Rural Digital Footprints

Information Gaps and Digital Representation

Dr. Jacob Thebault-Spieker
Assistant Professor, Information School
School of Computer, Data, and Information Science
University of Wisconsin – Madison

Rural Digital Footprints, Information Gaps, and Digital Representation

Three Main Goals

1. Introduce Digital Footprints and Information Gaps
2. One example of my work that helps us understand causes
3. Connect to conversations about AI and related technologies

Rural Digital Footprints, Information Gaps, and Digital Representation

Three Main Goals

1. **Introduce Digital Footprints and Information Gaps**
2. One example of my work that helps us understand causes
3. Connect to conversations about AI and related technologies

JANUARY 1994 • $1.95 PRINTED IN U.S.A.

Popular Mechanics

1994 DESIGN & ENGINEERING AWARDS
Best Products, Designs And Ideas Of The Year

UNDERSTANDING THE INFORMATION SUPERHIGHWAY

How You'll Shop, Bank, Learn, Be Entertained And More Via Interactive TV

THE GREATEST SOCIAL REVOLUTION SINCE THE AUTOMOBILE

WIKIPEDIA
The Free Encyclopedia

Web 2.0

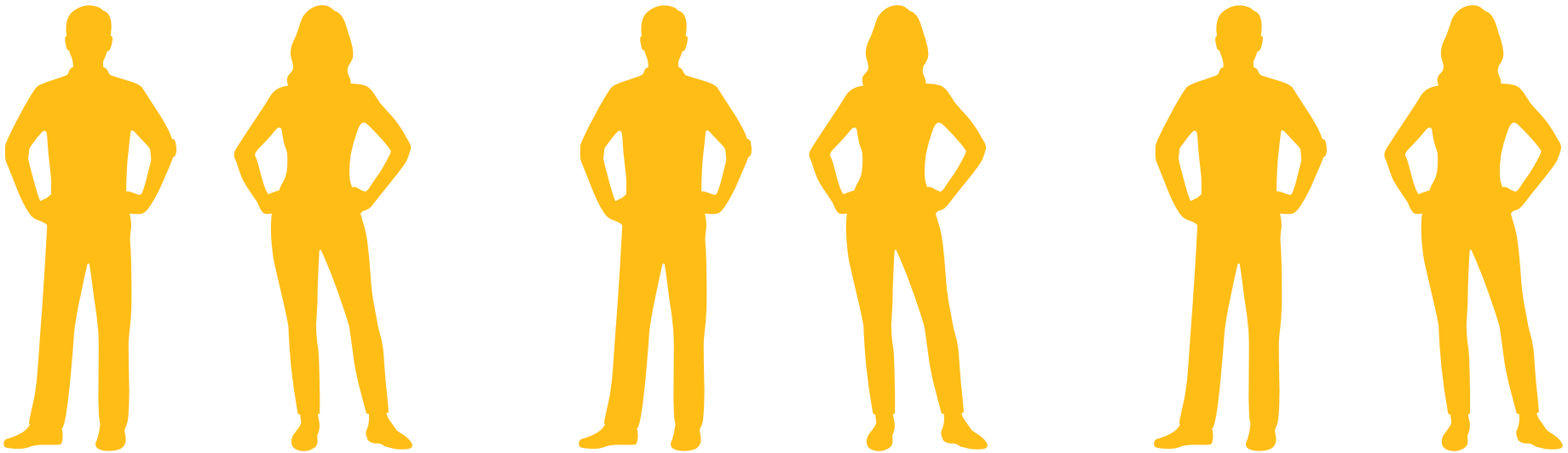
Article Talk

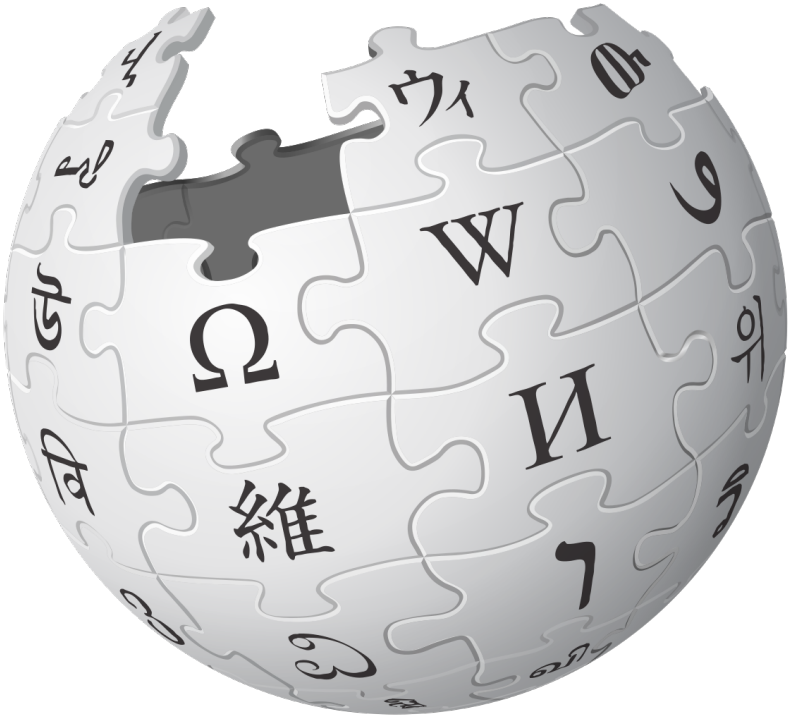
From Wikipedia, the free encyclopedia

Web 2.0 (also known as **participative** (or **participatory**)[1] **web** and **social web**)[2] refers to websites that emphasize user-generated content, ease of use, participatory culture and interoperability (i.e., compatibility with other products, systems, and devices) for end users.

Website	Domain name
Google Search	google.com
YouTube	youtube.com
Facebook	facebook.com
Instagram	instagram.com
X	twitter.com
Wikipedia	wikipedia.org
Yahoo!	yahoo.com
WhatsApp	whatsapp.com
Amazon	amazon.com
Reddit	reddit.com

"The social web" is slightly euphemistic




維

Ok, but "footprints"? "Gaps"?

WIKIPEDIA
The Free Encyclopedia

Search Wikipedia | Search

Black Earth, Wisconsin

20 languages

Article | Talk | Read | Edit | View history | Tools

From Wikipedia, the free encyclopedia

Coordinates: 43°8′12″N 89°44′51″W

Contents hide

- (Top)
- History
- Geography
- Demographics
  - 2010 census
  - 2000 census
- References
- External links

Black Earth is a village in Dane County, Wisconsin, United States. The population was 1,338 at the 2010 census. The village is located within the Town of Black Earth. It is part of the Madison Metropolitan Statistical Area.

