Wikipedia and the Urban-Rural Divide

Isaac Johnson
People, Space, & Algorithms (PSA) Computing Group
Northwestern University

Allen (Yilun) Lin, Toby Jia-Jun Li, Andrew Hall, Aaron Halfaker, Johannes Schöning, Brent Hecht
Not at Home on the Range: Peer Production and the Urban/Rural Divide

Isaac L. Johnson*, Yilun Lin*, Toby Jia-Jun Li†*, Andrew Hall*, Aaron Halfaker§, Johannes Schöning‡, Brent Hecht*

*GroupLens Research, Department of Computer Science, University of Minnesota, †Human-Computer Interaction Institute, Carnegie Mellon University, §Wikimedia Foundation, ‡Expertise Center for Digital Media, Hasselt University - tUL - iMinds

{ijohnson, ylin, hall, bhecht}@cs.umn.edu, tobyli@cs.cmu.edu, ahalfaker@wikimedia.org, johannes.schoening@uhasselt.be

ABSTRACT

Wikipedia articles about places, OpenStreetMap features, and other forms of peer-produced content have become critical sources of geographic knowledge for humans and machines. The importance of geographically-referenced peer-produced content, also known as peer production volunteered geographic information or peer production VGI [31, 57], has led some researchers to inquire as to whether the information is truly "at home." In this paper, we consider the effects of geographical distance on the quality of this content. We study the quality of Wikipedia articles about places in New York City, a world city, and a rural city in Montana, USA. We then consider the effects of distance on the accessibility of this content, both in the form of geographic distance and in the form of distance from the city centers. We find that, while the content is dispersed, it is generally of high quality, and that there is a strong correlation between the quality of the content and the distance from the city centers. We also find that there is a strong correlation between the quality of the content and the distance from the city centers. We discuss the implications of these findings for the future of peer production VGI, and for the future of geographic knowledge.
Wikipedia and the _______ divide...
Wikipedia and the **Gender** divide...

- Lam et al. (2011)
- Reagle and Rhue (2011)
- Stephens (2013)
- Menking and Erickson (2015)
- Wagner et al. (2015)
Wikipedia and the Language divide...

- Hecht and Gergle (2009)
- Graham et al. (2011)
- Sen et al. (2015)
Wikipedia and the **Urban-Rural** divide...
Wikipedia and the **Urban-Rural** divide...

Why is the urban-rural divide important in the context of Wikipedia?
Wikipedia and the **Urban-Rural** divide...

- It is fundamental to peer production

Why is the urban-rural divide important in the context of Wikipedia?
United States:
Urban-rural classifications of counties

Data: 2013 National Center for Health Statistics Urban–Rural Classification Scheme For Counties
United States: Urban-rural classifications of counties

Data: 2013 National Center for Health Statistics Urban–Rural Classification Scheme For Counties
United States: Urban-rural classifications of counties

Data: 2013 National Center for Health Statistics Urban–Rural Classification Scheme For Counties
United States: Urban-rural classifications of counties

Manhattan
1.6 million people
3087 Wikipedia articles

Loving County
82 people
5 Wikipedia articles
United States: Urban-rural classifications of counties

Manhattan: 26,822 people/km², 530 people per article

Loving County: 0.04 people/km², 16 people per article
Wikipedia and the **Urban-Rural** divide...

- It is fundamental to peer production
- It represents a strong cultural divide as well

Why is the urban-rural divide important in the context of Wikipedia?
United States:
Urban-rural classifications of counties

Data: 2013 National Center for Health Statistics Urban–Rural Classification Scheme For Counties
Wikipedia and the **Urban-Rural** divide...

- It is fundamental to peer production
- It represents a strong cultural divide as well
- It appears in related online platforms

Why is the urban-rural divide important in the context of Wikipedia?
Wikipedia and the Urban-Rural divide...

- It is fundamental to peer production
- It represents a strong cultural divide as well
- It appears in related online platforms
  - Social Media: Gilbert et al. (2008), Hecht and Stephens (2014)
  - OpenStreetMap: Zielstra et al. (2010), Mashadi et al. (2013)

Why is the urban-rural divide important in the context of Wikipedia?
Q: Is Wikipedia equally successful at describing rural and urban areas?

