

Using Wikipedia to Make the Digitized Newspapers of the National Digital Newspaper Program More Discoverable

Preliminary Findings

Donald Taylor

Wikipedian-in-Residence

Maryland Historic Newspapers Project

University of Maryland Libraries

Not your usual Wikipedian-in-Residence

Who is using *Chronicling America*?

linkypedia

linkypedia.inkdroid.org

Google

 linkypedia_α

*linkypedia helps you **explore** how your web content is used on wikipedia, **extract** crowd-sourced information about your web resources, and **connect** with wikipedians who are using your stuff. More about the rationale for linkypedia can be found [here](#).*

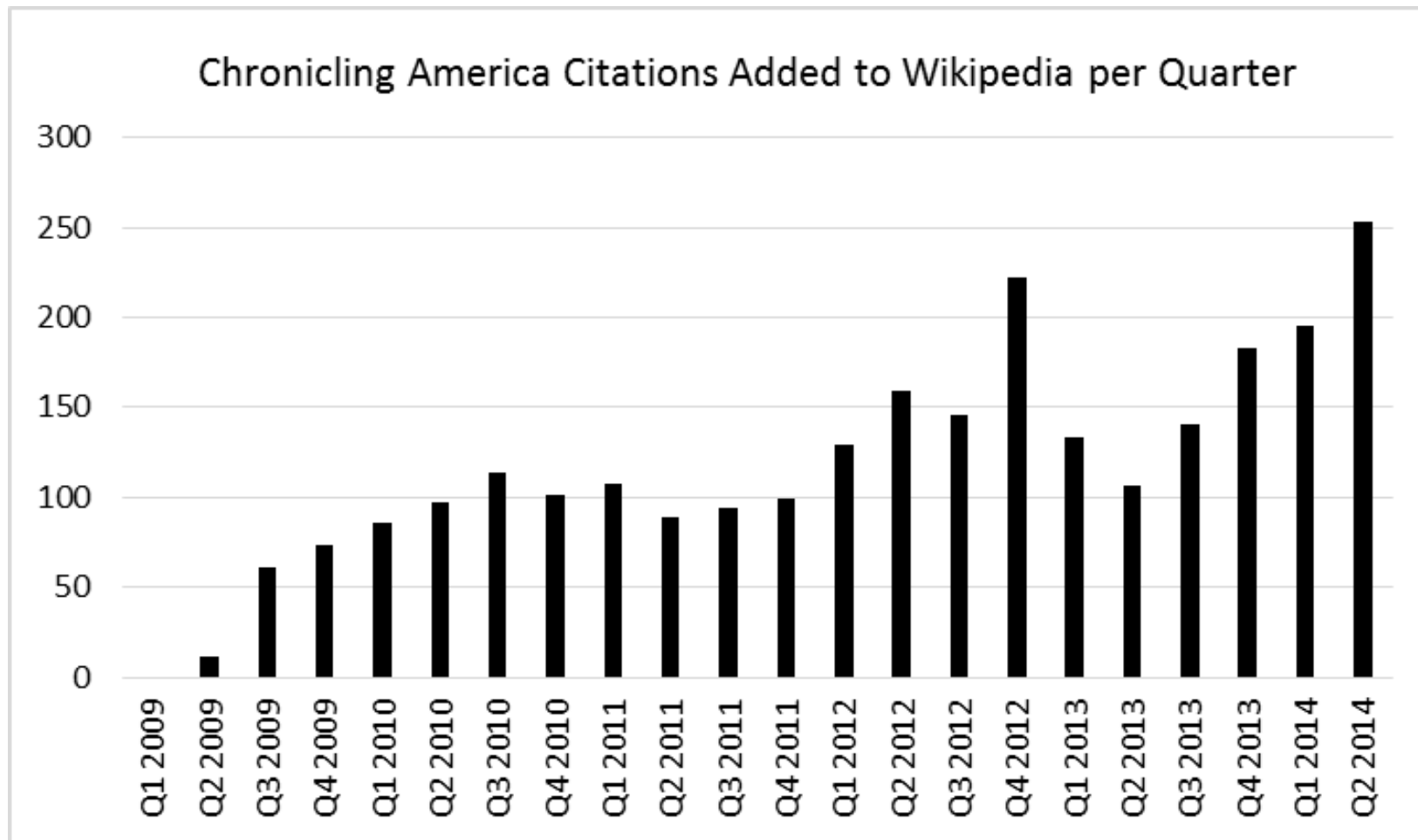
recently found: [Talk:Australia national soccer team results](#) referencing [National Library of Australia](#)

