JournalList Ecosystem Web Crawler

Last Updated: October 17, 2021

This document describes how to use a Python based web crawler to download all the trust.txt files that can be reached by traversing all the trust.txt files referenced from a root URL (by default https://www.journallist.net/). This document is intended for Mac computers running macOS and assumes some familiarity with the "terminal" application and the UNIX shell.

Preparing the environment:

1. Download the Python environment: https://www.python.org/downloads/
2. Install the Python requests module: https://docs.python-requests.org/en/master/user/install/#python-m-pip-install-requests
3. Download the yEd graphml editor/visualizer: https://www.yworks.com/products/yed/download#download

Open a terminal window and change directory to the Journalist directory. Pull the following files from the JournalList GitHub repository: https://github.com/brownwolf1355/JournalList

1. cron.sh - a bash shell script that runs the python webcrawler, processes the results through sqlite, and generates graphml files of the results.
2. webcrawler.py - a python script that recursively crawls trust.txt files to capture the state of the trust.txt ecosystem. It captures a copy of all of the trust.txt files it finds and generates a .csv file of the contents of all of them.
3. init.sql - the initialization sqlite script that creates the intermediate tables used in the following sql script.
4. symmetric - a sql script that generates .csv files containing the symmetric links in the trust.txt ecosystem and list of associations, publishers, and vendors discovered.
5. graphml.py - a python script that generates three graphml files containing the symmetric links, the asymmetric links, and the full ecosystem including both the symmetric and asymmetric links.
6. qa_trust_txt.py - a python script that parses a trust.txt file and lists any errors it contains.

Run the cron.sh shell script:

```
$ bash cron.sh
```

The shell and python scripts will create a directory "Webcrawl-YYYY-MM-DD" that will contain all the results of crawling the trust.txt files. This directory will contain the following:

1. The trust.txt files downloaded through the web crawling process. Each trust.txt file will be given a unique name "www.<domain>-trust.txt".

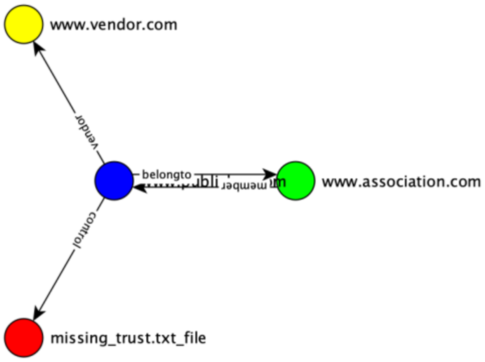
2. A log file useful for debugging, should anything go awry.
3. Several .csv files
    a. Webcrawl-YYYY-MM-DD.csv containing three columns: “srcurl,” “attr”, and “refurl”. Representing the Source URL, Attribute, and Referenced URL on each line of all downloaded trust.txt files.
    b. Webcrawl-YYYY-MM-DD-err.csv containing four columns: “srcurl”, “attr”, “refurl”, and “error”. Representing the Source URL, Attribute, Referenced URL, and the Error encountered when the Referenced URL trust.txt file is attempted to be downloaded, a blank “www.<domain>-trust.txt” file will be created in these cases. Errors that are identified as “HTTP Status Code:” or “Content type: text/html” indicate that the server was reached, but a trust.txt file doesn’t exist. Errors that are identified as “HTTP GET connection error exception occurred:” indicate that the server wasn’t able to be reached and may indicate an error in the spelling of the Referenced URL.
    c. Webcrawl-YYYY-MM-DD -symmetric.csv containing three columns: “srcurl,” “attr”, and “refurl”. Representing the Source URL, Attribute, and Referenced URL for all symmetric links (e.g., “member” <-> “belongto” or “control” <-> “controlledby” or “vendor” <-> “customer”). The Attribute will be one of “member”, “control”, or “vendor”.
    d. Webcrawl-YYYY-MM-DD -asymmetric.csv containing three columns: “srcurl,” “attr”, and “refurl”. Representing the Source URL, Attribute, and Referenced URL for all asymmetric links (e.g., “member” -> “belongto” reference without a corresponding “belongto” -> “member” reference). Note, these are the links listed in the Webcrawl-YYYY-MM-DD-err.csv file.
    e. Webcrawl-YYYY-MM-DD-associations.csv containing one column “srcurl” listing all of the Source URLs that have more “member” attributes than “belongto” entries and all of the Referenced URLs that have “control” attributes.
    f. Webcrawl-YYYY-MM-DD-publishers.csv containing one column “srcurl” listing all of the Source URLs that have more “belongto” attributes than “member” entries and all of the Referenced URLs that have “control” attributes.
    g. Webcrawl-YYYY-MM-DD-vendors.csv containing one column “srcurl” listing all of the Referenced URLs that have “vendor” attributes.
    h. Webcrawl-YYYY-MM-DD-symmetric.graphml containing the graphml diagram for the all symmetric links discovered.
    i. Webcrawl-YYYY-MM-DD-asymmetric.graphml containing the graphml diagram for the all asymmetric links discovered.
    j. Webcrawl-YYYY-MM-DD.graphml containing the graphml for the all links discovered, both symmetric and asymmetric.