- It is fundamental to peer production
- It represents a strong cultural divide as well
- It appears in related online platforms
  - Social Media: Gilbert et al. (2008), Hecht and Stephens (2014)
  - OpenStreetMap: Zielstra et al. (2010), Mashadi et al. (2013)

Why is the urban-rural divide important in the context of Wikipedia?
Q: Is Wikipedia equally successful at describing rural and urban areas?

A: The urban-rural divide has led to uneven (and particularly stubborn) geographies of information.

- It is fundamental to peer production.
- It represents a strong cultural divide as well.
- It appears in related online platforms:
  - Social Media: Gilbert et al. (2008), Hecht and Stephens (2014)
  - OpenStreetMap: Zielstra et al. (2012), Mashadi et al. (2013)
Q: Is Wikipedia equally successful at describing rural and urban areas?

A: The urban-rural divide has led to uneven (and particularly stubborn) geographies of information.
Q: Is Wikipedia equally successful at describing rural and urban areas?
A: The urban-rural divide has led to uneven (and particularly stubborn) geographies of information.
Methods

1. Collect geographic Wikipedia content
Peer-Production Datasets

218,709 geotagged articles
21.6 million outlinks
25.2 million inlinks

English Wikipedia
Peer-Production Datasets

218,709 geotagged articles

Chicago

This article is about the city in Illinois. For other uses, see Chicago (disambiguation).

Chicago (ˌʃɪˈkɔːrə or ˌʃəˈkoʊrə), officially the City of Chicago, is the third-most populous city in the United States. With over 2.7 million residents, it is the most populous city in the state of Illinois and the Midwestern United States, and the county seat of Cook County. The Chicago metropolitan area, often referred to as Chicagoland, has nearly 10 million people and is the third-largest in the U.S. [4]

Chicago was incorporated as a city in 1837, near a portage between the Great Lakes and the Mississippi River watershed, and grew rapidly in the mid-nineteenth century. [5] Positioned along Lake Michigan, the city is an international hub for finance, commerce, industry, technology, transportation, education, tourism, art, and culture.
Peer-Production Datasets

218,709 geotagged articles

Coordinates: 41°50′13″N 87°41′05″W
Methods

1. Collect geographic Wikipedia content
2. Label each article as urban or rural
United States: 
Urban-rural classifications of counties

Data: 2013 National Center for Health Statistics Urban–Rural Classification Scheme For Counties
United States: Urban-rural classifications of counties

Data: 2013 National Center for Health Statistics Urban–Rural Classification Scheme For Counties

Methods

1. Collect geographic Wikipedia content
2. Label each article as urban or rural
3. Analyze quantity
4. Analyze the process
5. Analyze quality
Regressions

Wikipedia Content ~ Percent Urban* + Control Variables

Quantity:
e.g., articles per capita

Process:
e.g., % of content by “human” editors

Quality:
e.g., % of articles of at least C-class quality

Household Median Income†
Median Age*
% White, Non-Latino*
2012 % Democratic Vote‡
% “White Collar” Jobs†

* 2010 US Census
† 2009-2013 American Communities Survey
‡ The Guardian
[Spatial] Regressions

Wikipedia Content ~ Percent Urban + Control Variables
English Wikipedia: Articles per capita

<table>
<thead>
<tr>
<th>Quantity</th>
<th>English Wiki Articles per 1k people</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>61</td>
</tr>
</tbody>
</table>

*Color breaks log-transformed*
English Wikipedia: Articles per capita

*Color breaks log-transformed*
English Wikipedia: Articles per capita

*Color breaks log-transformed

Appalachia

Quantity

Articles per 1k people

0.1

61
English Wikipedia: Articles per capita

*Color breaks log-transformed

Quantity

Articles per 1k people

South
### Spatial Regression Results

<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity (Wikipedia)</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
</tbody>
</table>

*** $p < 0.001$
** $p < 0.01$
* $p < 0.05$
## Spatial Regression Results

<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity (Wikipedia)</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
</tbody>
</table>