Black Earth, Wisconsin
Village


History [edit]

In 1851, the town board renamed the town Farmersville, but the name Black Earth was readopted in 1857.[6]

WIKIPEDIA
The Free Encyclopedia

Search Wikipedia | Search

Madison, Wisconsin

113 languages

Article | Talk | Read | Edit | View history | Tools

From Wikipedia, the free encyclopedia

Coordinates: 43°04′29″N 89°23′03″W

Contents hide

- (Top)
- History
- Geography
- Demographics
- Economy
- Arts and culture
- Sports
- Parks and recreation
- Government
- Education

This article is about the city. For the former town, see Madison (town), Wisconsin.

Madison is the capital city of the state of Wisconsin and the county seat of and largest city in Dane County. As of the 2020 census, the population was 269,840, making it the second most populous city in Wisconsin after Milwaukee, and the 80th most populous in the United States. Madison is named for American Founding Father and President James Madison.

Located on an isthmus and lands surrounding five lakes—Lake Mendota, Lake Monona, Lake Wingra, Lake Kegonsa and Lake Waubesa—the city is home to the University of Wisconsin–Madison, the Wisconsin State Capitol, the Overture Center for the Arts, and the Henry Vilas Zoo. Madison is home to an extensive network of parks and bike trails; it has the most parks and playgrounds per capita of any of the 100 largest U.S. cities and is one of five communities to have received a "Platinum Bicycle

Madison
State capital city

View of the Madison Isthmus and Lake Mendota from Picnic Point

Madison, Wisconsin

113 languages

From Wikipedia, the free encyclopedia

Coordinates: 43°04′29″N 89°23′03″W

This article is about the city. For the former town, see Madison (town), Wisconsin.

Madison is the capital city of the state of Wisconsin and the county seat of and largest city in Dane County. As of the 2020 census, the population was 269,840, making it the second most populous city in Wisconsin after Milwaukee, and the 80th most populous in the United States. Madison is named for American Founding Father and President James Madison.

Located on an isthmus and lands surrounding five lakes—Lake Mendota, Lake Monona, Lake Wingra, Lake Kegonsa and Lake Waubesa—the city is home to the University of Wisconsin–Madison, the Wisconsin State Capitol, the Overture Center for the Arts, and the Henry Vilas Zoo. Madison is home to an extensive network of parks and bike trails; it has the most parks and playgrounds per capita of any of the 100 largest U.S. cities and is one of five communities to have received a "Platinum Bicycle Friendly Community" rating from the League of American Bicyclists.[6][7] Madison is also home to nine National Historic Landmarks, including several buildings designed by architect Frank Lloyd Wright, such as his 1937 Jacobs I House, which is a UNESCO World Heritage Site.[8]

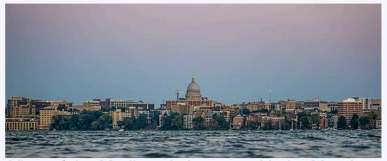
Residents of Madison are known as Madisonians.[9] Madison has long been a center for progressive political activity, protests, and demonstrations, and contemporary Madison is considered the most politically liberal city in Wisconsin.[10][11][12][13] The presence of the University of Wisconsin–Madison (the largest employer in the state) as well as other educational institutions has a significant impact on the economy, culture, and demographics of Madison.[12][13][14][15][16]

As of 2021, Madison is the fastest-growing city in Wisconsin.[17] Madison's economy features a large and growing technology sector, and the Madison area is home to the headquarters of Epic Systems, American Family Insurance, Exact Sciences, Promega, American Girl, Sub-Zero, Lands' End, Spectrum Brands, a regional office for Google, and the University Research Park,[18][19][20] as well as many biotechnology and health systems startups. Madison is a popular visitor destination, with tourism generating over $1 billion for Dane County's economy in 2018.[21]

History [edit]

Madison

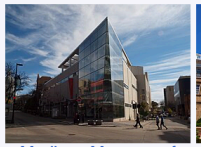
State capital city



View of the Madison Isthmus and Lake Mendota from Picnic Point

Wisconsin State Capitol

Gates of Heaven Synagogue



Madison Museum of Contemporary Art

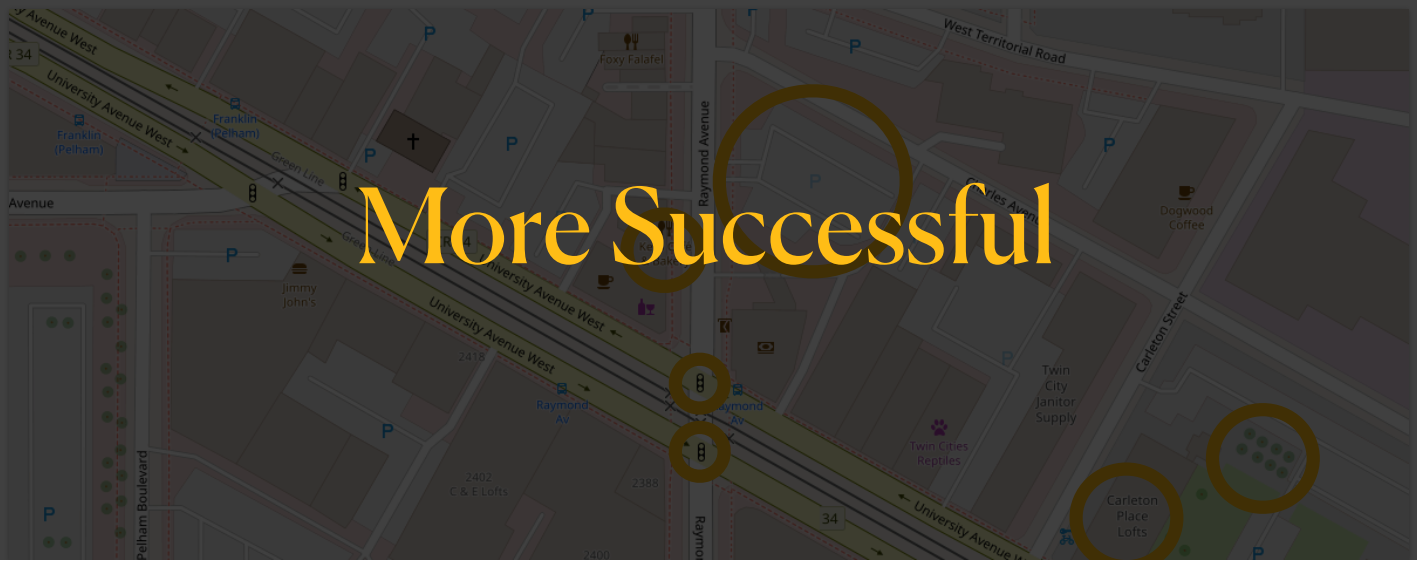
Bascom Hall at the University of Wisconsin–Madison