There are a couple ways to analyze these results:

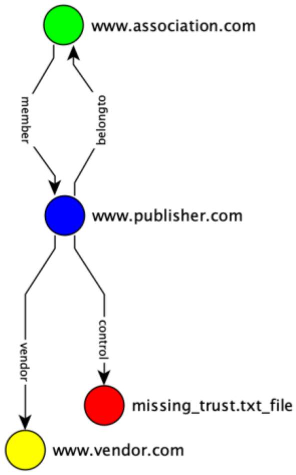
1. Through a yEd visual layout of the trust.txt ecosystem social graph
2. Through an Excel Pivot Table.

To visualize the trust.txt ecosystem social graph, open the graphml files int yEd. Because the generated graphml files do not have any imbedded layout, the graph will be displayed with all the nodes on top of each other:

Use yEd's layout tool to display the trust.txt social graph. I have found that the "Organic" layout with "Preferred Edge Length" set to 100 to be most useful. You may wish to manually place the graph legend (consisting of 4 nodes and 4 edges):

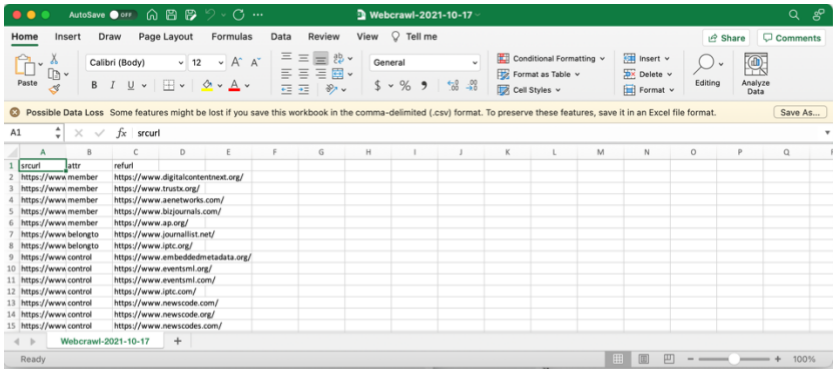


Use the selection tool to select these nodes and edges. You can obtain a better layout using the "Selection (Partial)" layout tool:



The "Component" layout tool has options for organizing the layout of the social graph.

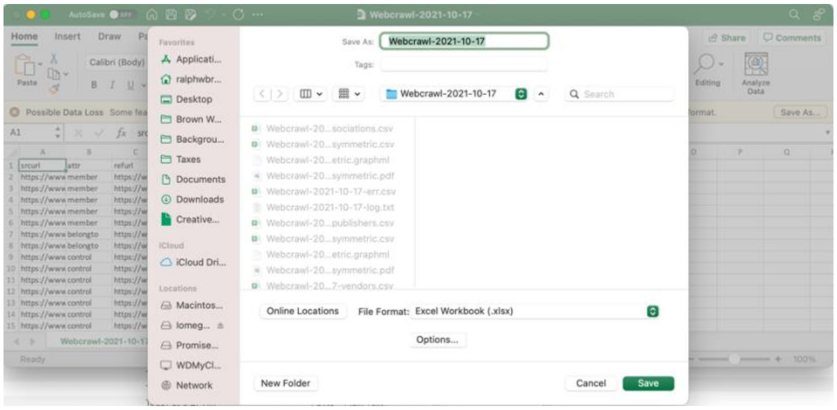
To create the Excel Pivot Table, open the file “Webcrawl-YYYY-MM-DD.csv” with Excel. It will notify you regarding Possible Data Loss.



Possible Data Loss Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format.

	A	B	C
1	srcurl	attr	refurl
2	https://www	member	https://www.digitalcontentnext.org/
3	https://www	member	https://www.trustx.org/
4	https://www	member	https://www.aenetworks.com/
5	https://www	member	https://www.bizjournals.com/
6	https://www	member	https://www.ap.org/
7	https://www	belongto	https://www.journallist.net/
8	https://www	belongto	https://www.iptc.org/
9	https://www	control	https://www.embeddedmetadata.org/
10	https://www	control	https://www.eventsml.org/
11	https://www	control	https://www.eventsml.com/
12	https://www	control	https://www.iptc.com/
13	https://www	control	https://www.newscode.com/
14	https://www	control	https://www.newscode.org/
15	https://www	control	https://www.newscodes.com/

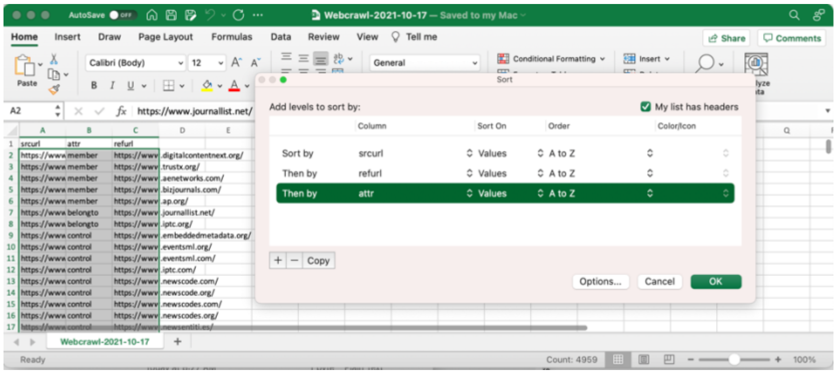
Save the file as an Excel Workbook “Webcrawl-YYYY-MM-DD.xlsx”.