The most urban areas have **3 times less articles per capita** than the most rural areas.
<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
<tr>
<td>(Wikipedia)</td>
<td>Length (Bytes) per capita</td>
<td>-0.41***</td>
<td>-0.18***</td>
<td>0.17***</td>
<td>0.19***</td>
<td>0.09***</td>
<td>0.13***</td>
</tr>
</tbody>
</table>

*** p < 0.001  
** p < 0.01  
* p < 0.05
<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity (Wikipedia)</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
<tr>
<td>Quantity (Wikipedia)</td>
<td>Length (Bytes) per capita</td>
<td>-0.41***</td>
<td>0.17***</td>
<td>0.19***</td>
<td>0.19***</td>
<td>0.09***</td>
<td>0.13***</td>
</tr>
</tbody>
</table>

Where did all of this content come from?
Orrtanna, Pennsylvania

From Wikipedia, the free encyclopedia

Orrtanna is a census-designated place (CDP) in Adams County, Pennsylvania, United States. The population was 173 at the 2010 census.[1]

Geography [edit]

Orrtanna is located in western Adams County at 39°50′53″N 77°21′27″W (39.848021, -77.357446),[2] at the eastern foot of South Mountain. The community is primarily in the northeast corner of Hamiltonian Township, with a small portion extending into the northwest corner of Highland Township.

According to the United States Census Bureau, the CDP has a total area of 0.19 square miles (0.5 km²), all of it land.[1]

Demographics [edit]

As of the census[3] of 2010, there were 173 people and 72 households residing in the CDP. The population density was 865 people per square mile (334/km²). There were 75 housing units at an average density of 375/sq mi (144.8/km²). The racial makeup of the CDP was 99.42% White and 0.58% from two or more races.

There were 72 households, of which 30.6% had children under the age of 18 living with them, 61.1% were married couples living together, 13.9% had a female householder with no husband present, and 25.0% were non-families. 20.8% of all households were made up of individuals, and 9.7% had someone living alone who was 65 years of age or older. The average household size was 2.40 and the average family size was 2.72.

In the CDP the population was spread out, with 17.92% under the age of 18, 5.2% from 18 to 24, 31.21% from 25 to 44, 28.32% from 45 to 64, and 17.34% who were 65 years of age or older. The median age was 42.3 years. For every 100 females there were 88 males. For every 100 females age 18 and over, there were 94.5 males.

The median income for a household in the CDP was $41,875, and the median income for a family was $54,063. Males had a median income of $40,000 versus $20,673 for females. The per capita income for the CDP was $21,257. About 10.6% of families and 19.9% of the population were below the poverty line, including 43.2% of those under the age of eighteen and none of those sixty five or over.

References [edit]

(cur | prev)  03:51, 23 October 2015  Chris the speller (talk | contribs)  m  .  .  . (6,129 bytes) (+3) .  .  . (<Demographics: comma(s) using AWB>) (undo)

(cur | prev)  16:32, 26 January 2015  Ken Gallagher (talk | contribs)  .  .  . (6,126 bytes) (+229 townships) (undo)

(cur | prev)  18:12, 17 August 2014  Yobot (talk | contribs)  m  .  .  . (5,897 bytes) (0) .  .  . (→Geography: WP:CHECKWIKI error fixes using AWB (10381)) (undo)

(cur | prev)  15:33, 30 July 2014  SporkBot (talk | contribs)  m  .  .  . (5,897 bytes) (+360) .  .  . (Substitute template per TfD outcome) (undo)

(cur | prev)  17:42, 26 October 2013  Hmains (talk | contribs)  m  .  .  . (5,537 bytes) (+50) .  .  . (copyedit, refine cat, and AWB general fixes using AWB) (undo)

(cur | prev)  19:21, 29 July 2013  Ken Gallagher (talk | contribs)  .  .  . (5,487 bytes) (+768) .  .  . (more geography) (undo)

(cur | prev)  18:51, 10 March 2013  Legobot (talk | contribs)  m  .  .  . (4,719 bytes) (-17) .  .  . (B Migrating 1 langlinks to WP:Wikidata - d:q278946) (undo)

(cur | prev)  02:25, 27 February 2013  Addbot (talk | contribs)  m  .  .  . (4,736 bytes) (-80) .  .  . (Bot: Migrating 5 interwiki links, now provided by Wikidata on d:q278946 (Report Errors)) (undo)