Flag

Seal

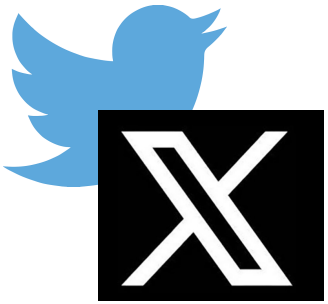
Less Successful

Sebec, ME



More Successful

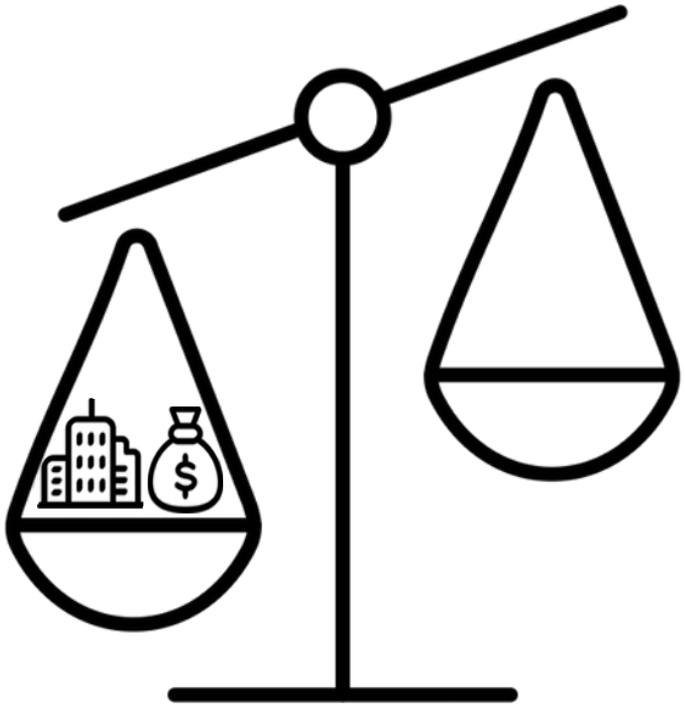
St. Paul, MN



Brent Hecht and Monica Stephens. 2014. A Tale of Cities: Urban Biases in Volunteered Geographic Information. In *Eighth International AAAI Conference on Weblogs and Social Media*. Retrieved October 20, 2016 from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8114

Linna Li, Michael F Goodchild, and Bo Xu. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science* 40, 2: 61–77.

Johnson, I., Lin, Y., Li, T., Hall, A., Halfaker, A., Schöning, J., and Hecht, B. Not at Home on the Range: Peer Production and the Urban/Rural Divide. *ACM Conference on Human Factors in Computing Systems 2016*.



Mordechai Haklay. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design* 37, 4: 682–703. https://doi.org/10.1068/b35097

Giovanni Quattrone, Afra Mashhadi, and Licia Capra. 2014. Mind the map: the impact of culture and economic affluence on crowd-mapping behaviours. In Proceedings of the 17th ACM conference on Computer Supported Cooperative Work & Social Computing, 934–944.

JANUARY 1994 • $1.95 PRINTED IN U.S.A.

Popular Mechanics

1994 DESIGN & ENGINEERING AWARDS
Best Products, Designs And Ideas Of The Year

UNDERSTANDING THE INFORMATION SUPERHIGHWAY

How You'll Shop, Bank, Learn, Be Entertained And More Via Interactive TV

THE GREATEST SOCIAL REVOLUTION SINCE THE AUTOMOBILE

The social web is

also "bypassing" rural communities

On hand, this is somewhat expected, fewer people means fewer people with interest/able to contribute

Key aspect of these platforms: theoretically, anyone, anywhere, can contribute

Question becomes:

If anyone, anywhere **can** contribute, how and why are rural areas underrepresented in these important information resources?

Rural Digital Footprints, Information Gaps, and Digital Representation

Three Main Goals

1. **Introduce Digital Footprints and Information Gaps**
2. One example of my work that helps us understand causes
3. Connect to conversations about AI and related technologies

Rural Digital Footprints, Information Gaps, and Digital Representation

Three Main Goals

1. Introduce Digital Footprints and Information Gaps
2. **One example of my work that helps us understand causes**
3. Connect to conversations about AI and related technologies

Specifically: what kinds of places do current volunteer contributors focus their effort on?