Save As: Webcrawl-2021-10-17

File Format: Excel Workbook (.xlsx)

Select all the contents of this worksheet and custom sort the contents, first by srcurl, then by refurl, and then by attr.

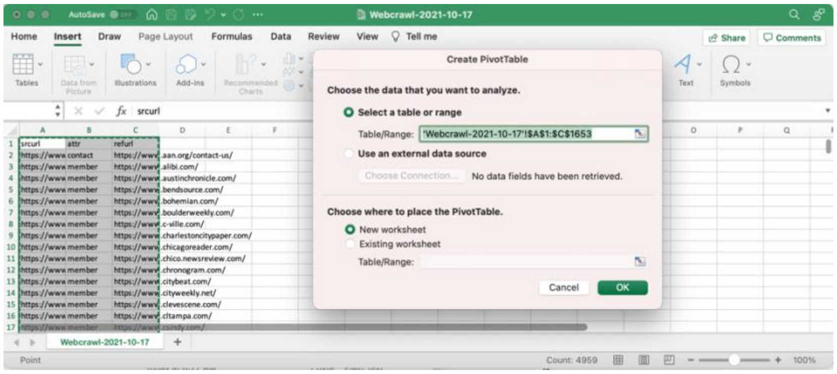


Add levels to sort by: ☑ My list has headers

	Column	Sort On	Order	Color/Icon
Sort by	srcurl	Values	A to Z	
Then by	refurl	Values	A to Z	
Then by	attr	Values	A to Z	

Count: 4959

Select all again and insert a PivotTable using selected range and placing it on a new worksheet.



Create PivotTable

Choose the data that you want to analyze.

● Select a table or range

Table/Range: 'Webcrawl-2021-10-17'!$A$1:$C$1653

○ Use an external data source

No data fields have been retrieved.

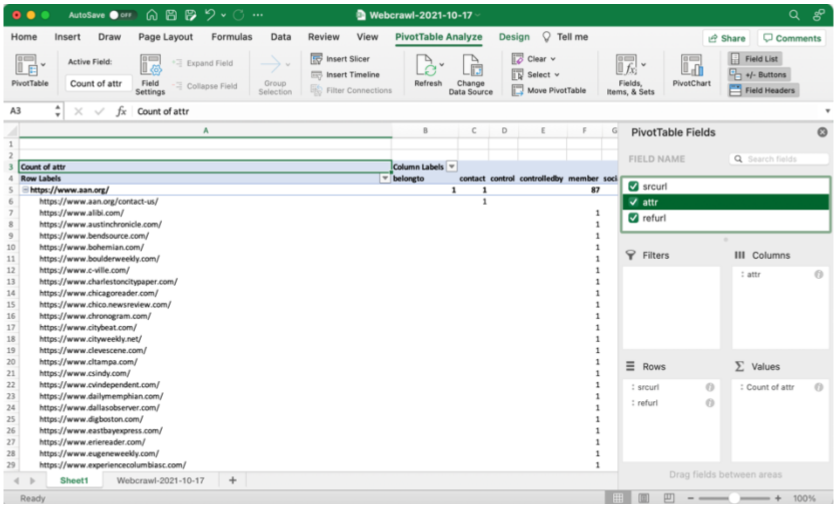
Choose where to place the PivotTable.

● New worksheet

○ Existing worksheet

Table/Range:

Add srcurl to the Rows of the PivotTable Fields and refurl below this. Add attr to the Columns of the PivotTable Fields and to the Sum Values.



A3 Count of attr

Count of attr	Column Labels					
Row Labels	belongto	contact	control	controlledby	member	soci
⊟ https://www.aan.org/	1	1			87	
https://www.aan.org/contact-us/		1				
https://www.alibi.com/					1	
https://www.austinchronicle.com/					1	
https://www.bendsource.com/					1	
https://www.bohemian.com/					1	
https://www.boulderweekly.com/					1	
https://www.c-ville.com/					1	
https://www.charlestoncitypaper.com/					1	
https://www.chicagoreader.com/					1	
https://www.chico.newsreview.com/					1	
https://www.chronogram.com/					1	
https://www.citybeat.com/					1	
https://www.cityweekly.net/					1	
https://www.clevescene.com/					1	
https://www.cltampa.com/					1	
https://www.csindy.com/					1	
https://www.cvindependent.com/					1	
https://www.dailymemphian.com/					1	
https://www.dallasobserver.com/					1	
https://www.digboston.com/					1	
https://www.eastbayexpress.com/					1	
https://www.eriereader.com/					1	
https://www.eugeneweekly.com/					1	
https://www.experiencecolumbiasc.com/					1	