-  National Library of Australia (126,090)
-  DNB, Katalog der Deutschen Nationalbibliothek (92,884)
-  bioguide.congress.gov (29,361)
-  Natural History Museum (27,022)
-  WorldCat.org: The World's Largest Library Catalog (25,740)
-  Australia Trove (24,555)
-  Oxford Dictionary of National Biography (19,569)
-  OCLC Research Linked Authority File (2.0) (16,023)
-  American Memory from the Library of Congress (12,966)
-  Papers Past (10,076)
-  HathiTrust Digital Library (7,752)
-  THOMAS (Library of Congress) (6,758)
-  Biodiversity Heritage Library (6,705)
-  National Archives and Records Administration (6,140)
-  The Online Books Page (5,779)
-  Chronicling America (5,538)
-  FDSys - US Government Printing Office (5,044)
-  World Digital Library (5,014)
-  Library of Congress Prints & Photographs Online Catalog (4,678)

Fork me on GitHub

Citation Distribution	
Number of citations	Number of articles
1	1248
2	263
3-5	167
6-10	57
11-20	21
21-50	17
51-76	4
Total	1,777

How many citations do editors make	
Number of citations	Editors
1	344
2	104
3-5	75
6-10	32
11-20	16
21-50	10
51-100	5
100-130	3
Total	589



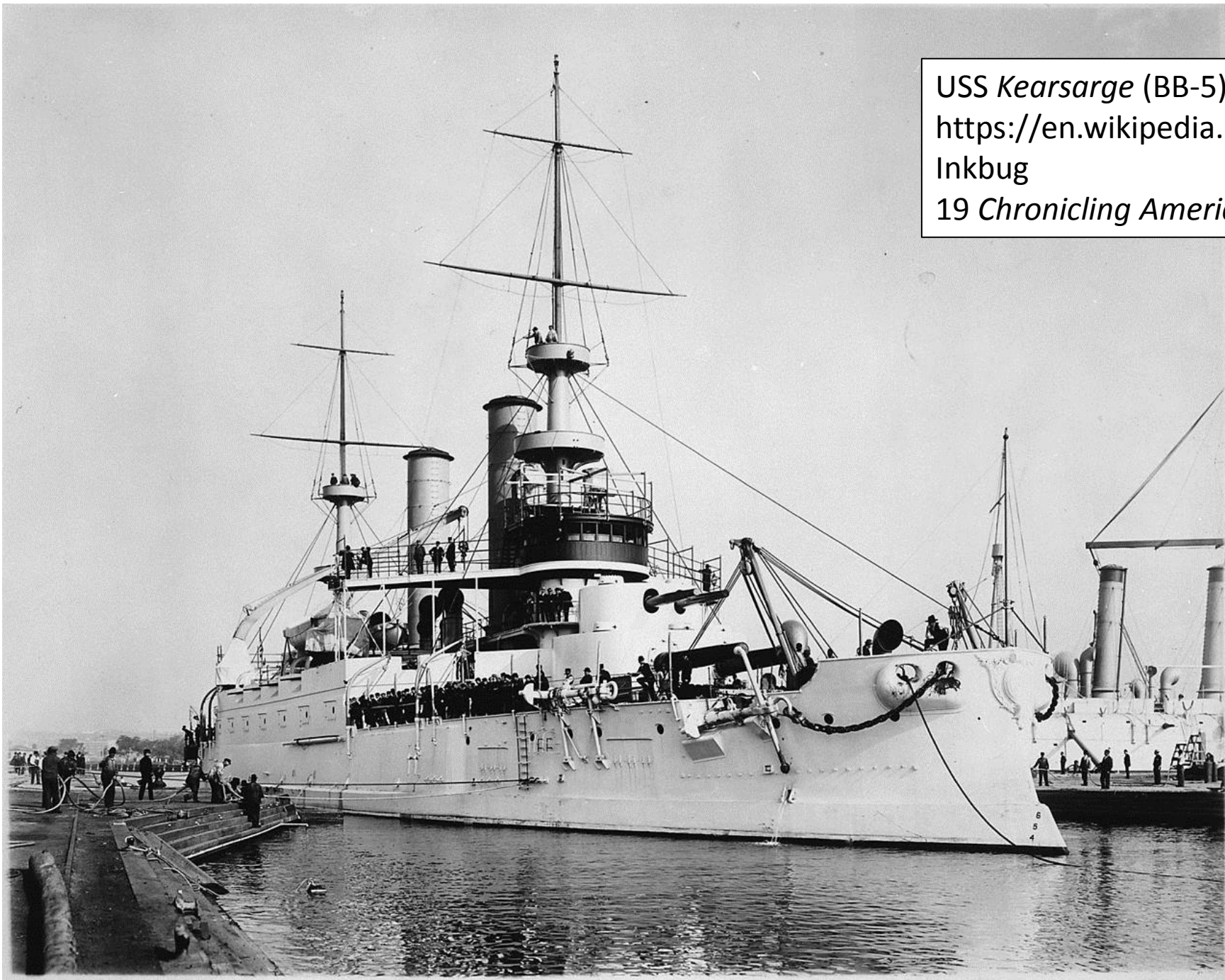
For the last year the rate of Chronicling America citation on Wikipedia has been increasing by about 36 per quarter.

