and medium sites, 24.8 percent of the referring traffic comes from other Chicago ecosystem sites. So small and medium sites are more than 11 times as reliant as large ones on traffic from within the Chicago market.

### **2. Social media, especially Facebook, are critically important to driving traffic.**

More than 23 percent of the referred visits in the ecosystem come from Facebook (18.9 percent) and Twitter (4.4 percent). The social media traffic is particularly important for smaller sites, as shown in the table below:

<u>Size of site</u>	<u>Avg. visits from Facebook</u>	<u>Avg. visits from Twitter</u>	<u>Total</u>
Large	14.5%	4.2%	18.7%
Medium	25.6%	5.4%	31.0%
Small	48.1%	3.6%	51.7%

Tumblr and Pinterest, other social media services that have attracted considerable attention recently, at this point contribute relatively little traffic to the ecosystem: an average of 0.3 percent for Tumblr and 0.2 percent for Pinterest.

### **3. Other than social media, the sites that refer traffic are different for different sizes of sites.**

**Among the large and medium sites** (most having traditional media ties), the list of major traffic drivers is dominated by partners and service providers such as homefinder.com (real estate), careerbuilder.com (employment) and legacy.com (death notices). Other sites showing up as traffic drivers include national sites that serve as hubs for different kinds of content: drudgereport.com (news), fark.com (offbeat/humor), bleacherreport.com (sports) and realclearpolitics.com (politics).

**Among small and medium sites**, the list of top referrers is more idiosyncratic, varying widely by site. The referrers are often “mini-hub” sites that aggregate news on certain topics. A few examples:

- [chicago-l.org](#), an aggregator of headlines about mass transit in Chicago, sent traffic to [progressillinois.com](#), a union-backed news site.
- [theatreinchicago.com](#), which covers the local theater scene, sent traffic to [newcitystage.com](#), operated by Newcity Communications.
- [curbed.com](#), a site focusing on real estate, sent traffic to [centersquarejournal.com](#), a hyperlocal site for the Lincoln Square neighborhood.

### **4. There is a strong relationship between links and traffic, at least for small sites.**

One of the reasons we wanted to do the analytics research was to develop a deeper understanding of the relationship between links and traffic. We started by calculating the “correlation coefficient,” a statistical measure that indicates how closely two variables are related to one another. Looking at the links and referred traffic for all 98 sites in our database, the correlation coefficient is .092. In statistics, that means that for any given pair of linked sites in our data set, the number of links received can explain only 9.2% of the variance in the amount of traffic referred.

We suspected, though, that the relationship between links and traffic would be easier to discern if we looked at sites based on their size. After all, the scale of the sites is very wide. The 18 large and medium sites, most of them related to established brands, averaged 254,400 referred visits

during the two-week period<sup>5</sup>. The 80 small sites, by contrast, averaged just 3,200 referred visits. So the average large or medium site gets 80 times as much referred traffic as the average small site.

Indeed, when we calculated the correlation coefficient based on the size of the referring site, we found:

- For the 18 large/medium sites, the number of outbound links explains 14.3 percent of the variation in the amount of referred traffic. This is a meaningful, but modest, relationship.
- For the 80 small sites, the number of outbound links explains 49.2 percent of the variance in the amount of referred traffic. In other words, it is even clearer among small sites that more links equate to driving more traffic.

These findings confirm that there is a statistical relationship between the number of links a site provides and the amount of traffic it drives – and that the relationship is especially strong for the smallest sites.

## **5. Sites that drive more traffic per link are those that link prolifically and prominently.**

One of our goals in this research was to identify the linking practices that are most effective in driving traffic. Unfortunately, our data sets are not complete enough to allow us to draw sweeping conclusions.<sup>6</sup> But we did want to seek out referring sites that seemed to be especially successful in driving traffic, and observe their linking practices. Since we had documented a strong relationship between links and traffic from the 80 small sites, we focused our attention there.