PivotTable Fields

☑ srcurl
☑ attr
☑ refurl

Columns: attr

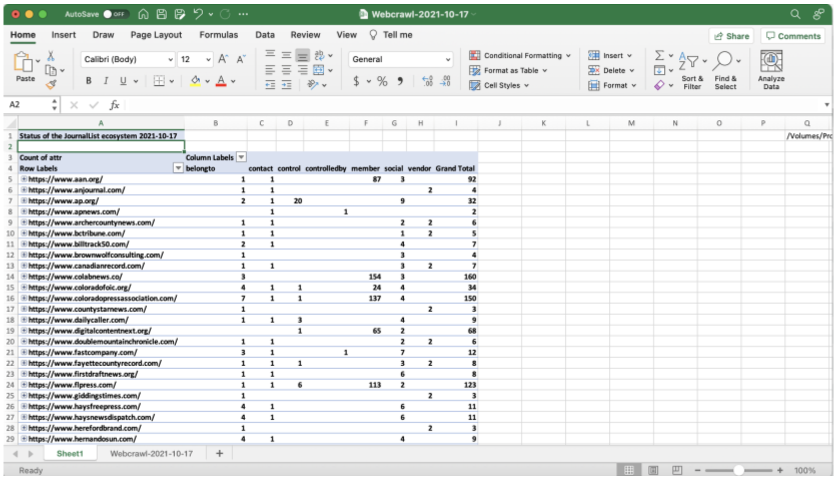
Rows: srcurl, refurl

Values: Count of attr

Collapse all the sub-rows by selecting cell A5 and “Collapse Field” under “PivotTable Analyze”. This will provide a summary of the different attributes used in each of the trust.txt files found.

Add a cell to capture the filename of the Excel workbook and create a title for the using the following formulas:

- Cell Q1
- Cell A1 =CONCAT("Status of the JournalList ecosystem ",MID(Q1,FIND("Webcrawl-",Q1)+9,10))

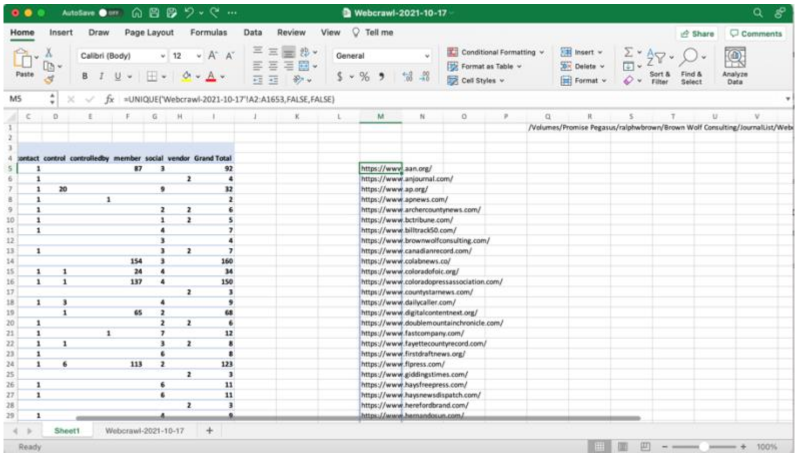


A2

	A	B	C	D	E	F	G	H	I
1	Status of the JournalList ecosystem 2021-10-17								
2									
3	Count of attr	Column Labels							
4	Row Labels	belongto	contact	control	controlledby	member	social	vendor	Grand Total
5	https://www.aan.org/	1	1			87	3		92
6	https://www.anjournal.com/	1	1					2	4
7	https://www.ap.org/	2	1	20			9		32
8	https://www.apnews.com/		1		1				2
9	https://www.archercountynews.com/	1	1				2	2	6
10	https://www.bctribune.com/	1	1				1	2	5
11	https://www.billtrack50.com/	2	1				4		7
12	https://www.brownwolfconsulting.com/	1					3		4
13	https://www.canadianrecord.com/	1	1				3	2	7
14	https://www.colabnews.co/	3				154	3		160
15	https://www.coloradofoic.org/	4	1	1		24	4		34
16	https://www.coloradopressassociation.com/	7	1	1		137	4		150
17	https://www.countystarnews.com/	1						2	3
18	https://www.dailycaller.com/	1	1	3			4		9
19	https://www.digitalcontentnext.org/			1		65	2		68
20	https://www.doublemountainchronicle.com/	1	1				2	2	6
21	https://www.fastcompany.com/	3	1		1		7		12
22	https://www.fayettecountyrecord.com/	1	1	1			3	2	8
23	https://www.firstdraftnews.org/	1	1				6		8
24	https://www.flpress.com/	1	1	6		113	2		123
25	https://www.giddingstimes.com/	1						2	3
26	https://www.haysfreepress.com/	4	1				6		11
27	https://www.haysnewsdispatch.com/	4	1				6		11
28	https://www.herefordbrand.com/	1						2	3
29	https://www.hernandosun.com/	4	1				4		9