Over the last five quarters the growth rate of Chronicling America citation has increased by 150 percent.

What are they using *Chronicling America* for?

WP:NOR

In inclusionism vs. deletionism
newspapers = notability



USS *Kearsarge* (BB-5) (keel laid: 1896, scrapped: 1955)
[https://en.wikipedia.org/wiki/USS_Kearsarge_\(BB-5\)](https://en.wikipedia.org/wiki/USS_Kearsarge_(BB-5))
Inkbug
19 *Chronicling America* citations of 62 total



Belle Gunness (November 11, 1859 – April 28, 1908)

https://en.wikipedia.org/wiki/Belle_Gunness

74.83.126.88

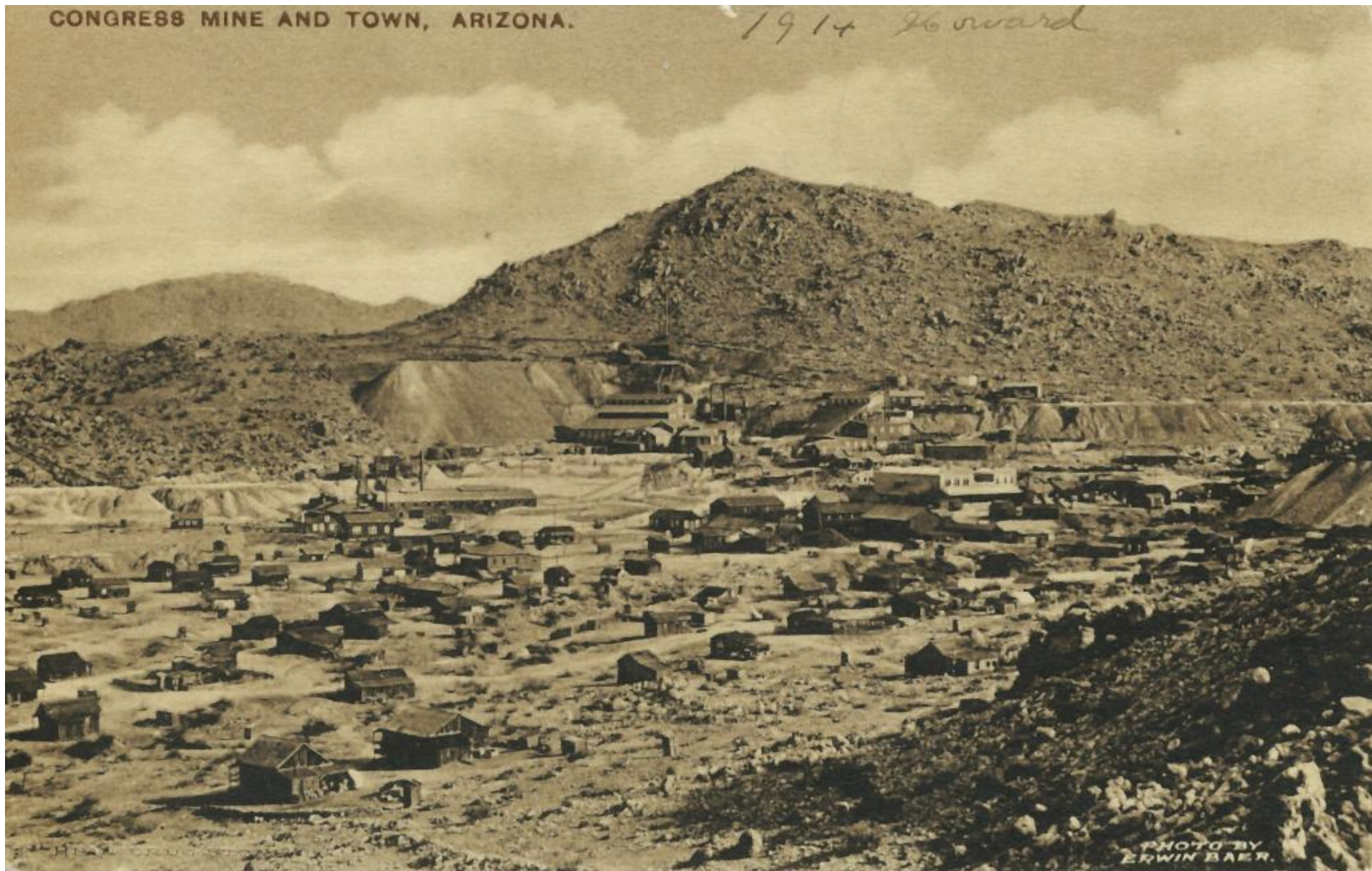
27 Chronicling America citations of 44 total



Adele Ritchie (December 21, 1874 – April 24, 1930)
12 *Chronicling America* citations of 31 total



McCants Stewart (July 11, 1877 – April 14, 1919)
16 *Chronicling America* citations of 25 total



Congress Mine – 48 Chronicling America citations from seven different newspapers

What is to be done?

The Human Factor

50 percent of NDNP awardee respondents have not used Wikipedia because it simply did not occur to them.

Other reasons:

- * Concerns about the reputation, integrity, or reliability of Wikipedia
- * Concerns over conflict of interest in editing Wikipedia
- * Our institution does not have the project resources
- * Lack of relevant expertise