We counted 535 outbound links from small sites in our database. On a per-link basis, there was an average of 16.1 referred visits. Three sites drove more than the average number of visits per link:

- gapersblock.com (106 links, 41.1 visits per link)
- centersquarejournal.com (15 links, 39.8 visits per link)
- catalyst-chicago.org (17 links, 33.3 visits per link)

Our ability to understand the reasons these sites drive more traffic per link is limited by a lack of data about which specific links drove how many visits. This information is available through the sites' analytics systems, however – so this could be an avenue for future research.

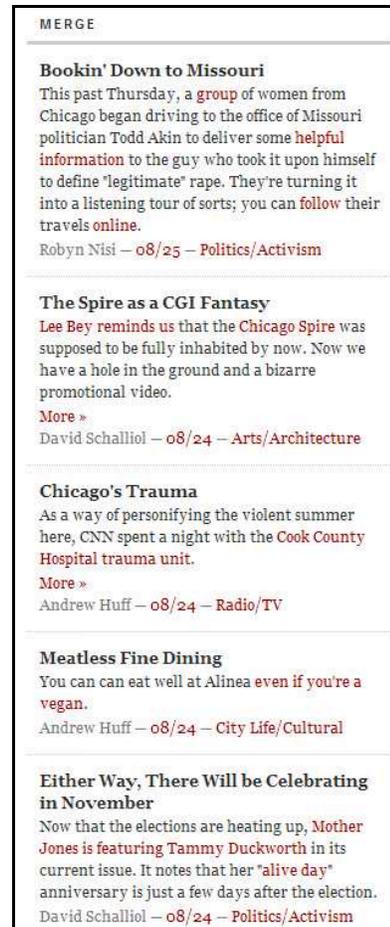
---

<sup>5</sup> We grouped the large and medium sites together to have a large enough set of sites to analyze – there were only three large sites in our data set.

<sup>6</sup> Key limitations: We do not have complete referring-URL information. And neither our links data nor our referring URL data include the specific Web addresses of the linking and linked-to pages. So even for pairs of sites in both data sets, we have a number of links and a number of referred visits, but no guarantee that those links drove those specific visits.

Still, looking at these three websites, we can see that all three prioritize linking externally. On one day in late August, we counted the number of links from the home page of the three sites to specific content (as opposed to home pages) on other websites. We counted 15 links from [centersquarejournal.com](http://centersquarejournal.com) (most of them to sister sites [edgevillebuzz.com](http://edgevillebuzz.com) and [roscoeviewjournal.com](http://roscoeviewjournal.com)) and 16 links from [gapersblock.com](http://gapersblock.com). There were only two links from the home page of [catalyst-chicago.org](http://catalyst-chicago.org) – but the site’s most frequently updated section, Catalyst Notebook, had five more.

Gapersblock.com, in particular, is known for its front-page “Merge” feature (screen capture at right), which is a constantly updated feed of links to content elsewhere on the Web. Unlike some sites, which require the user to click through an intermediate page to get to linked content featured on the home page, [gapersblock.com](http://gapersblock.com) enables users to go straight off to the other site – accepting the risk that users will go immediately from the home page to another site and not return.