Studied in OpenStreetMap
Continental US
Individual contributors

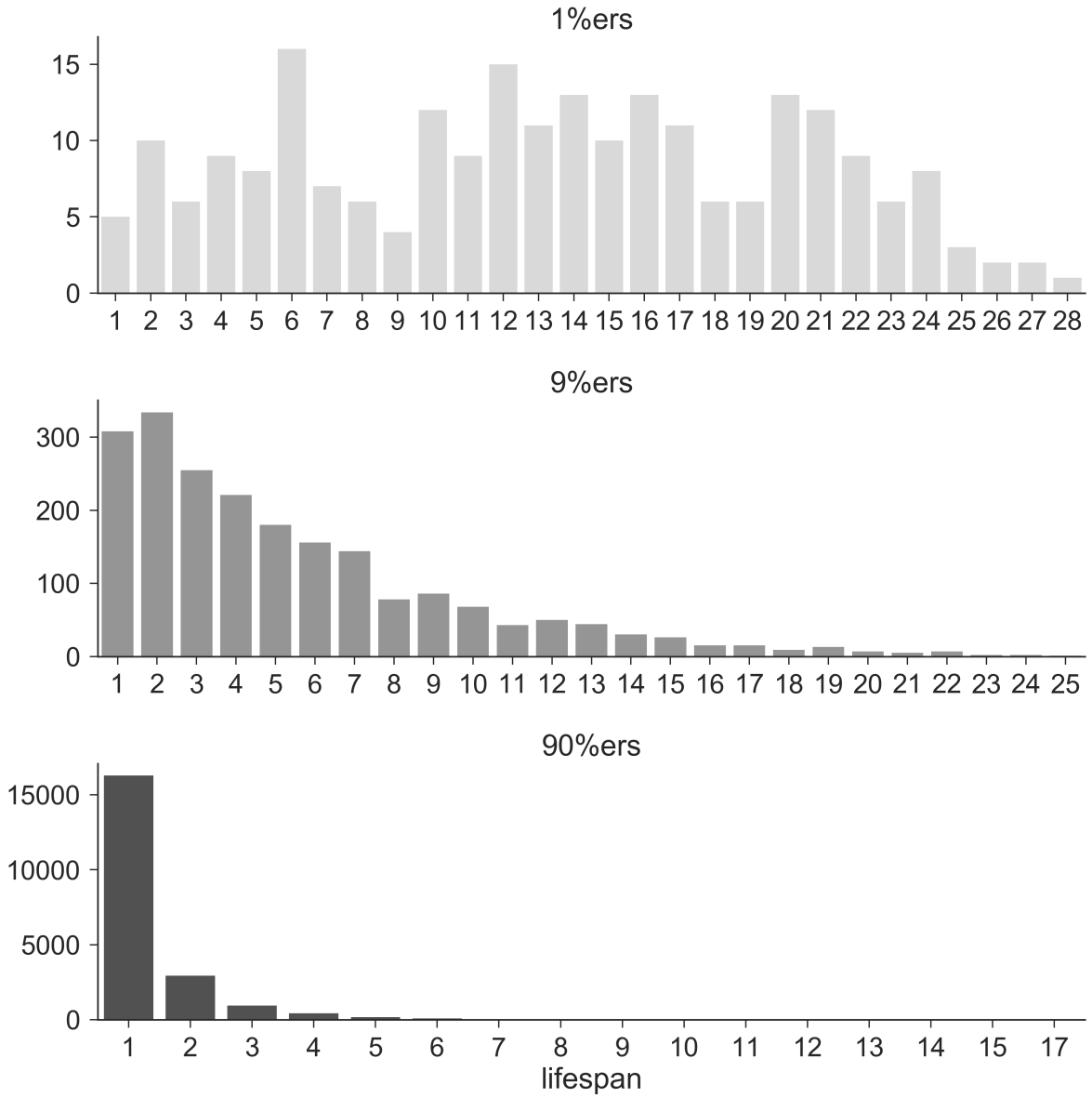
Two variables, over time:

(3mo intervals)

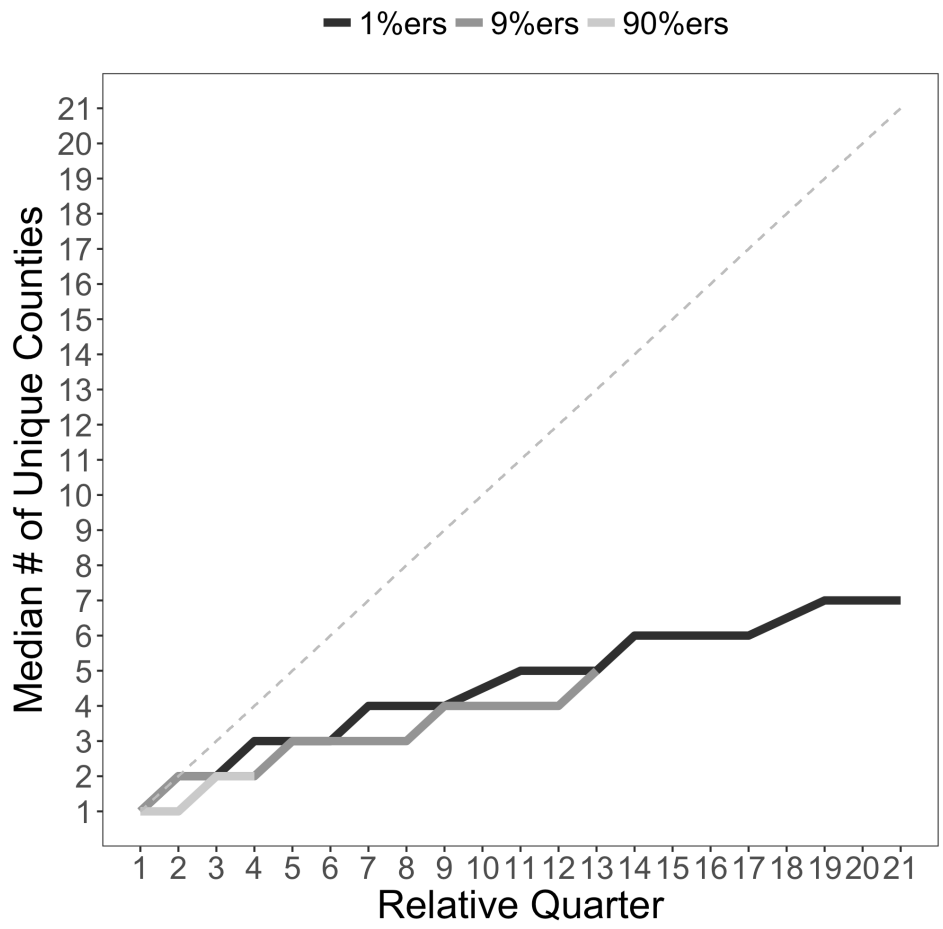
- **"Focus" county**: where contributors focus
- **% Rural**: what portion of contributors' effort happens in rural counties?

Categories of Volunteer Contributors

- Top 1% of contributors: produced 68% of the content
- Middle 9% of contributors: produced 27% of the content
- Bottom 90% of contributors: produced ~4% of the content

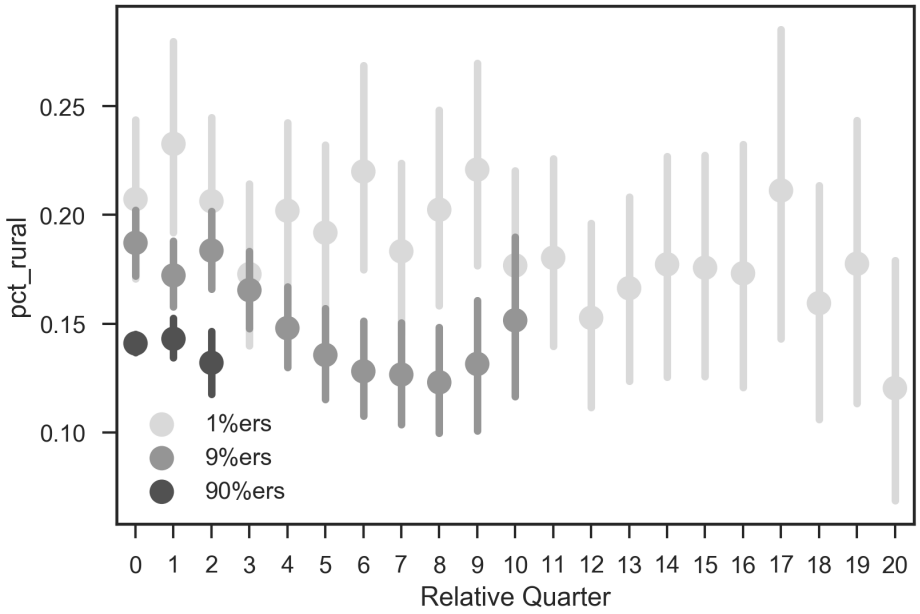

1%ers
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
9%ers
300
200
100
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
90%ers
15000
10000
5000
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 17
lifespan

