



LIBRARIANS ARE WIKIPEDIANS TOO

2 AUGUST 2012

Max Klein

Webjunction Webinar

Twitter: #wjwebinar @notconfusing

IMPORTANCE

SIZE OF WIKIPEDIA

Top Largest Websites

1. Google
2. Facebook
3. Youtube
4. Yahoo!
5. Baidu
6. Wikipedia
7. Windows Live!
8. QQ
9. Amazon
10. LinkedIn

>500,000,000
Monthly Views

6th Largest Site

89,000 Active
Editors

OF WHICH ARE NONPROFIT

Top Largest Websites

1. Google
2. Facebook
3. Youtube
4. Yahoo!
5. Baidu
6. Wikipedia
7. Windows Live!
8. QQ
9. Amazon
10. LinkedIn

No user
surveillance

CONNECTION

MISSIONS

“Imagine a world in which every single human being can freely share in the sum of all knowledge. That's our commitment.”

Library or Wikipedia?

Complete

Free

Access

QUALITY

Wikipedia had a reputation for being unreliable but Wikipedians are dedicated to quality sourcing

The latest data shows users rating scores in this order

1. “Well-sourced”
2. “Readable”
3. “Complete”
4. “Neutral”

And all are rising over time

[1]

http://en.wikipedia.org/wiki/User:Protonk/Article_Feedback

Highest article rating is “well sourced”[1]

Improving over time

CITATIONS

- 411,274 citations of books
- 244,236 citations of journals
- 57,868 citations of encyclopedias
- 342,470 of newspapers
- **1,055,845** total **print citations**^[2]
- **1,169,495** citations of **web**^[2]

[1]

http://en.wikipedia.org/wiki/Category:Articles_lacking_sources

[2]

<http://en.wikipedia.org/wiki/User:Maximilianklein/Citations>

94.2% contain at least one reference^[1]

Print citations are still heavily used

COMMUNITY

Even though online, Wikipedians consider themselves part of a community [1] of 89,000. (They even have their own newspaper [2].)

Regardless of your size you could engage locally and digitally.

[1]

<http://en.wikipedia.org/wiki/Wikipedia:Community>

[2]

<http://en.wikipedia.org/wiki/Wikipedia:Geonotice>

Wikipedians see themselves altruistically [1]

It's possible to contact local editors [2]

ENGAGEMENT

RESEARCH FLOW, PRE-INTERNET



RESEARCH FLOW, NOW



RESEARCH FLOW, FUTURE



RESEARCH FLOW, FUTURE



LINKING WIKIPEDIA TO LIBRARIES

It's not contentious that research should end at the sources.

The open question is how to create the links between Wikipedia and Libraries.

Who's linking?

What needs linking?

Where to link?

COLLABORATION SO FAR

- **Institutions** have edited Wikipedia on their own special collections.
- Editathons bring **Wikipedians** to the sources.
- Workshops train **new editors** on how to create edit Wikipedia.
- Bots (robots) are **software** that automatically edit Wikipedia.
- Media Donations are **bulk uploads** of media to *the commons*.

Direct Editing

Editathons

Workshops

Bots

Media Donation

DIRECT EDITING

The model of Librarians editing Wikipedia and linking to their own special collections was probably the earliest attempt at linking.

For instance: University of Washington Libraries [1].

Advantages: can be completed by a paid workforce.

Disadvantages: Wikipedia community can be hostile to *conflict of interest editors* if they (unknowingly) break Wikipedia policy.

[1]

<http://interactivearchivist.archivists.org/case-studies/wikipedia-at-uw/#lessons>

Link traffic to special collections sites

Done by paid staff

Wikipedia community can be hostile if not executed perfectly

EDITATHONS

Invite Wikipedians and patrons alike into the library for an event. Possibly giving access to rare materials.

For instance: Princeton Libraries[1], NYPL[2]

Advantages: Brings newcomers to the library. Exposes rare materials.

Disadvantages: Varying quantity and quality of work. Requires event planning.

[1]

http://outreach.wikimedia.org/w/index.php?title=File:Q_WLL_talk_slideshow.pdf

[2]

http://en.wikipedia.org/wiki/Wikipedia:The_Musical

Planned events

Brings foot traffic

Expose rare materials

Variable results

WORKSHOPS

Encouraging newcomers to edit by hosting workshops on editing techniques.

For instance: Harvard Workshop[1].

Advantages: Focus on new editors can bring in otherwise hesitant participants.

Disadvantages: Does not ensure ensuing work.

[1]

http://en.wikipedia.org/wiki/Wikipedia:VLL_at_Harvard

Garner new participants

Low inertia

No guaranteed action

BOTS

(Ro)bots are software that can edit Wikipedia at huge scale.

For instance: VIAFbot[1].

Advantages: Hugely scalable if the problem can be solved algorithmically.

Disadvantages: Requires very specific data focus and skills.

[1]

<http://en.wikipedia.org/wiki/Wikipedia:VIAF>

Scalable

Useful for data integration projects

Must be approved by community

Requires programming skills

MEDIA DONATION

Donations of Public Domain (or CC-BY-SA) images, and texts
Wikimedia commons help support free online materials.

For instance: National Archives^[1]

Advantages: Creates promotional press. Can enrich media by unforeseeable reuse.

