PEW / INTERNET
PEW INTERNET & AMERICAN LIFE PROJECT

DATA MEMO

BY: Director Lee Rainie and Bill Tancer, General Manager of Global Research at Hitwise[1]
April 2007

36% of online American adults consult Wikipedia
It is particularly popular with the well-educated and current college-age students

More than a third of American adult internet users (36%) consult the citizen-generated online encyclopedia Wikipedia, according to a new nationwide survey by the Pew Internet & American Life Project. And on a typical day in the winter of 2007, 8% of online Americans consulted Wikipedia.

There has been ongoing controversy about the reliability of articles on Wikipedia. Still, the Pew Internet Project survey shows that Wikipedia is far more popular among the well-educated than it is among those with lower levels of education. For instance, 50% of those with at least a college degree consult the site, compared with 22% of those with a high school diploma. And 46% of those age 18 and older who are current full- or part-time students have used Wikipedia, compared with 36% of the overall internet population.

In addition, young adults and broadband users have been among those who are earlier adopters of Wikipedia. While 44% of those ages 18-29 use Wikipedia to look for information, just 29% of users age 50 and older consult the site. In a similar split, 42% of home broadband users look for information on Wikipedia, while just 26% of home dial-up users do so.

A Profile of Wikipedia Users

Do you ever use the internet to look for information on Wikipedia?	*% of adult internet users who say "yes"*
Men	39%
Women	34%
Whites	37%
Blacks	27%
English-speaking Hispanics	36%
Age 18-29	44%
Age 30-49	38%
Age 50-64	31%
Age 65+	26%
High school diploma	22%
Some college	36%
College grad +	50%
< $30,000 household income	32%
$30,000-$50,000	35%
$50,000-$75,000	39%
>$75,000	42%
Dial-up connection at home	26%
Broadband at home	42%

Source: Pew Internet & American Life Project Survey Feb. 15-March 7, 2007. N=1,492 internet users. Margin of error is ±3%.

[1] See http://hitwise.com/

All told, the use of Wikipedia is more popular on a typical day than some of the more prominent activities tracked by the Pew Internet & American Life Project, including online purchasing, visiting dating websites, making travel reservations, using chat rooms, and participating in online auctions.

The popularity of Wikipedia

Several Web traffic measuring firms say that Wikipedia is one of the most heavily visited sites on the internet. comScore Media Metrix, the Web traffic measuring firm, regularly says that monthly traffic to the cluster of Wikipedia sites positions them in the top 10 global sites.[2]

Likewise, data from Hitwise consistently registers Wikipedia in the top 10 most popular sites on the entire Web. Moreover, in the cluster of sites that are focused on educational and reference material, Wikipedia is by far the most popular site, drawing nearly six times more traffic than the next-closest site.

Top 20 Educational and Reference Websites Week Ending March 17, 2007	*Share of traffic in the category that week*
1. Wikipedia (http://www.wikipedia.org/)	24.33%
2. Yahoo! Answers (http://answers.yahoo.com/)	4.23%
3. Dictionary.com (http://dictionary.reference.com/)	3.79%
4. Answers.com (http://www.answers.com/)	3.53%
5. SparkNotes (http://www.sparknotes.com/)	1.62%
6. Google Scholar (http://scholar.google.com/http://scholar.google.com/)	1.31%
7. Google Book Search (http://books.google.com/)	1.09%
8. Find Articles (http://www.findarticles.com/)	.99%
9. U.S. National Library of Medicine (http://www.nlm.nih.gov/)	.99%
10. Merriam-Webster Online (http://www.m-w.com/)	.85%

Source: Hitwise, U.S. Internet Visits (market share) for week ending March 17, 2007

Why is Wikipedia so popular?

Hitwise data suggest several reasons for the popularity of Wikipedia: First, there is the sheer amount of material on the site, covering everything from ancient history to current events and popular culture. Because authors can create and edit entries at any time and in

[2] See, for instance:

http://www.comscore.com/press/release.asp?id=1193

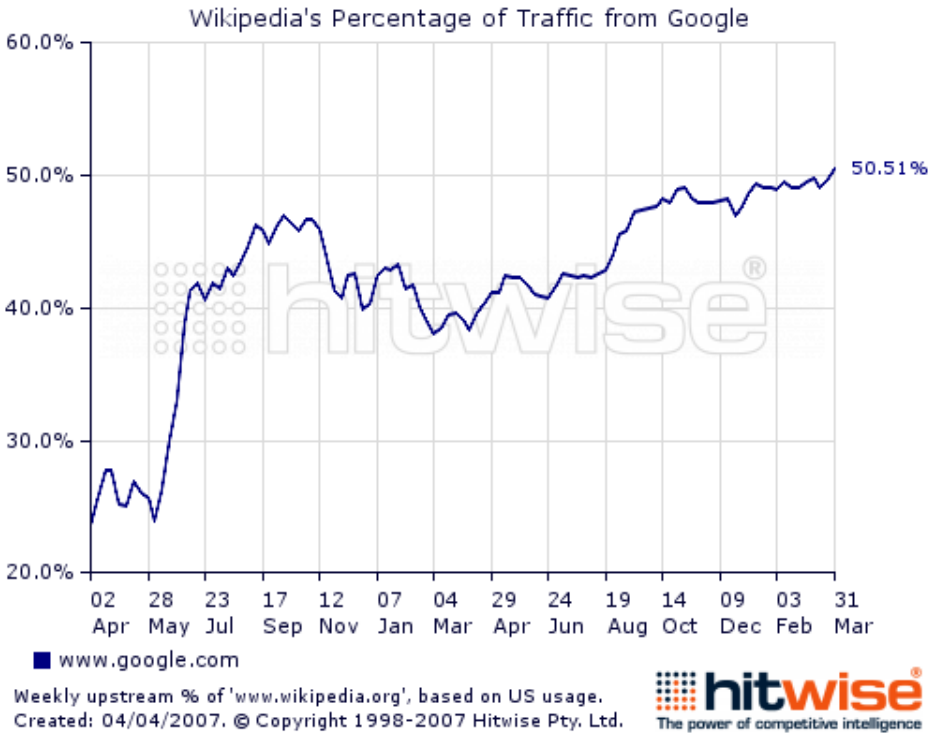
http://www.comscore.com/press/release.asp?id=1242

http://www.comscore.com/press/release.asp?id=1246

any location, new entries about emerging tech terms, celebrities or political scandals can be published or updated within a matter of minutes. In total, Wikipedia claims over 5.3 million entries, 1.6 million of them in English.

Second, Wikipedia's dramatic growth is strongly correlated with Americans' affection for search engines. Over 70% of the visits to Wikipedia in the week ending March 17 came from search engines, according to Hitwise data.

Wikipedia's article structure helps explain this. Many of the pieces in the encyclopedia are full of links to other Wikipedia articles and other material on the Web. One of the prime factors in Google's search results algorithm is the number of links connected to a given webpage. Thus, Wikipedia's culture of encouraging internal and external links gives Wikipedia material a shot at very high Google rankings on many keyword queries. In fact, Wikipedia has become the #1 external site visited after Google's search page, receiving over half of its traffic from the search engine.



Convenience is likely a big factor, too

In addition to having some logistical advantages in drawing traffic, Wikipedia is probably popular for other reasons as well. Research on other subjects by the Pew Internet Project has shown that the convenience of using sites makes a difference. In our "Science Online" report,[3] we found that science knowledge-seeking online is driven a lot by convenience -- more so than a sense that science information online is more accurate than

[3] See http://www.pewinternet.org/PPF/r/191/report_display.asp

other sources. Convenience mattered to 71% of those seeking science information, compared with 13% of those who said using the internet gave them more accurate information. In addition, our research about seekers of science information showed that users do a lot of checking using other offline and online sources of science information beyond the sites they find easily through search engines.

This finding also mirrors work we have done about the way people get health information and get political information online. The ease of using the internet to find information quickly seems a constant for many kinds of data hunts online.

About the Pew Internet & American Life Project

The Pew Internet & American Life Project is a non-partisan, non-profit initiative of the Pew Research Center that does research on the social impact of the internet. It is funded by the Pew Charitable Trusts. The Projects takes no position on policy issues.

The findings in this report come from a survey conducted between February 15 and March 7, 2007 among a representative sample of 2,200 adults, 18 and older.

More information about the Pew Internet Project is available at: www.pewinternet.org.

About Hitwise

Hitwise is an online competitive intelligence service that provides its 1,200 global clients with daily insights on how their customers interact with a broad range of competitive websites, and how their competitors use different tactics to attract online customers.

Through relationships with ISPs around the world, Hitwise's patented methodology captures the anonymous online usage, search, and conversion behavior of 25 million Internet users. Hitwise is a privately held company headquartered in New York City and operates in the United States, United Kingdom, Australia, New Zealand, Hong Kong, and Singapore. More information about Hitwise is available at www.hitwise.com.

Pew Internet Project Survey Questions

February 2007 Tracking Survey Final Topline 3/12/07

Data for February 15 – March 7, 2007

Princeton Survey Research Associates International
for the Pew Internet & American Life Project

Sample: *n* = 2,200 adults 18 and older
Interviewing dates: 02.15.07 – 03.07.07
Margin of error is plus or minus 2 percentage points for results based on total sample [n=2,200]
Margin of error is plus or minus 3 percentage points for results based on internet users [n=1,492]

Q6a Do you use the internet, at least occasionally?
Q6b Do you send or receive email, at least occasionally?