(cur | prev)  02:51, 21 December 2012  EmausBot (talk | contribs)  m  .  .  . (4,816 bytes) (+1 .  .  . (r2.7.2+) (Robot: Adding uz:Ortanna) (undo)

(cur | prev)  14:40, 13 August 2012  Magioladitis (talk | contribs)  m  .  .  . (4,854 bytes) (+12) .  .  . (→Geography: WP:CHECKWIKI error fixes using AWB (10381)) (undo)
Editor Groups

- Automatic
  - Bots

03:51, 23 October 2015 Chris the speller (talk | contribs) m . . (6,129 bytes)
(+3) . . (Demographics: comma(s) using AWB) (undo)

16:32, 26 January 2015 Ken Gallager (talk | contribs) . . (6,126 bytes) (+229
townships) (undo)

18:12, 17 August 2014 Yobot (talk | contribs) m . . (5,897 bytes) (0) . .
(WP:CHECKWIKI error fixes using AWB (10381)) (undo)

15:33, 30 July 2014 SporkBot (talk | contribs) m . . (5,897 bytes) (+360) . .
(Substitute template per TfD outcome) (undo)

17:42, 26 October 2013 Hmains (talk | contribs) m . . (5,537 bytes) (+50) . .
(copyedit, refine cat, and AWB general fixes using AWB) (undo)

19:21, 29 July 2013 Ken Gallager (talk | contribs) . . (5,487 bytes) (+768) . .
(more geography) (undo)

18:51, 10 March 2013 Legobot (talk | contribs) m . . (4,719 bytes) (-17) . . (B
Migrating 1 langlinks to WP:Wikidata - d:q278946) (undo)

02:25, 27 February 2013 Addbot (talk | contribs) m . . (4,736 bytes) (-80) . .
(Bot: Migrating 5 interwiki links, now provided by Wikidata on d:q278946 (Report Errors)) (undo)

02:51, 21 December 2012 EmausBot (talk | contribs) m . . (4,816 bytes) (+1
. (r2.7.2+) (Robot: Adding uz:Orrtanna) (undo)

14:40, 13 August 2012 Magioladitis (talk | contribs) m . . (4,793 bytes) (+83
Bot: Migrating 1 langlinks, now provided by Wikidata on d:q278946 (Report Errors)) (undo)
Editor Groups

- **Automatic**
  - **Bots**

- **Semi-Automatic (Bot-like)**
  - **AutoWikiBrowser**

---

19:21, 29 July 2013   Ken Gallager (talk | contribs)  . . (5,487 bytes) (+768) . . (more geography) (undo)

18:51, 10 March 2013   Legobot (talk | contribs)  m . . (4,719 bytes) (-17) . . (Bot: Migrating 1 langlinks to WP:Wikidata - d:q278946) (undo)

02:25, 27 February 2013   Addbot (talk | contribs)  m . . (4,736 bytes) (-80) . . (Bot: Migrating 5 interwiki links, now provided by Wikidata on d:q278946 (Report Errors)) (undo)

02:51, 21 December 2012   EmausBot (talk | contribs)  m . . (4,816 bytes) (+16) . . (r2.7.2+) (Robot: Adding uz:Orrtanna) (undo)

14:40, 13 August 2012   Magioladitis (talk | contribs)  m . . (4,800 bytes) (-8) . . (Removed empty comments and/or general fixes, removed: <!-- --> using AWB (8252)) (undo)

01:50, 18 November 2011   208.76.113.2 (talk)  . . (4,808 bytes) (-2) . . (undo)

01:46, 18 November 2011   208.76.113.2 (talk)  . . (4,810 bytes) (-51) . . (undo)

01:15, 18 November 2011   208.76.113.2 (talk)  . . (4,861 bytes) (+8) . . (undo)

13:09, 4 October 2011   Lightbot (talk | contribs)  . . (4,853 bytes) (-13) . . (Geography: Delink non-obscure units. Conversions. Report bugs to Lightmouse) (undo)

19:37, 5 September 2011   Hmains (talk | contribs)  m . . (4,866 bytes) (-46) . . (References: copyedit, refine cat, and AWB general fixes, removed: Category:Census-designated places in Pennsylvania using AWB) (undo)
Editor Groups