Disadvantages: May not link back to donating institution.

[1]

<http://en.wikipedia.org/wiki/Wikipedia:NARA>

Generates good press

Is often remixed

Does not necessarily create links

LIBRARIAN OPPORTUNITIES

Individual librarians are also in need on Wikipedia.

Reliable sources noticeboard is a discussion forum on assessing the quality of sources[1].

“Category:Unreferenced” is a list of articles lacking sourcing entirely[2].

The Teahouse is a new, friendly approach to peer support on Wikipedia[3].

[1]

http://en.wikipedia.org/wiki/Wikipedia:Reliable_sources/Noticeboard

[2]

http://en.wikipedia.org/wiki/Category:Articles_lacking_sources

[3]

<http://en.wikipedia.org/wiki/Wikipedia:Teahouse>

Give advice on references

Improve articles

Grow an inside perspective

LIBRARY BENEFITS

In any of these cases, these collaborations are going to bring in more foot or digital traffic.

You're likely surrounded by Wikipedians but only the introduction hasn't been made yet.

More foot
traffic

More digital
traffic

THE SIMPLEST EXAMPLE

I'm going to add the OCLC number for the Origin of Species to the article as structured data.

What benefit will this have?

“| oclc = 352242”

http://en.wikipedia.org/wiki/On_the_Origin_of_Species

One
code
snippet

FORENSICS

THE TALK PAGE

Sometimes known as the “discussion page” this is the location for editors to talk about the article.

Duality: every page is two pages.

Metapage

Permanently
attached

Talk to other
editors

THE TALK PAGE – EXAMPLES

Egyptian Revolution

What would be some of the major talk today around the Egyptian Revolution article?

http://en.wikipedia.org/wiki/2011_Egyptian_revolution

PAGE RATINGS

Quality

Featured Article

A

Good Article

B

C

Start

Stub

Importance

Top

High

Mid

Low

Bottom

No

PAGE RATINGS – EXAMPLES

Egyptian Revolution

What score would you guess?

http://en.wikipedia.org/wiki/Talk:2011_Egyptian_revolution

What percentage are stubs?

http://en.wikipedia.org/wiki/Wikipedia:Version_1.0_Editorial_Team/Assessment#Quality_scale

VIEW HISTORY

The entire revision history for a page is always stored on Wikipedia.

That also includes talk pages.

Complete History

For Every Page

Forever

VIEW HISTORY- EXAMPLES

Egyptian Revolution 2011

How soon after the first protest
did the first draft come?

http://en.wikipedia.org/w/index.php?title=Talk:2011_Egyptian_revolution&action=history

USER PAGES

The user page itself is the public face that the user chooses to show.

The attached talk page of a user page acts as the message system.

Private messaging is impossible by design.

Profile

Messaging System

Transparency

USER PAGES—EXAMPLES

User:The Egyptian Liberal

What was s/he discussing?

http://en.wikipedia.org/wiki/User:The_Egyptian_Liberal

WHAT LINKS HERE

Part of Wikipedia's special structure is that links are **bidirectional**.

We take links forwards all the time.
Why not backwards?

All pages

Counting tool

New browsing
mode

WHAT LINKS HERE— EXAMPLES

Pages

What do you think links to Egyptian Revolution?

http://en.wikipedia.org/wiki/Special:WhatLinksHere/2011_Egyptian_revolution

Templates

Infoboxes are a type of template.

They are the boxes at the right with structured data

What about “infobox book”

http://en.wikipedia.org/wiki/Template:Infobox_book

PAGE VIEWS

Simple page view counter displayed over time.

All pages

All languages

Up to 90 days prior

PAGE VIEWS – EXAMPLES

Pages

On the July 18th release of a new
Joyce Biography, what effect on
Ulysses

[http://stats.grok.se/en/latest/Ulysses
\(novel\)](http://stats.grok.se/en/latest/Ulysses(novel))

THEORY

FIVE PILLARS

An
Encyclopedia

Neutral
Point of View

Free
Content

Civility

No Firm
Rules

http://en.wikipedia.org/wiki/Wikipedia:Five_pillars

ZEROETH LAW OF WIKIPEDIA

“The problem with Wikipedia is that it only works in practice. In theory, it can never work.” ~Anon

VIAFBOT

VIAFBOT - IDEA

VIAF – the **V**irtual **I**nternational **A**uthority **F**ile – is a matching algorithm for over 20 national name authority files.

VIAFbot proposes to edit 260,000 articles, adding reciprocal links to where VIAF links to Wikipedia.

Getting approval on large projects on Wikipedia is difficult because there is no central office, instead you must find *community consensus*.

VIAFbot is scheduled for completion by end of August[1].

[1]

http://en.wikipedia.org/wiki/Wikipedia:Authority_control_integration_proposal

An aggregating authority file

Uses and links to Wikipedia data

Would be valuable if Wikipedia reciprocated links

IDEAS

WE'RE ONLY HALF DONE

Wikipedia Loves Libraries is a
Wikipedian-centered approach
to the two collaborating.

We want more Library voices and
ideas.

What can you imagine in this space?

Formally <2
years old

More library
voices can shape
projects

COMMENTS AND QUESTIONS

Next steps at:

http://outreach.wikimedia.org/wiki/Wikipedia_Loves_Libraries