Focus County



Finding 1: People are consistent about the counties they contribute in

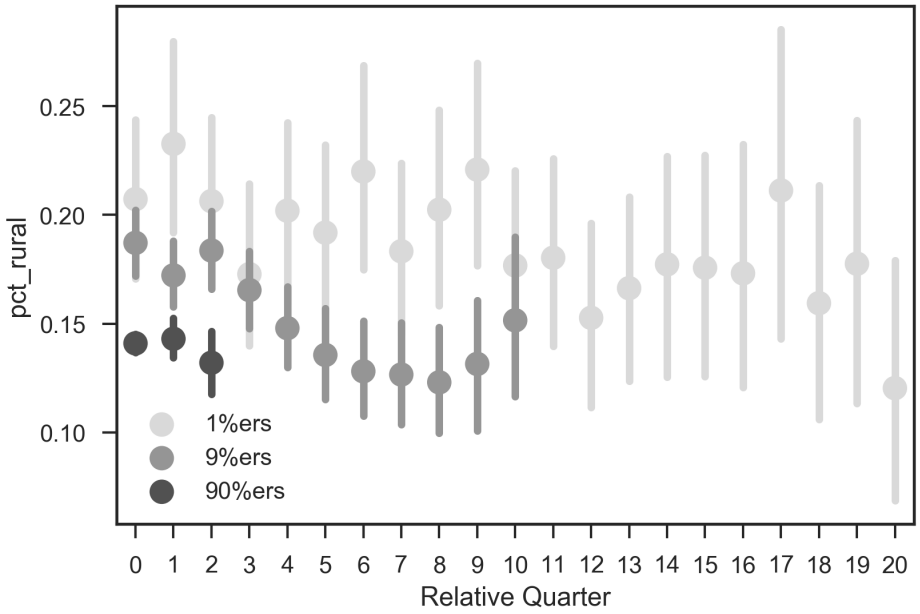
% Rural


pct_rural
0.25
0.20
0.15
0.10
1%ers
9%ers
90%ers
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Relative Quarter

Finding 2: People are consistent about the *kinds of counties* they contribute in as well (mostly not rural)

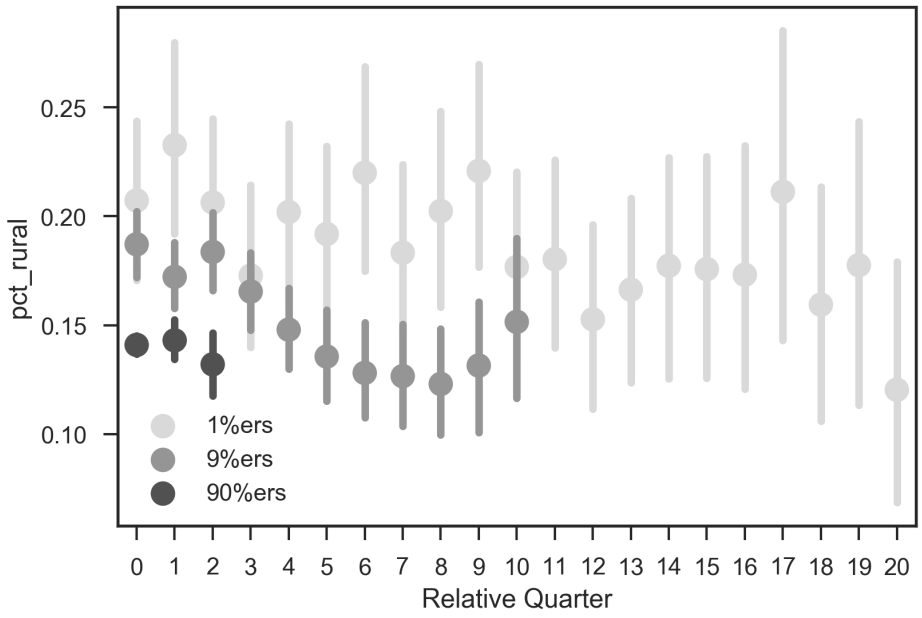
% Rural

- relative to:
  - county populations
  - # of rural/high-poverty counties
  - # of contributors focusing in rural/high-poverty areas



% Rural

- relative to:
  - county populations
  - # of rural/high-poverty counties
  - # of contributors focusing in rural/high-poverty areas