Create a list of the organization URLs with that have published a trust.txt file by placing the following in cell M5 of the piviot table:

=UNIQUE('[Webcrawl-YYYY-MM-DD.xlsx]Webcrawl- YYYY-MM-DD!A2:A<last row>,FALSE,FALSE)

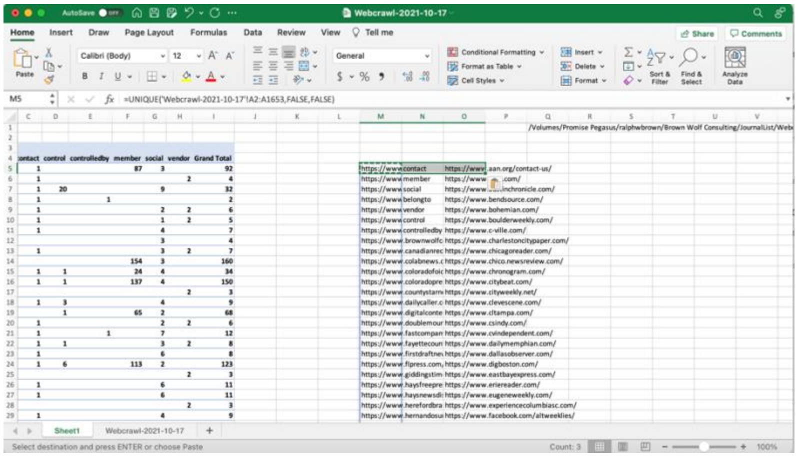


M5 =UNIQUE('Webcrawl-2021-10-17'!A2:A1653,FALSE,FALSE)

/Volumes/Promise Pegasus/ralphwbrown/Brown Wolf Consulting/JournalList/Web

	M
5	https://www.aan.org/
6	https://www.anjournal.com/
7	https://www.ap.org/
8	https://www.apnews.com/
9	https://www.archercountynews.com/
10	https://www.bctribune.com/
11	https://www.billtrack50.com/
12	https://www.brownwolfconsulting.com/
13	https://www.canadianrecord.com/
14	https://www.colabnews.co/
15	https://www.coloradofoic.org/
16	https://www.coloradopressassociation.com/
17	https://www.countystarnews.com/
18	https://www.dailycaller.com/
19	https://www.digitalcontentnext.org/
20	https://www.doublemountainchronicle.com/
21	https://www.fastcompany.com/
22	https://www.fayettecountyrecord.com/
23	https://www.firstdraftnews.org/
24	https://www.flpress.com/
25	https://www.giddingstimes.com/
26	https://www.haysfreepress.com/
27	https://www.haysnewsdispatch.com/
28	https://www.herefordbrand.com/
29	https://www.hernandosun.com/

Copy this cell to the two adjacent cells



M5 =UNIQUE('Webcrawl-2021-10-17'!A2:A1653,FALSE,FALSE)

/Volumes/Promise Pegasus/ralphwbrown/Brown Wolf Consulting/JournalList/Web

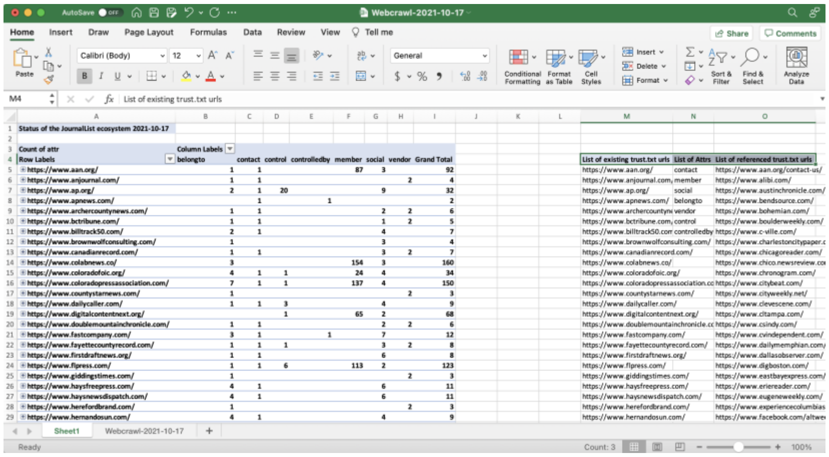
	M	N	O
5	https://www	contact	https://www.aan.org/contact-us/
6	https://www	member	https://www [illegible] .com/
7	https://www	social	https://www [illegible] inchronicle.com/
8	https://www	belongto	https://www.bendsource.com/
9	https://www	vendor	https://www.bohemian.com/
10	https://www	control	https://www.boulderweekly.com/
11	https://www	controlledby	https://www.c-ville.com/
12	https://www	.brownwolfc	https://www.charlestoncitypaper.com/
13	https://www	.canadianrec	https://www.chicagoreader.com/
14	https://www	.colabnews.c	https://www.chico.newsreview.com/
15	https://www	.coloradofoic	https://www.chronogram.com/
16	https://www	.coloradopre	https://www.citybeat.com/
17	https://www	.countystarn	https://www.cityweekly.net/
18	https://www	.dailycaller.c	https://www.clevescene.com/
19	https://www	.digitalconte	https://www.cltampa.com/
20	https://www	.doublemour	https://www.csindy.com/
21	https://www	.fastcompan	https://www.cvindependent.com/
22	https://www	.fayettecoun	https://www.dailymemphian.com/
23	https://www	.firstdraftne	https://www.dallasobserver.com/
24	https://www	.flpress.com,	https://www.digboston.com/
25	https://www	.giddingstim	https://www.eastbayexpress.com/
26	https://www	.haysfreepre	https://www.eriereader.com/
27	https://www	.haysnewsdi	https://www.eugeneweekly.com/
28	https://www	.herefordbra	https://www.experiencecolumbiasc.com/
29	https://www	.hernandosu	https://www.facebook.com/altweeklies/