- **Automatic**
  - Bots

- **Semi-Automatic (Bot-like)**
  - AutoWikiBrowser

- **Manual (Human)**
  - Anonymous Editors
Chris the speller (talk | contribs) m . . (6,129 bytes) (+3) . . (Demographics: comma(s) using AWB) (undo)

Ken Gallager (talk | contribs) . . (6,126 bytes) (+229 townsships) (undo)

Yobot (talk | contribs) m . . (5,897 bytes) (0) . . (Geography: WP:CHECKWIKI error fixes using AWB (10381)) (undo)

SporkBot (talk | contribs) m . . (5,897 bytes) (+360) . . (Substitute template per TfD outcome) (undo)

Hmains (talk | contribs) m . . (5,537 bytes) (+50) . . (copyedit, refine cat, and AWB general fixes using AWB) (undo)

Ken Gallager (talk | contribs) . . (5,487 bytes) (+768) . . (more geography) (undo)

Legobot (talk | contribs) m . . (4,719 bytes) (-17) . . (Migrating 1 langlinks to WP:Wikidata - d:q278946) (undo)

Addbot (talk | contribs) m . . (4,736 bytes) (-80) . . (Bot: Migrating 5 interwiki links, now provided by Wikidata on d:q278946 (Report Errors)) (undo)

EmausBot (talk | contribs) m . . (4,816 bytes) (+1 . (r2.7.2+) (Robot: Adding uz:Orrtanna) (undo)

Magioladitis (talk | contribs) . . (4,798 bytes) (-14) . . (Substitute template per TfD outcome) (undo)
Editor Groups

● Automatic
  ○ Bots

● Semi-Automatic (Bot-like)
  ○ AutoWikiBrowser

● Manual (Human)
  ○ Anonymous Editors
  ○ User Accounts

Editor Groups

- **Automatic**
  - Bots
- **Semi-Automatic (Bot-like)**
  - AutoWikiBrowser
- **Manual (Human)**
  - Anonymous Editors
  - User Accounts
    - “Fly-by” editor
    - Locally focused

---

* ...no local editors for Orrtanna unfortunately

---

The proportion of bot-like content generated by “fly-by” editors is 40%. For “locally-focused” editors, it is 16%.
<table>
<thead>
<tr>
<th>Time</th>
<th>Username</th>
<th>Action</th>
<th>Bytes</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>03:51, 23 Oct 2015</td>
<td>Chris the speller</td>
<td>(talk</td>
<td>contribs) m (6,129 bytes) (+3) Demographics: comma(s) using AWB (undo)</td>
<td>m</td>
</tr>
<tr>
<td>16:32, 26 Jan 2015</td>
<td>Ken Gallager</td>
<td>6,126 bytes (+229) (townships) (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>18:12, 17 Aug 2014</td>
<td>Yobot</td>
<td>m (5,897 bytes) (0) Geography: WP:CHECKWIKI error fixes using AWB (10381) (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>15:33, 30 Jul 2014</td>
<td>SporkBot</td>
<td>m (5,897 bytes) (+360) Substitute template per TfD outcome (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>17:42, 26 Oct 2013</td>
<td>Hmains</td>
<td>m (5,537 bytes) (+50) copyedit, refine cat, and AWB general fixes using AWB (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>19:21, 29 Jul 2013</td>
<td>Ken Gallager</td>
<td>m (5,487 bytes) (+768) (more geography) (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>18:51, 10 Mar 2013</td>
<td>Legobot</td>
<td>m (4,719 bytes) (-17) Migrating 1 langlinks to WP:Wikidata - d:q278946 (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>02:25, 27 Feb 2013</td>
<td>Addbot</td>
<td>m (4,736 bytes) (-80) Bot: Migrating 5 interwiki links, now provided by Wikidata on d:q278946 (Report Errors) (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>02:51, 21 Dec 2012</td>
<td>EmausBot</td>
<td>m (4,816 bytes) (+106) (Robot: Adding uz:Orrtanna) (undo)</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>14:40, 13 Aug 2012</td>
<td>Magioladitis</td>
<td>m (4,455 bytes) (2035 bytes)</td>
<td>m</td>
<td></td>
</tr>
</tbody>
</table>
## Process Results