	population	# **counties**
rural	**15%**	**63%**

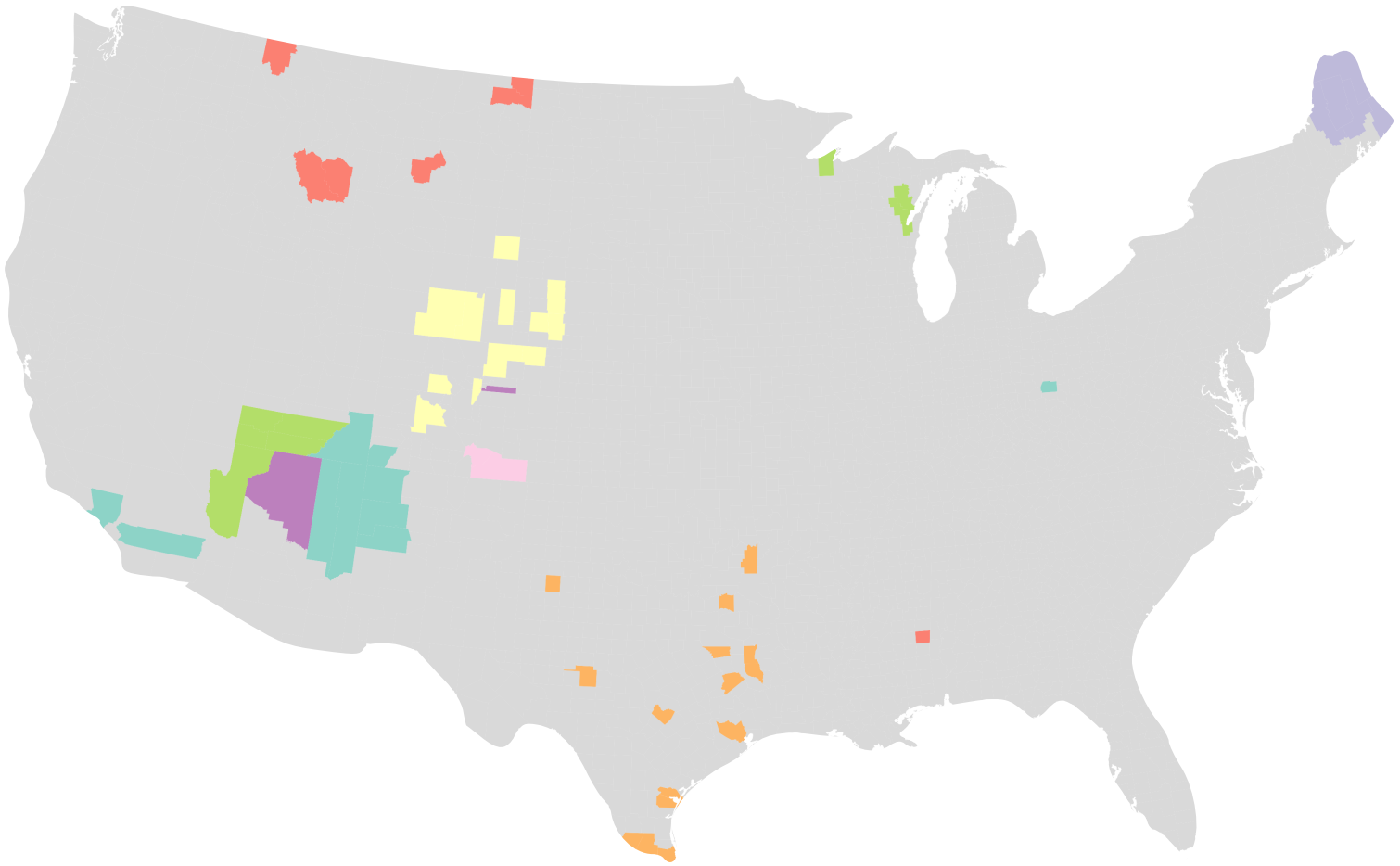
Contextualizing Findings

- relative to:
  - county populations
  - # of rural/high-poverty counties
  - # of contributors focusing in rural/high-poverty areas

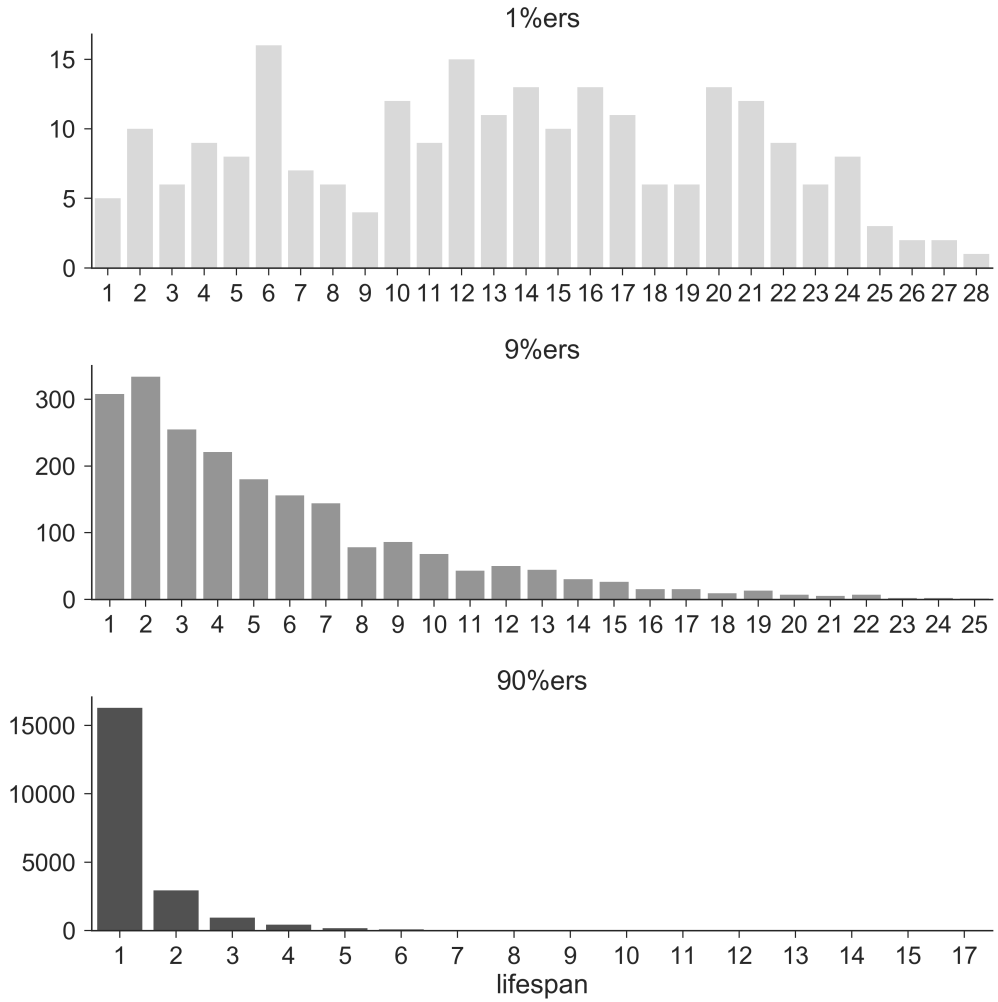
	avg. pct_rural
county populations	=
# **of counties**	-
# **focused contributors**	-

= proportional, - disproportionately low

Finding 3: In general, the rural contribution rates in this data are disproportionately low



Finding 4: The most prolific contributors often focusing on "attractive" rural places, not average rural places


1%ers
15
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
9%ers
300
200
100
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
90%ers
15000
10000
5000
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 17
lifespan

People who focus in *urban* areas **remain in the system longer**

	Urban Focused	Rural Focused	*sig. difference*
mean # of quarters	1.9	3.05	$p < 0.01$

Consistency + longevity suggest that urban areas receive more content, longer.