Select destination and press ENTER or choose Paste

Count: 3

Put headers above these in cells M4, N4, and O4 using the following:

List of existing trust.txt urls	List of Attrs	List of referenced trust.txt urls

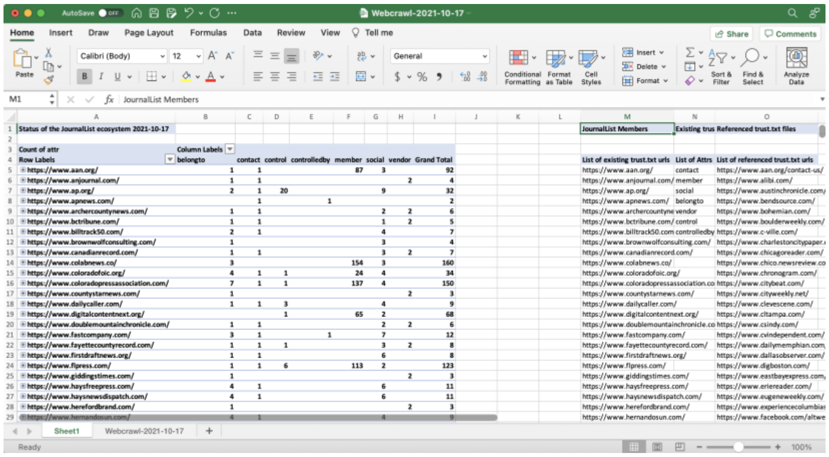


M4 | *fx* List of existing trust.txt urls

	A	B	C	D	E	F	G	H	I	M	N	O
1	Status of the JournalList ecosystem 2021-10-17											
2												
3	Count of attr	Column Labels										
4	Row Labels	belongto	contact	control	controlledby	member	social	vendor	Grand Total	List of existing trust.txt urls	List of Attrs	List of referenced trust.txt urls
5	https://www.aan.org/	1	1			87	3		92	https://www.aan.org/	contact	https://www.aan.org/contact-us/
6	https://www.anjournal.com/	1	1					2	4	https://www.anjournal.com,	member	https://www.alibi.com/
7	https://www.ap.org/	2	1	20			9		32	https://www.ap.org/	social	https://www.austinchronicle.com/
8	https://www.apnews.com/		1		1				2	https://www.apnews.com/	belongto	https://www.bendsource.com/
9	https://www.archercountynews.com/	1	1				2	2	6	https://www.archercountyn	vendor	https://www.bohemian.com/
10	https://www.bctribune.com/	1	1				1	2	5	https://www.bctribune.com,	control	https://www.boulderweekly.com/
11	https://www.billtrack50.com/	2	1				4		7	https://www.billtrack50.com	controlledby	https://www.c-ville.com/
12	https://www.brownwolfconsulting.com/	1					3		4	https://www.brownwolfconsulting.com/		https://www.charlestoncitypaper.c
13	https://www.canadianrecord.com/	1	1				3	2	7	https://www.canadianrecord.com/		https://www.chicagoreader.com/
14	https://www.colabnews.co/	3				154	3		160	https://www.colabnews.co/		https://www.chico.newsreview.co
15	https://www.coloradofoic.org/	4	1	1		24	4		34	https://www.coloradofoic.org/		https://www.chronogram.com/
16	https://www.coloradopressassociation.com/	7	1	1		137	4		150	https://www.coloradopressassociation.co		https://www.citybeat.com/
17	https://www.countystarnews.com/	1						2	3	https://www.countystarnews.com/		https://www.cityweekly.net/
18	https://www.dailycaller.com/	1	1	3			4		9	https://www.dailycaller.com/		https://www.clevescene.com/
19	https://www.digitalcontentnext.org/			1		65	2		68	https://www.digitalcontentnext.org/		https://www.cltampa.com/
20	https://www.doublemountainchronicle.com/	1	1				2	2	6	https://www.doublemountainchronicle.cc		https://www.csindy.com/
21	https://www.fastcompany.com/	3	1		1		7		12	https://www.fastcompany.com/		https://www.cvindependent.com/
22	https://www.fayettecountyrecord.com/	1	1	1			3	2	8	https://www.fayettecountyrecord.com/		https://www.dailymemphian.com/
23	https://www.firstdraftnews.org/	1	1				6		8	https://www.firstdraftnews.org/		https://www.dallasobserver.com/
24	https://www.flpress.com/	1	1	6		113	2		123	https://www.flpress.com/		https://www.digboston.com/
25	https://www.giddingstimes.com/	1						2	3	https://www.giddingstimes.com/		https://www.eastbayexpress.com/
26	https://www.haysfreepress.com/	4	1				6		11	https://www.haysfreepress.com/		https://www.eriereader.com/
27	https://www.haysnewsdispatch.com/	4	1				6		11	https://www.haysnewsdispatch.com/		https://www.eugeneweekly.com/
28	https://www.herefordbrand.com/	1						2	3	https://www.herefordbrand.com/		https://www.experiencecolumbias
29	https://www.hernandosun.com/	4	1				4		9	https://www.hernandosun.com/		https://www.facebook.com/altwe