<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quantity</strong></td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>% Human Content</td>
<td><strong>0.26</strong>*</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.08**</td>
<td>-0.03</td>
<td>0.12***</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td>% Local-Focus Content</td>
<td><strong>0.31</strong>*</td>
<td>-0.07***</td>
<td>-0.06***</td>
<td>0.07***</td>
<td>0.01</td>
<td>0.22***</td>
</tr>
</tbody>
</table>

* *** p < 0.001  ** p < 0.01  * p < 0.05
<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity (Wikipedia)</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
<tr>
<td>Process (Wikipedia)</td>
<td>% Human Content</td>
<td>0.26***</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.08**</td>
<td>-0.03</td>
<td>0.12***</td>
</tr>
<tr>
<td>Process (Wikipedia)</td>
<td>% Local-Focus Content</td>
<td>0.31***</td>
<td>-0.07***</td>
<td>-0.06***</td>
<td>0.07***</td>
<td>0.01</td>
<td>0.22***</td>
</tr>
</tbody>
</table>

The proportion of human-generated content in the most urban areas is 95% while in the most rural areas it is 78%.
The proportion of “locally”-generated content in the most **urban** areas is **38%** while in the most **rural** areas it is **4%**.

<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process (Wikipedia)</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
<tr>
<td>Process (Wikipedia)</td>
<td>% Human Content</td>
<td>0.26***</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.08**</td>
<td>-0.03</td>
<td>0.12***</td>
</tr>
<tr>
<td>Process (Wikipedia)</td>
<td>% Local-Focus Content</td>
<td>0.31***</td>
<td>-0.07***</td>
<td>-0.06***</td>
<td>0.07***</td>
<td>0.01</td>
<td>0.22***</td>
</tr>
</tbody>
</table>

*** p < 0.001
** p < 0.01
* p < 0.05
English Wikipedia: % Human-Generated Content

37% - 98%

*Color breaks log-transformed
### What is the quality of rural peer-production information?