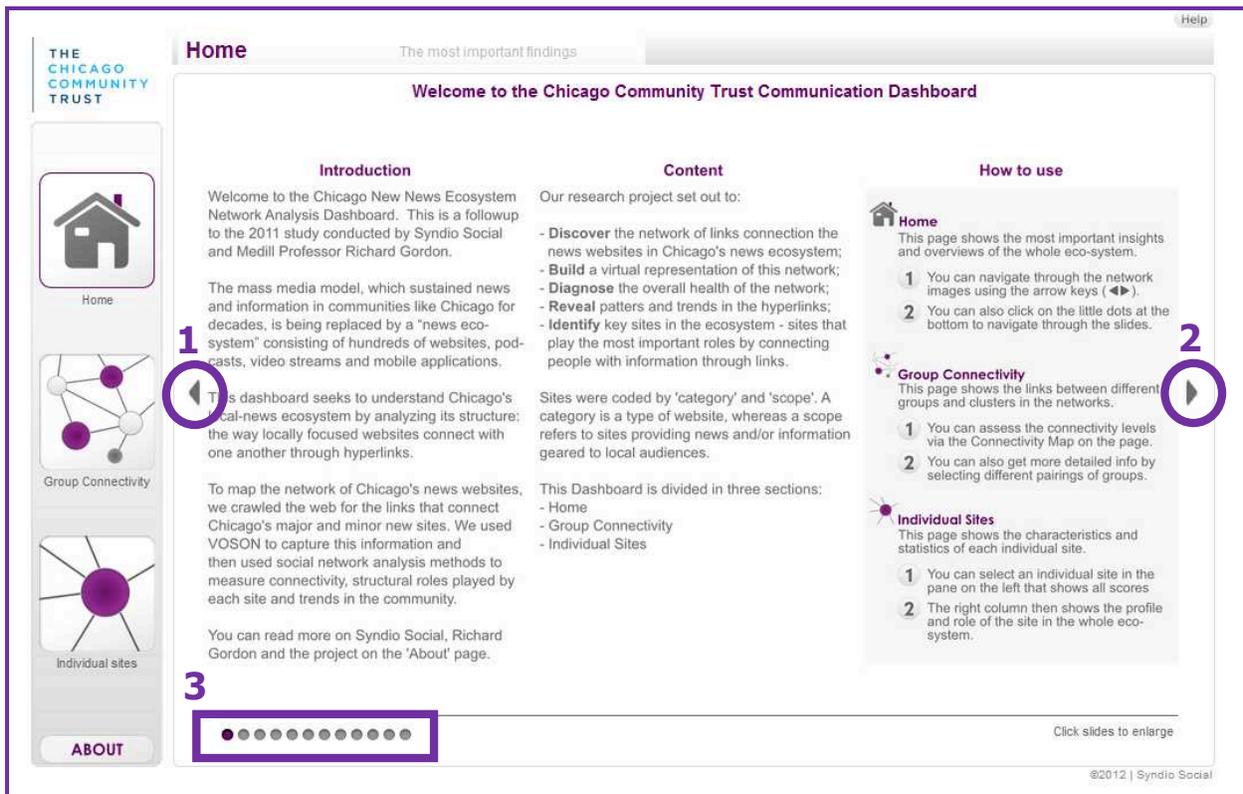
**Fig. 14: “Merge” section on the home page of [gapersblock.com](http://gapersblock.com)** and the ecosystem – especially to small sites.

The fact that [gapersblock.com](http://gapersblock.com) seems to drive a high number of links to other sites, it has a particularly important role in referring traffic through the ecosystem – especially to small sites.

# Exploring the Ecosystem with the Interactive Dashboard

After the first study, we realized that many people wanted to drill deeper into the links data than was possible in a printed report, so this time we created an interactive online dashboard available at <http://public.syndiosocial.com/ChiLinks>. Please note that this dashboard incorporates only the research on links, not on referred traffic.

The **home page** includes an overview of the research, through a series of visualizations navigated through using arrows (circled below, 1 and 2) or the small circles at the lower left (see square, 3). When you view visualizations, you can magnify the image by clicking on it and using the browser zoom tools (ctrl-plus to zoom in, Ctrl-minus to zoom out)

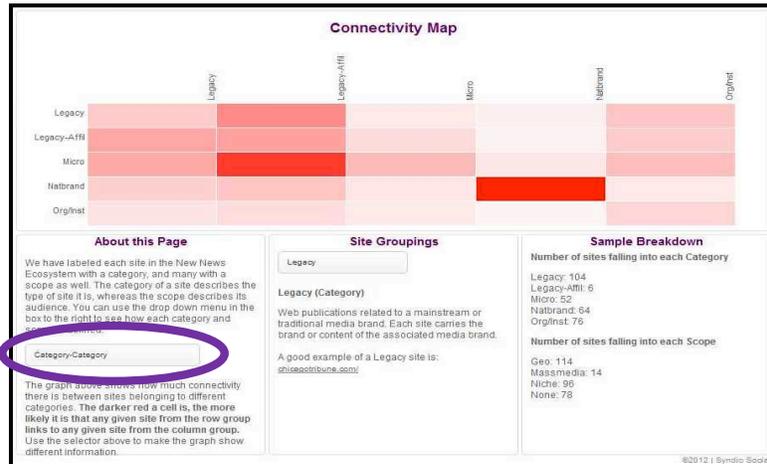


**Group Connectivity** provides information about links among different kinds of sites. This section uses the Category and Scope groupings described earlier in this report. In this section, you can see what kinds of sites are most likely (and least likely) to link to what other kinds of sites.