Above this in cell M1, N1, and O1 create the following headers:

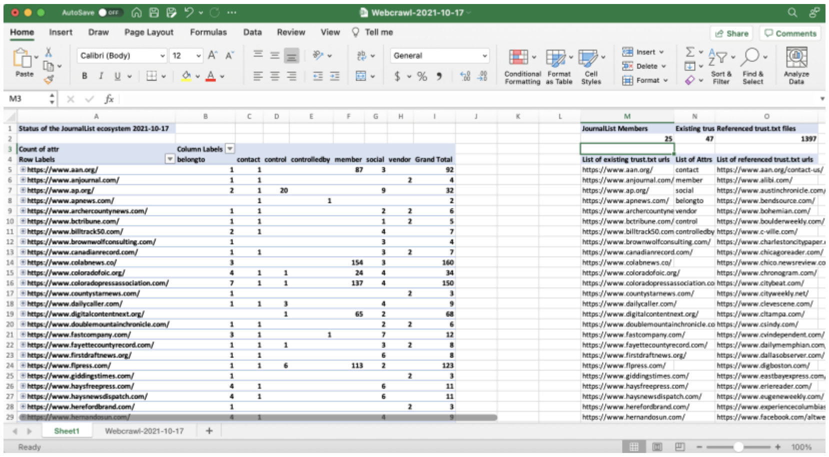
JournalList Members	Existing trust.txt files	Referenced trust.txt files



M1 | *fx* JournalList Members

	A	B	C	D	E	F	G	H	I	M	N	O
1	Status of the JournalList ecosystem 2021-10-17									JournalList Members	Existing trus	Referenced trust.txt files
2												
3	Count of attr	Column Labels										
4	Row Labels	belongto	contact	control	controlledby	member	social	vendor	Grand Total	List of existing trust.txt urls	List of Attrs	List of referenced trust.txt urls
5	https://www.aan.org/	1	1			87	3		92	https://www.aan.org/	contact	https://www.aan.org/contact-us/
6	https://www.anjournal.com/	1	1					2	4	https://www.anjournal.com/	member	https://www.alibi.com/
7	https://www.ap.org/	2	1	20			9		32	https://www.ap.org/	social	https://www.austinchronicle.com/
8	https://www.apnews.com/		1		1				2	https://www.apnews.com/	belongto	https://www.bendsource.com/
9	https://www.archercountynews.com/	1	1				2	2	6	https://www.archercountyne	vendor	https://www.bohemian.com/
10	https://www.bctribune.com/	1	1				1	2	5	https://www.bctribune.com/	control	https://www.boulderweekly.com/
11	https://www.billtrack50.com/	2	1				4		7	https://www.billtrack50.com	controlledby	https://www.c-ville.com/
12	https://www.brownwolfconsulting.com/	1					3		4	https://www.brownwolfconsulting.com/		https://www.charlestoncitypaper.c
13	https://www.canadianrecord.com/	1	1				3	2	7	https://www.canadianrecord.com/		https://www.chicagoreader.com/
14	https://www.colabnews.co/	3				154	3		160	https://www.colabnews.co/		https://www.chico.newsreview.co
15	https://www.coloradofoic.org/	4	1	1		24	4		34	https://www.coloradofoic.org/		https://www.chronogram.com/
16	https://www.coloradopressassociation.com/	7	1	1		137	4		150	https://www.coloradopressassociation.co		https://www.citybeat.com/
17	https://www.countystarnews.com/	1						2	3	https://www.countystarnews.com/		https://www.cityweekly.net/
18	https://www.dailycaller.com/	1	1	3			4		9	https://www.dailycaller.com/		https://www.clevescene.com/
19	https://www.digitalcontentnext.org/			1		65	2		68	https://www.digitalcontentnext.org/		https://www.cltampa.com/
20	https://www.doublemountainchronicle.com/	1	1				2	2	6	https://www.doublemountainchronicle.co		https://www.csindy.com/
21	https://www.fastcompany.com/	3	1		1		7		12	https://www.fastcompany.com/		https://www.cvindependent.com/
22	https://www.fayettecountyrecord.com/	1	1	1			3	2	8	https://www.fayettecountyrecord.com/		https://www.dailymemphian.com/
23	https://www.firstdraftnews.org/	1	1				6		8	https://www.firstdraftnews.org/		https://www.dallasobserver.com/
24	https://www.flpress.com/	1	1	6		113	2		123	https://www.flpress.com/		https://www.digboston.com/
25	https://www.giddingstimes.com/	1						2	3	https://www.giddingstimes.com/		https://www.eastbayexpress.com/
26	https://www.haysfreepress.com/	4	1				6		11	https://www.haysfreepress.com/		https://www.eriereader.com/
27	https://www.haysnewsdispatch.com/	4	1				6		11	https://www.haysnewsdispatch.com/		https://www.eugeneweekly.com/
28	https://www.herefordbrand.com/	1						2	3	https://www.herefordbrand.com/		https://www.experiencecolumbias
29	https://www.hernandosun.com/	4	1				4		9	https://www.hernandosun.com/		https://www.facebook.com/altwe

