

# Wikipedia

ASDAL 2011

**Steve Sowder**

Head of Systems and Media Department

Andrews University

sowder@andrews.edu

Wikipedia is:

“a free content encyclopedia that anyone can edit”

# What is an Encyclopedia?

- A book or set of books giving information on many subjects or on many aspects of one subject and typically arranged alphabetically
- A multi-volume summary of the world and everything contained in it, organized alphabetically by subject.
- A comprehensive reference work with articles on a range of topics
- A rock band in Yeovil, England

# History

- Earliest surviving: *Naturalis Historia* of Pliny the Elder. AD 77-79
- The *Etymologiae* (around AD 630) by Saint Isador of Seville.

# The Universal Encyclopedia

- Paul Otlet (1868-1944)

Worked at extracting and organizing knowledge

Saw technology as a means of disassembling, synthesizing and distributing knowledge on an international scale.

Technology of his day included: loose-leaf binders, index cards, and microphotography.

The 3x5 card would be the basis of his most profound insights.

During his career he “helped found a universal bibliographic database/encyclopedia, a international library and museum, and numerous international associations”.

## Paul Otlet (2)

- Otlet proposed a universal classification scheme (UDC) based on the Dewey Decimal System that anticipated new information technology. It spanned over 2,000 pages.
- It also provided a primitive query language that allowed a subset of the catalog to be specified.
- A search service was provided until the early 1970's

# H.G. Wells

- Also interested in a universal reference work.
- In 1936 he proposed a world encyclopedia which he called a “World Brain”.
- It would be appealing to all mankind
- It would remain open to corrective criticism
- It would be skeptical of myths (no matter how venerated.)
- Nothing ever came of the “World Brain”

# A Short Digression

There are three historical orders in managing information.

1. How physical objects are arranged (e.g. shelving of a book)
2. Use of metadata and a catalog (Otlet's innovations)
3. Today's disorder (fluid, ad hoc, temporary, disposable)

Weinberger: "The third order takes the territory subjugated by classification and liberates it. Instead of forcing it into categories, it tags it".

Reagle: The notion of the "crowd" creating and editing knowledge is profoundly disturbing to some.



# A Short Digression

Being able to uniquely identify a resource on the Web continues to be important. But, the way we manage most online information is no longer so carefully organized. To some, this is thought to be a feature, not a bug.

# Project Xanadu

Ted Nelson's project Xanadu

1965: Started as “zippered lists” that would allow compound documents to be formed from pieces of other documents

1968: Software was released, but was not complete

1972: First demonstration

1998 Ted Nelson released the source code as Project Udanax

Now exists as downloadable software as

“A spatial viewer for hypertext in 3D based on the open transliterary standard”  
(see <http://www.xanadu.com/>)

# Project Gutenberg

Started in 1971 by Michael Hart to provide access to print information.

First document: He typed in a copy of the Declaration of Independence that he happened to have in his backpack.

First books were entered by volunteers typing but, eventually OCR took over.

Because OCR is not 100% accurate each page is proofread by at least two volunteers. Coordinating volunteers is a project in itself.

Each text is simple ASCII text. No HTML. No PDF.

# Nupedia

In January, 2000 Larry Sanger and Jimmy Wales talked about an Internet encyclopedia to be called “Nupedia”.

Their goal was for Nupedia to become the world largest Encyclopedia. It was to be built by experts contributing articles.

But, writing the articles took time and experts that signed up were not providing them very fast. It was three months before the first article was received and went through the complex editorial process.

In January 2001 there was frustration at the slow pace of acquiring articles. The idea of a wiki as a possible remedy was suggested.

Wikipedia.com went live on Jan 15. 2001.

In September 2003, with 24 complete articles and 74 in progress, Nupedia’s server crashed. It was never restored.

# Wikipedia

## The Five pillars

- **Wikipedia is an online encyclopedia**
- **Written from a Neutral point of view**
- **Free content that anyone can edit**
- **Editors should be respectful and civil to each other**
- **Wikipedia does not have firm rules**

# Wikipedia Today

- Articles in English: 3,660,191
- Pages: 24,207,875
- Total Edits Made: 468,101,403
- Uploaded files: 845,212
- Registered Users: 14,750,179
- Administrators: 1,791
- Over 17,000,000 articles in 270 languages
- 400 Million Unique Hits in March, 2011

# What's the Problem?

- Bias

Individual

Systemic bias in coverage

Notability of topics

Liberal

American and Corporate

Gender

# What's the Problem?

- **Reliability**

Deliberate false information

Sometimes inserted in order to test the system and demonstrate its alleged unreliability.

Accidental false information

Thought to be true by the editor.  
More of ignorance than malice



# What's the Problem?

- **Authority**

In 2004, a Philip Bradley said that he would not use Wikipedia and was "not aware of a single librarian who would. The main problem is the lack of authority. With printed publications, the publishers have to ensure that their data are reliable, as their livelihood depends on it. But with something like this, all that goes out the window."

In 2006, the Library Journal, using a panel of librarians, "the toughest critics of reference materials, whatever their format", asked "long standing reviewers" to evaluate three areas of Wikipedia (popular culture, current affairs, and science), and concluded: "While there are still reasons to proceed with caution when using a resource that takes pride in limited professional management, many encouraging signs suggest that (at least for now) Wikipedia may be granted the librarian's seal of approval"

# What's the Problem?

- **Authority - Academia**

A 2006 study by 55 academics suggests that the accuracy of Wikipedia is high. However, the results should not be seen as support for Wikipedia as a totally reliable resource.

A search in ScienceDirect looking for international scientific journals that cited Wikipedia in their references. Found:

Year article published	No. wikipedia cites
2004	9
2005	31
2006	133
2007	330
2008	451
2009	614
2010 (as of June 5)	478

# Which Leads us to:

- Paul Revere and Sarah Palin

# Summary

- Image a Wikipedia otherwise...

# Recommended Reading

- Everything is Miscellaneous: The power of the New Digital Disorder – David Weinberger, Times Books, 2007
- Good Faith Collaboration – Joseph Michael Reagle, Jr., MIT Press, 2010
- I Want My Wikipedia! – Barry Miller, et al., Library Journal (online), April 2006
- The Treachery of Keywords - William Badke, Online Magazine, May/June, v. 34, Issue 3, 2011, p52-54
- Wikipedia search “Encyclopedia”
- Wikipedia search “Five Pillars”
- Wikipedia search “Paul Otlet”
- Wikipedia search “Reliability of Wikipedia”

# Questions?

<http://www.andrews.edu/library/ASDAL/wikipedia.pptx>

Steve Sowder – [sowder@andrews.edu](mailto:sowder@andrews.edu)