<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity (Wikipedia)</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
<tr>
<td>Process (Wikipedia)</td>
<td>% Human Content</td>
<td>0.26</td>
<td>0.02</td>
<td>0.18**</td>
<td>0.08**</td>
<td>-0.03</td>
<td>0.12***</td>
</tr>
<tr>
<td>Process (Wikipedia)</td>
<td>% Local-Focus Content</td>
<td>0.31***</td>
<td>-0.07***</td>
<td>-0.06***</td>
<td>0.07***</td>
<td>0.01</td>
<td>0.22***</td>
</tr>
<tr>
<td>Class</td>
<td>Criteria</td>
<td>Reader’s experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>★ FA</td>
<td>The article has attained <strong>featured article</strong> status by passing an official review. More detailed criteria</td>
<td>Professional, outstanding, and thorough; a definitive source for encyclopedic information.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>○ A</td>
<td>The article is well organized and essentially complete, having been reviewed by impartial reviewers from this WikiProject or elsewhere. Good article status is not a requirement for A-Class. More detailed criteria</td>
<td>Very useful to readers. A fairly complete treatment of the subject. A non-expert in the subject would typically find nothing wanting.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>✔ GA</td>
<td>The article has attained <strong>good article</strong> status by passing an official review. More detailed criteria</td>
<td>Useful to nearly all readers, with no obvious problems; approaching (but not equalling) the quality of a professional encyclopedia.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>The article is mostly complete and without major problems, but requires some further work to reach <strong>good article standards</strong>. More detailed criteria</td>
<td>Readers are not left wanting, although the content may not be complete enough to satisfy a serious student or researcher.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>The article is substantial, but is still missing important content or contains much irrelevant material. The article should have some references to reliable sources, but may still have significant problems or require substantial cleanup. More detailed criteria</td>
<td>Useful to a casual reader, but would not provide a complete picture for even a moderately detailed study.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start</td>
<td>An article that is developing, but which is quite incomplete. It might or might not cite adequate reliable sources. More detailed criteria</td>
<td>Provides some meaningful content, but most readers will need more.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stub</td>
<td>A very basic description of the topic. However, all very-bad-quality articles will fall into this category. More detailed criteria</td>
<td>Provides very little meaningful content; may be little more than a dictionary definition. Readers probably see insufficiently developed features of the topic and may not see how the features of the topic are significant.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class</td>
<td>Criteria</td>
<td>Reader's experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>★ FA</td>
<td>The article has attained <strong>featured article</strong> status by passing an official review.</td>
<td>Professional, outstanding, and thorough; a definitive source for encyclopedic information.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>More detailed criteria</strong></td>
<td>[show]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>○ A</td>
<td>The article is well organized and essentially complete, having been reviewed by impartial reviewers from this WikiProject or elsewhere. Good article status is not a requirement for A-Class.</td>
<td>Very useful to readers. A fairly complete treatment of the subject. A non-expert in the subject would typically find nothing wanting.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>More detailed criteria</strong></td>
<td>[show]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>○ GA</td>
<td>The article has attained <strong>good article</strong> status by passing an official review.</td>
<td>Useful to nearly all readers, with no obvious problems; approaching (but not equalling) the quality of a professional encyclopedia.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>More detailed criteria</strong></td>
<td>[show]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>The article is mostly complete and without major problems, but requires some further work to reach <strong>good article standards</strong>.</td>
<td>Readers are not left wanting, although the content may not be complete enough to satisfy a serious student or researcher.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>More detailed criteria</strong></td>
<td>[show]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>The article is substantial, but is still missing important content or contains much irrelevant material. The article should have some references to reliable sources, but may still have significant problems or require substantial cleanup.</td>
<td>Useful to a casual reader, but would not provide a complete picture for even a moderately detailed study.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>More detailed criteria</strong></td>
<td>[show]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start</td>
<td>An article that is developing, but which is quite incomplete. It might or might not cite adequate reliable sources.</td>
<td>Provides some meaningful content, but most readers will need more.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>More detailed criteria</strong></td>
<td>[show]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stub</td>
<td>A very basic description of the topic. However, all very-bad-quality articles will fall into this category.</td>
<td>Provides very little meaningful content; may be little more than a dictionary definition. Readers probably see insufficiently developed features of the topic and may not see how the features of the topic are significant.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>More detailed criteria</strong></td>
<td>[show]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class</td>
<td>Criteria</td>
<td>Reader’s experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FA</strong></td>
<td>The article has attained <strong>featured article</strong> status by passing an official review.</td>
<td>Professional, outstanding, and thorough; a definitive source for encyclopedic information.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A</strong></td>
<td>The article is well organized and essentially complete, having been reviewed by impartial reviewers from this WikiProject or elsewhere. Good article status is not a requirement for A-Class.</td>
<td>Very useful to readers. A fairly complete treatment of the subject. A non-expert in the subject would typically find nothing wanting.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GA</strong></td>
<td>The article has attained <strong>good article</strong> status by passing an official review.</td>
<td>Useful to nearly all readers, with no obvious problems; approaching (but not equalling) the quality of a professional encyclopedia.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>The article is mostly complete and without major problems, but requires some further work to reach <strong>good article standards</strong>.</td>
<td>Readers are not left wanting, although the content may not be complete enough to satisfy a serious student or researcher.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>The article is substantial, but is still missing important content or contains much irrelevant material. The article should have some references to reliable sources, but may still have significant problems or require substantial cleanup.</td>
<td>Useful to a casual reader, but would not provide a complete picture for even a moderately detailed study.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Start</strong></td>
<td>An article that is developing, but which is quite incomplete. It might or might not cite adequate reliable sources.</td>
<td>Provides some meaningful content, but most readers will need more.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stub</strong></td>
<td>A very basic description of the topic. However, all very-bad-quality articles will fall into this category.</td>
<td>Provides very little meaningful content; may be little more than a dictionary definition. Readers probably see insufficiently developed features of the topic and may not see how the features of the topic are significant.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Quality Results