Need to generate institutional buy-in, capability, enthusiasm

Wikipedia needs to be in your strategic plan

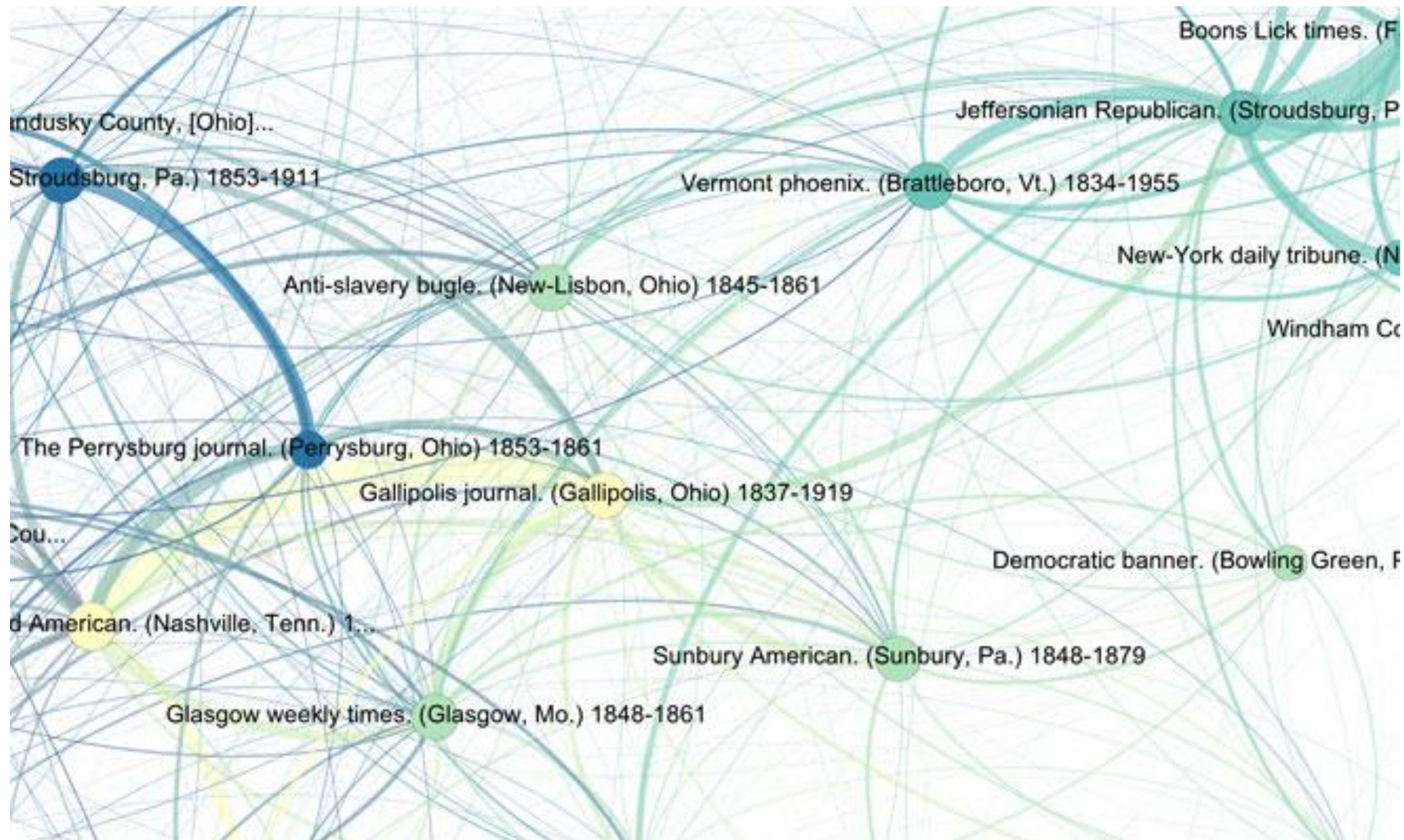
Tools

Chronicling America is Big Data

- * 8 million pages
- * If you read *War and Peace* in a month, it would take you 600 years to read *Chronicling America*
- * Viral Texts output can't even be opened on a PC
- * Tools are the only means
- * Viral Texts found thousands of previously unknown articles

Viral Texts

Ryan Cordell, Elizabeth Maddock Dillon and David Smith
Northeastern University's NULab for Texts, Maps, and Networks



Recurrence is a signal

Named Entity Recognition

About 895,000 results (0.32 seconds)

Official Kitsap County Web Site

www.kitsapgov.com/ ▾ Kitsap County ▾

Kitsap County Government. Elected Officials, Departments, programs, budgets, etc. Demographic and statistical information, Parks and Recreation. Visit Kitsap ...

[Parcel Search](#) - [Kitsap County Assessor](#) - [Jobs](#) - [Departments](#)

Kitsap County Assessor - Kitsap County Government

www.kitsapgov.com/assr/ ▾ Kitsap County ▾

6 days ago - Welcome to the Kitsap County Assessor's Office website!

News for Kitsap County



Teen held on \$1 million bail in death of Wash. girl, 6

CBS News - 5 days ago

Prosecutors in **Kitsap County**, Wash., filed court documents Monday saying they had sufficient evidence to hold Gabriel Gaeta, 17, in the ...