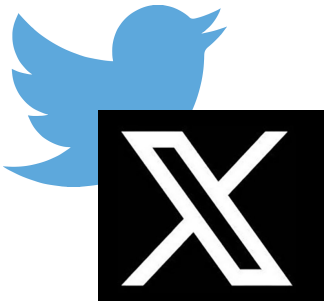
To summarize, people are consistent in the places *and kinds of places* they contribute

People who focus in urban places stick around longer

People who focus in *rural* places often focusing on "attractive" rural places

Take-away:

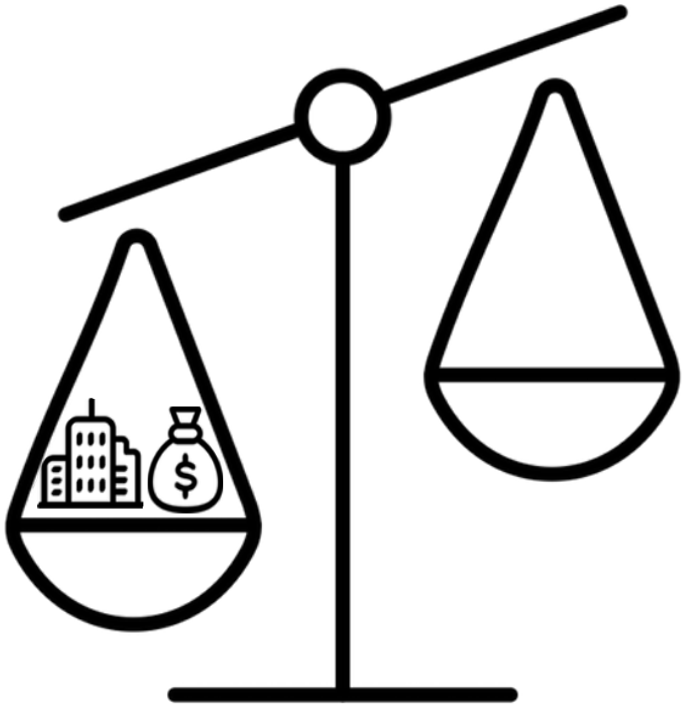
Default behavioral patterns of current contributors facilitate rural information gaps



Brent Hecht and Monica Stephens. 2014. A Tale of Cities: Urban Biases in Volunteered Geographic Information. In *Eighth International AAAI Conference on Weblogs and Social Media*. Retrieved October 20, 2016 from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8114

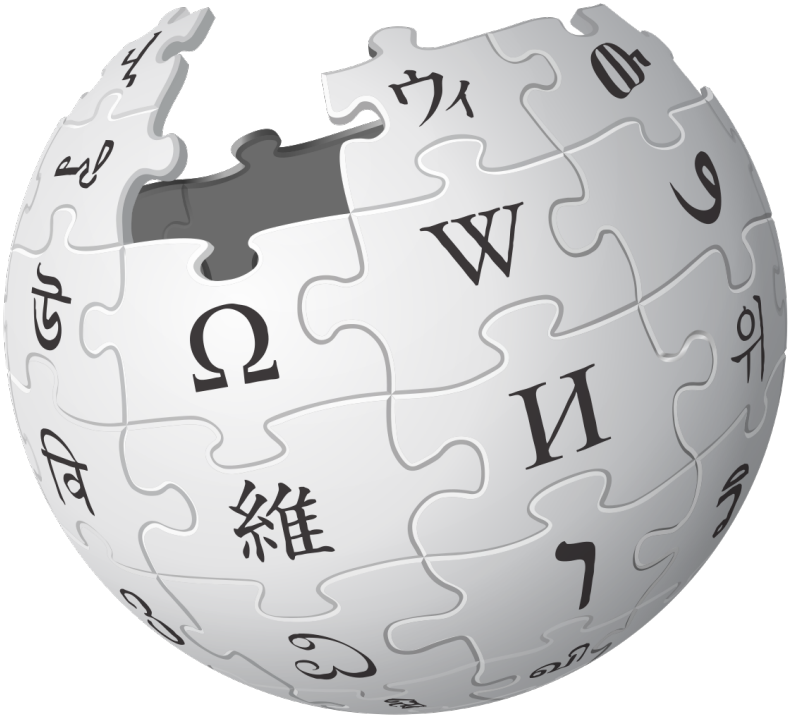
Linna Li, Michael F Goodchild, and Bo Xu. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science* 40, 2: 61–77.

Johnson, I., Lin, Y., Li, T., Hall, A., Halfaker, A., Schöning, J., and Hecht, B. Not at Home on the Range: Peer Production and the Urban/Rural Divide. *ACM Conference on Human Factors in Computing Systems 2016*.



Mordechai Haklay. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design* 37, 4: 682–703. https://doi.org/10.1068/b35097

Giovanni Quattrone, Afra Mashhadi, and Licia Capra. 2014. Mind the map: the impact of culture and economic affluence on crowd-mapping behaviours. In Proceedings of the 17th ACM conference on Computer Supported Cooperative Work & Social Computing, 934–944.


維

In short: organic, volunteer-created information resources *will* have rural information gaps

Rural Digital Footprints, Information Gaps, and Digital Representation

Three Main Goals

1. Introduce Digital Footprints and Information Gaps
2. **One example of my work that helps us understand causes**
3. Connect to conversations about AI and related technologies

Rural Digital Footprints, Information Gaps, and Digital Representation

Three Main Goals

1. Introduce Digital Footprints and Information Gaps
2. One example of my work that helps us understand causes
3. **Connect to conversations about AI and related technologies**

TESLA MOTORS CLUB

A news blog, community, forum and marketplace for Tesla and EV owners and enthusiasts covering Model 3, Model S, Model X, Model Y, Roadster, Cybertruck, Semi, and more.

Tesla Owners Can Edit Maps to Improve Summon Routes

It appears that Tesla is pulling data in realtime from OSM.