The number of JournalList members can be selected from the pivot table, number of existing published trust.txt files is obtained by =COUNTA(M5#), the number of referenced, but not published trust.txt files is obtained by =COUNTA(O5#). The list of attributes used across all the existing trust.txt files is show in Column N.



Status of the JournalList ecosystem 2021-10-17

Count of attr	Column Labels							
Row Labels	**belongto**	**contact**	**control**	**controlledby**	**member**	**social**	**vendor**	**Grand Total**
https://www.aan.org/	1	1			87	3		92
https://www.anjournal.com/	1	1					2	4
https://www.ap.org/	2	1	20			9		32
https://www.apnews.com/		1		1				2
https://www.archercountynews.com/	1	1				2	2	6
https://www.bctribune.com/	1	1				1	2	5
https://www.billtrack50.com/	2	1				4		7
https://www.brownwolfconsulting.com/	1					3		4
https://www.canadianrecord.com/	1	1				3	2	7
https://www.colabnews.co/	3				154	3		160
https://www.coloradofoic.org/	4	1	1		24	4		34
https://www.coloradopressassociation.com/	7	1	1		137	4		150
https://www.countystarnews.com/	1						2	3
https://www.dailycaller.com/	1	1	3			4		9
https://www.digitalcontentnext.org/			1		65	2		68
https://www.doublemountainchronicle.com/	1	1				2	2	6
https://www.fastcompany.com/	3	1		1		7		12
https://www.fayettecountyrecord.com/	1	1	1			3	2	8
https://www.firstdraftnews.org/	1	1				6		8
https://www.flpress.com/	1	1	6		113	2		123
https://www.giddingstimes.com/	1						2	3
https://www.haysfreepress.com/	4	1				6		11
https://www.haysnewsdispatch.com/	4	1				6		11
https://www.herefordbrand.com/	1						2	3
https://www.hernandosun.com/	4	1				4		9

JournalList Members	Existing trus	Referenced trust.txt files
25	47	1397

List of existing trust.txt urls	List of Attrs	List of referenced trust.txt urls
https://www.aan.org/	contact	https://www.aan.org/contact-us/
https://www.anjournal.com/	member	https://www.alibi.com/
https://www.ap.org/	social	https://www.austinchronicle.com/
https://www.apnews.com/	belongto	https://www.bendsource.com/
https://www.archercountyne	vendor	https://www.bohemian.com/
https://www.bctribune.com/	control	https://www.boulderweekly.com/
https://www.billtrack50.com	controlledby	https://www.c-ville.com/
https://www.brownwolfconsulting.com/		https://www.charlestoncitypaper.
https://www.canadianrecord.com/		https://www.chicagoreader.com/
https://www.colabnews.co/		https://www.chico.newsreview.co
https://www.coloradofoic.org/		https://www.chronogram.com/
https://www.coloradopressassociation.co		https://www.citybeat.com/
https://www.countystarnews.com/		https://www.cityweekly.net/
https://www.dailycaller.com/		https://www.clevescene.com/
https://www.digitalcontentnext.org/		https://www.cltampa.com/
https://www.doublemountainchronicle.co		https://www.csindy.com/
https://www.fastcompany.com/		https://www.cvindependent.com/
https://www.fayettecountyrecord.com/		https://www.dailymemphian.com,
https://www.firstdraftnews.org/		https://www.dallasobserver.com/
https://www.flpress.com/		https://www.digboston.com/
https://www.giddingstimes.com/		https://www.eastbayexpress.com,
https://www.haysfreepress.com/		https://www.eriereader.com/
https://www.haysnewsdispatch.com/		https://www.eugeneweekly.com/
https://www.herefordbrand.com/		https://www.experiencecolumbias
https://www.hernandosun.com/		https://www.facebook.com/altwe