Wash. teen, 17, raped and murdered 6-year-old neighbor ...

New York Daily News - 5 days ago

Charges in slaying of Bremerton girl delayed for mental ...

The Seattle Times (blog) - 5 days ago

More news for **Kitsap County**

Kitsap County, Washington - Wikipedia, the free encyclopedia

en.wikipedia.org/wiki/Kitsap_County,_Washington ▾ Wikipedia ▾

Kitsap County is located in the State of Washington. As of the 2010 census, its population was 251,133. Its county seat is Port Orchard, and its largest city is ...

[History](#) - [Geography](#) - [Demographics](#) - [Politics](#)

Kitsap Peninsula located Across from Seattle,WA - Home of ...

www.visitkitsap.com/ ▾

National Park Service Awards County/VKP Grant to Develop Water Trails Management Program. The National Park Service has awarded **Kitsap County** a grant ...

Kitsap Sun

www.kitsapsun.com/ ▾ Kitsap Sun ▾

Current news for **Kitsap**, WA and its surrounding communities brought to you by the ...

More racks are available for installation in other parts of the **county**.



Kitsap County

US County in Washington

Kitsap County is located in the State of Washington. As of the 2010 census, its population was 251,133. Its county seat is Port Orchard, and its largest city is Bremerton. [Wikipedia](#)

Area: 566 sq miles (1,466 km²)

Founded: January 16, 1857

County seat: [Port Orchard](#)

Population: 254,991 (2012)

Unemployment rate: 5.5% (Apr 2014)

Points of interest

[View 15+ more](#)



USS Hornet



Point No Point



Puget Sound Naval Shi...



Tillicum Village



Hood Canal

Colleges and Universities



Olympic College
Bremerton



Bainbridge Graduate Institute
Bainbridge Island



Northwest College of Art
Poulsbo



Everest College-B...
Bremerton

SuggestBot

Articles you might like to edit, from SuggestBot [[edit source](#)]

SuggestBot predicts that you will enjoy editing some of these articles. Have fun!

Add sources

[Rodney King](#)
[Military history of the Soviet Union](#)
[Chandra Wickramasinghe](#)
[United States Air Forces in Europe - Air Forces Africa](#)
[The Longest Time](#)
[Louis Maurice Adolphe Linant de Bellefonds](#)

Cleanup

[Captain America in other media](#)
[Liquidity premium](#)
[Mashhour Ahmed Mashhour](#)

Expand

[Société d'Études du Canal de Suez](#)
['Urabi Revolt](#)
[Geostrategy](#)

Unencyclopaedic

[Helmuth von Moltke the Elder](#)
[On War](#)
[Modern history](#)

Wikify

[Operation Telescope](#)
[History of Russia \(1892–1917\)](#)
[Dissolution of the Ottoman Empire](#)

Orphan

[Herman Beukema](#)
[Stolypin car](#)
[Joseph Inginac](#)

Merge

[History of Serbia \(1804–1918\)](#)
[Planetary protection](#)
[Treaty of Balta Liman](#)

Stub

[Kassassin](#)
[South Country Central School District](#)
[CD79](#)
[Michael Blumlein](#)
[The Rose of Mooncoin](#)
[Ex tunc](#)

SuggestBot picks articles in a number of ways based on other articles you've edited, including straight text similarity, following wikilinks, and matching your editing patterns against those of other Wikipedians. It tries to recommend only articles that other Wikipedians have marked as needing work. Your contributions make Wikipedia better — thanks for helping.

If you have *feedback* on how to make SuggestBot better, please tell me on [SuggestBot's talk page](#). Thanks from [Nettrom \(talk\)](#),

SuggestBot's caretaker. -- [SuggestBot \(talk\)](#) 13:32, 26 June 2014 (UTC)

You're not going to build one thing and have it work –
it's iterative.

1. What doesn't work about Wikipedia for you?
2. What doesn't work about Chronicling America for you?
3. What ideas do you have that could make them better?

@donaldtaylorii dwtii@umd.edu