OpenStreetMap Blog home (recent posts)

OpenStreetMap Blog

Apple Maps

Apple's new maps for iOS6 make use of OpenStreetMap in some parts of the world. We're not sure how extensive this use is, but it's fair to say they are mostly using *other* sources. Apple have used TomTom as a key supplier of data for example. This means that inaccuracies in apple maps are probably not the fault of OpenStreetMap (contrary to some commentary!) However OpenStreetMap *is* mentioned in apple's credits, and we have spotted some areas where we think we can see our data in use. For example here in Islamabad:

The Data Provenance Initiative: A Large Scale Audit of Dataset Licensing & Attribution in AI

Shayne Longpre [1†] **Robert Mahari** [1,2] **Anthony Chen** [3] **Naana Obeng-Marnu** [1,4]
Damien Sileo [5] **William Brannon** [1,4] **Niklas Muennighoff** [6] **Nathan Khazam** [7]
Jad Kabbara [1,4] **Kartik Perisetla** **Xinyi (Alexis) Wu** [8] **Enrico Shippole** **Kurt Bollacker** [7]
Tongshuang Wu [9] **Luis Villa** [10] **Sandy Pentland** [1] **Sara Hooker** [11]

[1] MIT [2] Harvard Law School [3] UC Irvine [4] MIT Center for Constructive Communication
[5] Inria, Univ. Lille Center [6] Contextual AI [7] ML Commons [8] Olin College
[9] Carnegie Mellon University [10] Tidelift [11] Cohere For AI

Name	Pct
Encyclopedias	21.5%
wikipedia.org	14.6%
wikihow.com	2.7%
dbpedia	1.4%
Social Media	15.9%
reddit	6.2%
twitter	4.0%
quora	1.6%
Entertainment	8.5%
opensubtitles.org	2.5%
imdb.com	1.6%
travel guides	1.3%
Code	5.7%
stackexchange.com	2.0%
github	1.2%
opus software projects	0.9%

Website	Domain name
Google Search	google.com
YouTube	youtube.com
Facebook	facebook.com
Instagram	instagram.com
X	twitter.com
Wikipedia	wikipedia.org
Yahoo!	yahoo.com
WhatsApp	whatsapp.com
Amazon	amazon.com
Reddit	reddit.com

COMMENT | 31 October 2023

Garbage in, garbage out: mitigating risks and maximizing benefits of AI in research

Artificial-intelligence tools are transforming data-driven science — better ethical standards and more robust data curation are needed to fuel the boom and prevent a bust.

By Brooks Hanson, Shelley Stall ✉, Joel Cutcher-Gershenfeld, Kristina Vrouwenvelder, Christopher Wirz, Yuhan (Douglas) Rao & Ge Peng

When it comes to AI,
information gaps in,
information gaps out

AI tools are very likely to be ineffective and risky for rural communities, because AI tools do not have rural information

Airline held liable for its chatbot giving passenger bad advice - what this means for travellers

23 February 2024

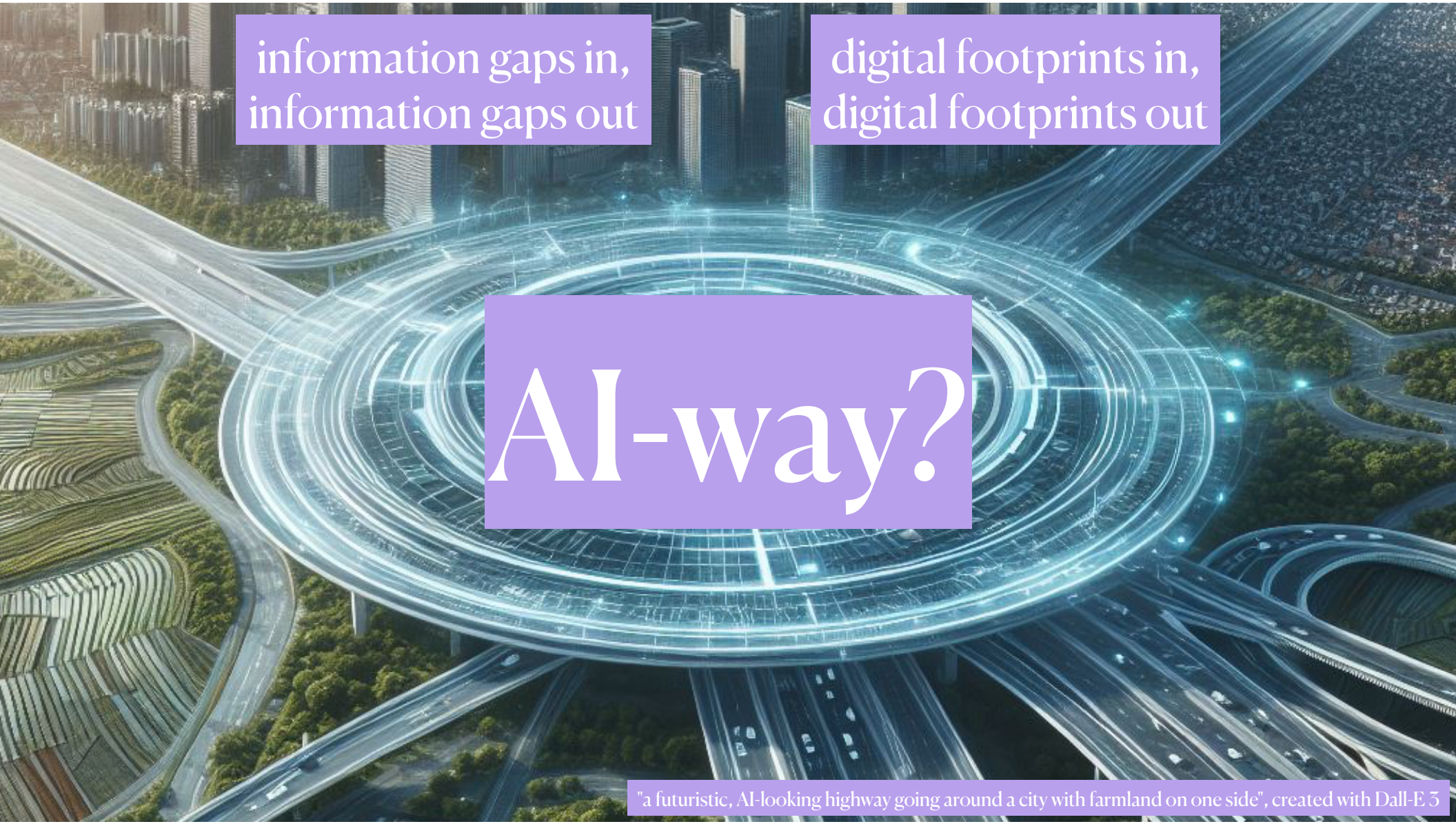
Share

By Maria Yagoda, Features correspondent

So to wrap up

The social web is

also "bypassing" rural communities



"a futuristic, AI-looking highway going around a city with farmland on one side", created with Dall-E 3

Thank you
Questions?

Dr. Jacob Thebault-Spieker, Information School, UW-Madison
e: jacob.thebaultspieker@wisc.edu
w: http://jacob.thebault-spieker.com
mastodon: @jts@hci.social / Twitter(X): @jaketangosierra