<table>
<thead>
<tr>
<th>Type</th>
<th>Metric</th>
<th>More Urban</th>
<th>Richer</th>
<th>Older</th>
<th>More Democratic</th>
<th>More White, Non-Latino</th>
<th>More White-Collar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity (Wikipedia)</td>
<td>Articles per capita</td>
<td>-0.42***</td>
<td>-0.16***</td>
<td>0.19***</td>
<td>0.14***</td>
<td>0.06***</td>
<td>0.07***</td>
</tr>
<tr>
<td>Process (Wikipedia)</td>
<td>% Human Content</td>
<td>0.26***</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.08**</td>
<td>-0.03</td>
<td>0.12***</td>
</tr>
<tr>
<td>Quality (Wikipedia)</td>
<td>% C-class or greater</td>
<td>0.27***</td>
<td>0.00</td>
<td>0.00</td>
<td>0.15***</td>
<td>0.03</td>
<td>0.13***</td>
</tr>
</tbody>
</table>
The median percent of **C-class or higher articles** in the most **urban** areas is **9%** while in the most **rural** areas it is **0%**.
Orrtanna, Pennsylvania

From Wikipedia, the free encyclopedia

Orrtanna is a census-designated place (CDP) in Adams County, Pennsylvania, United States. The population was 173 at the 2010 census.[1]

Geography [ edit]

Orrtanna is located in western Adams County at 39°50′53″N 77°21′27″W (39.848021, -77.357446),[2] at the eastern foot of South Mountain. The community is primarily in the northeast corner of Hamilton Township, with a small portion extending into the northwest corner of Highland Township.

According to the United States Census Bureau, the CDP has a total area of 0.19 square miles (0.5 km²), all of it land.[1]

Demographics [ edit]

As of the census[3] of 2010, there were 173 people and 72 households residing in the CDP. The population density was 865.6 people per square mile (334.1/km²). There were 75 housing units at an average density of 375.8/sq mi (144.8/km²). The racial makeup of the CDP was 99.42% White and 0.58% from two or more races.

There were 72 households, out of which 30.6% had children under the age of 18 living with them, 61.1% were married couples living together, 13.9% had a female householder with no husband present, and 25.0% were non-families. 20.8% of all households were made up of individuals, and 9.7% had someone living alone who was 65 years of age or older. The average household size was 2.40 and the average family size was 2.72.

In the CDP the population was spread out, with 17.9% under the age of 18, 5.2% from 18 to 24, 31.2% from 25 to 44, 28.3% from 45 to 64, and 17.3% who were 65 years of age or older. The median age was 42.3 years. For every 100 females there were 88 males. For every 100 females age 18 and over, there were 94.5 males.

The median income for a household in the CDP was $41,875, and the median income for a family was $54,063. Males had a median income of $40,000 versus $20,673 for females. The per capita income for the CDP was $21,257. About 10.6% of families and 19.9% of the population were below the poverty line, including 43.2% of those under the age of eighteen and none of those sixty five or over.

References [ edit]

English Wikipedia:
Percentage of articles that are C-class or higher

*Color breaks log-transformed
English Wikipedia: Percentage of articles that are C-class or higher

*Color breaks log-transformed
English Wikipedia:
Percentage of articles that are C-class or higher

*Color breaks log-transformed
Summary of Results

Quantity: more content per person in rural areas
Summary of Results

Quantity: more content per person in rural areas

Process: content generation by humans much higher in urban areas
Summary of Results

Quantity: more content per person in rural areas

Process: content generation by humans much higher in urban areas

Quality: higher quality content in urban areas
Peer-production is less like peer-production in rural areas.
Bots are not great for quality but rural areas need them.
China: Total population per prefecture

*Color breaks log-transformed*
Rural China

China: Total population per prefecture

*Color breaks log-transformed
Chinese Wikipedia: Articles per capita

Quantity

Articles per 1k people

*Color breaks log-transformed
We need better bots.
Is Wikipedia equally successful at describing rural and urban areas?
Big thank you to...

Yilun (Allen) Lin  Andrew Hall

Toby Jia-Jun Li  Brent Hecht  +  Many others!

Aaron Halfaker  Johannes Schöning
Isaac Johnson
email: isaacj@u.northwestern.edu
website: www.isaacjoh.com

Project site:
https://meta.wikimedia.org/wiki/Research:Peer_Production_and_the_Urban-Rural_Divide
Credits