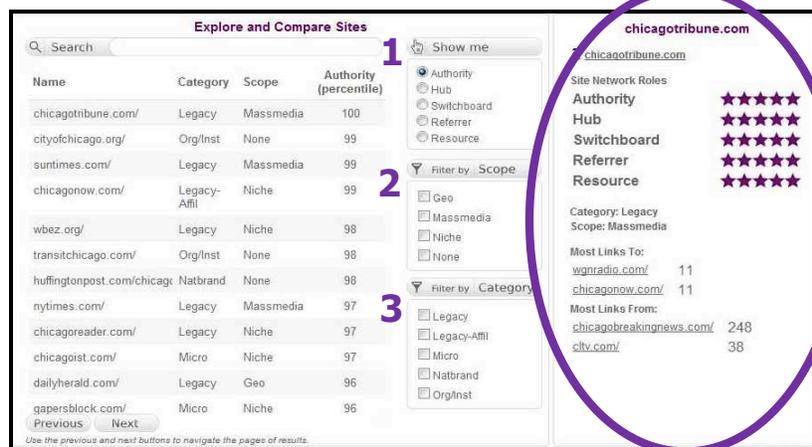


**Individual Sites** provides additional details about each of the 301 sites in our database. You can look at a ranked list of the top Hubs, Authorities, Switchboards, Referrers and Resources. You can also choose to look only at sites in specific Category and Scope groupings.

The **Group Connectivity** panel allows you to see the likelihood of linking between different types of sites. The darker the shade of red, the greater the likelihood of linking between those types of sites. If you mouse over the rectangles you can see the likelihood that sites of the type listed on the left side will link to a site of the type listed along the top. Outlined in an oval below is a pull-down menu enabling you to switch between Category, Scope and combinations of the two.



The **Individual Sites** panel displays ranked lists of sites based on network metrics (Hubs, Authorities, etc.) and by Category and Scope. Sites are ranked by percentile. In the center panel, you can sort and filter by the network criteria you are interested in (1), as well as the Scope (2) and/or Category (3).



When you click on a site name in the ranked list on the left, you will see a brief profile of that site (the oval above). The profile shows how highly rated the site is as a Hub, Authority, Switchboard, Referrer and Resource. It also shows a few of the sites it links to the most and is linked from the most.

## Looking to the Future

Having completed two research projects, we have learned a lot about the structure of the Chicago news ecosystem and the patterns of linking among its websites. We believe the results can be useful to website publishers, academic researchers and organizations like The Chicago Community Trust that are interested in strengthening the local news ecosystem.

Here are some ideas for next steps:

### For publishers

- **Link more to other sites in the ecosystem.** Sites that link out a lot also can get links back. At a national level, huffingtonpost.com has demonstrated this – and Huffington Post Chicago has quickly become one of the most influential sites in the local market.
- **Emphasize social media as a traffic driver.** Especially for small sites, it is now as important to have a Facebook page as a home page. Devote time and attention to winning “likes” of your Facebook page, and update the page frequently.
- **Find ways to aggregate and distribute headlines across sites in the ecosystem.** The proliferation of locally focused websites suggests a need to expose audiences to content they would not otherwise be aware of. For instance, a Chicago ecosystem widget with the most recent headlines from the region’s websites could be created and made available to all sites in Chicagoland.

### For researchers

- **Continue data collection and analysis for this research project.** We are interested in finding researchers who would collaborate with us on additional work with Web analytics. With additional help, we could work with publishers to get identically structured data from a much larger number of sites.
- **Conduct comparable research in multiple markets.** Our research has demonstrated the value of network analysis in understanding the structure of the ecosystem of news and information sites in a geographic region. Comparable studies in multiple markets would enable us to compare news ecosystems to one another and identify ways of improving these ecosystems.