	USES INTERNET	DOES NOT USE INTERNET
Current	71	29
December 2006	70	30
November 2006	68	32
August 2006	70	30
Feb-April 2006	73	27
Jan/Feb 2006	73	27
Nov/Dec 2005	66	34
September 2005	72	28
May/June 2005	68	32
February 2005	67	33
January 2005	66	34
November 2004	61	39
May/June 2004	63	37
February 2004	63	37
November 2003	64	36
July 2003	63	37
June 2003	62	38
April/May 2003	63	37
March 3-11, 2003	62	38
February 2003	64	36
December 2002	57	43
November 2002	61	39
October 2002	59	41
September 2002	61	39
July 2002	59	41
March/May 2002	58	42
January 2002	61	39

Q7 Did you happen to use the internet YESTERDAY?

Based on internet users [N=1,492]

	YES, USED INTERNET YESTERDAY	NO, DID NOT USE INTERNET YESTERDAY	DON'T KNOW/ REFUSED
Current	69	31	*
December 2006	65	34	*
November 2006	64	36	*
August 2006	66	34	*
Feb-April 2006	66	33	*
Nov/Dec 2005	63	36	*
September 2005	65	34	*
February 2005	60	40	*
January 2005	58	42	*
November 2004	61	39	*
May/June 2004	53	46	1
February 2004	55	44	*
November 2003	54	45	*
July 2003	52	47	1
June 2003	55	44	*
April/May 2003	58	42	*
March 3-11, 2003	60	40	0
February 2003	60	40	*
December 2002	56	44	*
November 2002	57	43	*
October 2002	57	43	0
September 2002	58	42	*
June 26-July 26, 2002	53	47	*
March/May 2002	57	43	*
January 2002	59	41	*

WEB1 Please tell me if you ever use the internet to do any of the following things. Do you ever use the internet to…/Did you happen to do this **yesterday**, or not?[4]

	total have ever DONE THIS	---------- DID have not done YESTERDAY this		don't know/ refused
Look for information on Wikipedia				
Current	36	8	62	2

[4] Prior to January 2005, question wording was "Please tell me if you ever do any of the following when you go online. Do you ever...?/Did you happen to do this yesterday, or not?"

Methodology

This report is based on the findings of a daily tracking survey on Americans' use of the Internet. The results in this report are based on data from telephone interviews conducted by Princeton Survey Research Associates between February 15 to March 7 2007, among a sample of 2,200 adults, 18 and older. For results based on the total sample, one can say with 95% confidence that the error attributable to sampling and other random effects is plus or minus 2.3 percentage points. For results based Internet users (n=1,492), the margin of sampling error is plus or minus 2.8 percentage points. In addition to sampling error, question wording and practical difficulties in conducting telephone surveys may introduce some error or bias into the findings of opinion polls.

The sample for this survey is a random digit sample of telephone numbers selected from telephone exchanges in the continental United States. The random digit aspect of the sample is used to avoid "listing" bias and provides representation of both listed and unlisted numbers (including not-yet-listed numbers). The design of the sample achieves this representation by random generation of the last two digits of telephone numbers selected on the basis of their area code, telephone exchange, and bank number.

New sample was released daily and was kept in the field for at least five days. The sample was released in replicates, which are representative subsamples of the larger population. This ensures that complete call procedures were followed for the entire sample. At least 10 attempts were made to complete an interview at sampled households. The calls were staggered over times of day and days of the week to maximize the chances of making contact with a potential respondent. Each household received at least one daytime call in an attempt to find someone at home. In each contacted household, interviewers asked to speak with the youngest male currently at home. If no male was available, interviewers asked to speak with the youngest female at home. This systematic respondent selection technique has been shown to produce samples that closely mirror the population in terms of age and gender. All interviews completed on any given day were considered to be the final sample for that day. The final response rate on this survey is 29.3%.

Non-response in telephone interviews produces some known biases in survey-derived estimates because participation tends to vary for different subgroups of the population, and these subgroups are likely to vary also on questions of substantive interest. In order to compensate for these known biases, the sample data are weighted in analysis. The demographic weighting parameters are derived from a special analysis of the most recently available Census Bureau's March 2006 Annual Social and Economic Supplement. This analysis produces population parameters for the demographic characteristics of adults age 18 or older, living in households that contain a telephone. These parameters are then compared with the sample characteristics to construct sample weights. The weights are derived using an iterative technique that simultaneously balances the distribution of all weighting parameters.