### For organizations concerned about the Chicago news environment

- **Encourage more linking through incentives.** The Chicago Community Trust used the findings in the first study as input to its decisions on grants under the Community News Matters program. Incentives based on the volume of outbound links could encourage more publishers to link outward.

- **Encourage local publishers to collaborate on a headline-sharing service** such as the widget suggested above for publishers.
- **Promote link-sharing partnerships between traditional media and “new news” sites.** J-Lab, the Institute for Interactive Journalism, has implemented a “Networked Journalism Project” to nurture and support this kind of partnership in other markets.
- **Take network “snapshots” on a regular basis.** The news ecosystem evolves over time as new sites are launched and others become dormant or defunct. Looking at the network regularly through a process like this one can yield deeper understanding of which sites are most critical to a healthy news ecosystem.
- **Consider establishing a data repository of Web analytics data for local websites.** Our examination of this data demonstrates that it is possible to aggregate analytics for multiple sites in order to gain a deeper understanding of the connections in the ecosystem. While some site managers are reluctant to share their data, we believe this kind of data collection – and the creation of benchmarks – can yield multiple insights relevant to individual sites as well as the network as whole. We think it’s possible to come up with confidentiality terms that publishers can live with in return for gaining additional insights into the behavior of their audiences.

## About the Authors

### **Rich Gordon**

Rich Gordon is professor and director of digital innovation at the Medill School of Journalism, Media, Integrated Marketing Communication at Northwestern University. His research and teaching focus on interactive publishing, including online audience development, social media and online communities. He co-developed, with Northwestern Prof. Noshir Contractor, an undergraduate course focusing on the application of network theory to media and journalism.

Before coming to Northwestern, Prof. Gordon was the first online director for the Miami Herald Publishing Co., overseeing editorial and business operations. Before that he worked as reporter, bureau chief and editor for newspapers in Virginia and Florida, where he was one of the early leaders in computer-assisted reporting.

### **Zachary Johnson**

Zachary Johnson is the Co-Founder and CEO of Syndio Social. In January 2009, he initiated the expansion of the company by partnering with Professor Noshir Contractor to provide scalable network analysis solutions. He has secured and led projects for numerous clients, including Procter & Gamble.

Mr. Johnson studied network analysis at Northwestern University under the tutelage of Prof. Contractor. Prior to founding Syndio Social, he applied network analysis methodologies to numerous independent music marketing projects. His most successful collaboration to date is with pop-star Mike Posner. Mr. Johnson advised the digital distribution of Posner's debut mix tape, "A Matter of Time," which has established him as an up-and-coming artist. Posner's hit single "Cooler Than Me," reached the tops of both Billboard charts and the iTunes music store.



## Linking Audiences to News II: 2012 Network Analysis of Chicago Websites

By **Rich Gordon** (Northwestern University)  
and **Zachary Johnson** (Syndio Social)  
September 2012

© 2012 - Syndio Social

All rights reserved. Any redistribution or reproduction of part or all of the contents in any form is prohibited other than the following: a) you may print or download to a local hard disk extracts for your personal and non-commercial use only; b) you may copy the content to individual third parties for their personal use, but only if you acknowledge Syndio Social as the source of the material. You may not, except with our express written permission, distribute or commercially exploit the content.

### THE CHICAGO COMMUNITY TRUST

AND AFFILIATES

225 North Michigan Avenue  
Suite 2200  
Chicago, IL 60601  
312.616.8000  
[www.cct.org](http://www.cct.org)

The Chicago Community Trust, our region's community foundation, partners with donors to leverage their philanthropy in ways that transform lives and communities. The Trust, together with its donors, grants more than \$100 million annually to nonprofits working to develop new audiences to sustain the region's vibrant arts organizations, protect the human services safety net for those hardest hit by the recession, stem the devastating effects of foreclosures on our neighborhoods, elevate teaching to meet world-class standards, and improve conditions for healthy and active lifestyles. By connecting the generosity of donors with the needs of the community, the Trust ensures that our region thrives today and for future generations.