Wikipedia Survey – First Results

Working Draft

Current Version: 0.3
Date: 9 April 2009

Ruediger Glott (UNU-MERIT), glott@merit.unu.ed
Philipp Schmidt (UNU-MERIT), schmidt@merit.unu.edu
Rishab Ghosh (UNU-MERIT)

This file contains an ongoing work in progress. We will add further analysis of the Wikipedia survey data to it periodically. This document is not suitable for citation. Do not distribute without clearly indicating that it represents preliminary results that might be updated in the future. Please contact us if you have any questions.

QUESTIONS DEALT WITH IN THIS REPORT (numbering refers to questionnaire)

Question A1 – gender
Question A2 – Age
Question A7 - tertiary education
Question C14 - If you do note donate, why not?
Question D1A - Why don't you contribute to Wikipedia?
Question D1B - I would be much likelier to contribute if ..

QUESTIONS OF HIGH INTEREST
Question A4 – Country
Question A9 - has partner
Question A10 - has children
Question B13 - How many hours?
Question B17 - Why do you contribute?

Data Overview

The following sections present first results of the first global Wikipedia Survey, based on a debugged database (i.e cleaned from erroneous entries such as extreme old or young ages, or evidently wrong combination of responses, such as having a PhD but being 14 years of age, and the like). For the purpose of this first data analysis we have also removed the response to the Russian version of the survey because it is still unclear why this group of the Wikipedians is over-represented. These data will be re-integrated and analysed after we have evidence that the Russian survey has not been manipulated (so far we did not find any hints to that, e.g. different records from the same IP address or similar answering patterns, but this examination is ongoing at the moment).

Overall, the database that has been used for this first overview contains 130,576 cases.

Wikipedia Language Versions

Table 1 shows the relative shares of the different language versions in the total response. The lion share of responses was provided by users of the English Wikipedia version, followed by the German and Spanish version.

```
+-------+--------+---------+
| N     | Lang   | %       |
+-------+--------+---------+
| 43912 | en     | 33.6295 |
| 22989 | de     | 17.6058 |
| 20144 | es     | 15.4270 |
|  8706 | nl     |  6.6674 |
|  6155 | pt     |  4.7137 |
|  5082 | ja     |  3.8920 |
|  3613 | zhhant |  2.7670 |
|  3569 | fr     |  2.7333 |
|  3516 | it     |  2.6927 |
|  3427 | zhhans |  2.6245 |
|  2907 | pl     |  2.2263 |
|  2319 | cs     |  1.7760 |
|  1567 | ar     |  1.2001 |
|   969 | th     |  0.7421 |
|   598 | vi     |  0.4580 |
|   562 | ca     |  0.4304 |
|   213 | id     |  0.1631 |
|   154 | ta     |  0.1179 |
|   113 | el     |  0.0865 |
|    47 | eo     |  0.0360 |
|    14 | af     |  0.0107 |
+-------+--------+---------+
 N = 130576
```

Table 1: Respondents by Wikipedia language version

Countries

Table 2 provides an overview of the countries where the respondents live, illustrating that Wikipedia is used in at least 228 countries in the world.[1]

```
+-------+--------+-------+          |-----------------------|
| N     | Country| %     |          | N: 51 - 200           |
+-------+--------+-------+          |-----------------------|
| 20084 | DE     | 15.46 |          | HU, AE, RU, SV, GT, TR, |
| 18413 | US     | 14.17 |          | RS, SK, PK, HR, PR, DZ, |
|  6842 | NL     |  5.27 |          | BO, BG, PA, JO, PY, EE, |
|  6010 | BR     |  4.63 |          | MA, KW, SI, HN, LU, UA, |
|  5996 | ES     |  4.61 |          | LT, KR, BA, LK, IR, NI, |
|  5431 | GB     |  4.18 |          | MG, BH, IS,           |
|  5128 | JP     |  3.95 |          |-----------------------|
|  4946 | MX     |  3.81 |          | N: 11 - 50            |
|  3693 | IT     |  2.84 |          |-----------------------|
|  3450 | CA     |  2.66 |          | PS, CI, MO, GE, SY, TN, |
|  3416 | FR     |  2.63 |          | LV, LB, BD, IQ, AD, QA, |
|  3175 | AR     |  2.44 |          | OM, TT, MT, LY, CY, CU, |
|  2952 | PL     |  2.27 |          | AL, MK, NG, MU, NP, UM, |
|  2848 | PG     |  2.19 |          | JM, KE, SD, BY, IM, YE, |
|  2531 | AU     |  1.95 |          | AW, AQ, AN, LI, MD, AX, |
|  2431 | CZ     |  1.87 |          | ZW, RE, FJ, MC,       |
|  2399 | BE     |  1.85 |          |-----------------------|
|  2252 | TW     |  1.73 |          | N: <= 10             |
|  2009 | AT     |  1.55 |          |-----------------------|
|  1996 | IN     |  1.54 |          | AO, BM, AS, BZ, VI, AF, |
|  1828 | CL     |  1.41 |          | UG, BN, BB, MV, GH, SN, |
|  1690 | CO     |  1.30 |          | AZ, DM, BS, CD, JE, AM, |
|  1479 | SE     |  1.14 |          | GP, MN, BF, KH, BW, VG, |
|  1417 | CH     |  1.09 |          | NC, BT, GI, EH, ME, ZM, |
|  1048 | VE     |  0.81 |          | PF, ET, DJ, CV, KN, VA, |
|   989 | TH     |  0.76 |          | KY, GL, GG, CF, AG, MZ, |
|   977 | HK     |  0.75 |          | SR, CM, BJ, BI, KZ, MM, |
|   875 | PE     |  0.67 |          | MQ, KP, AI, GU, FK, BL, |
|   643 | NZ     |  0.49 |          | LC, FO, CX, WS, CC, TM, |
|   627 | CN     |  0.48 |          | ML, TZ, GQ, LA, RW, TL, |
|   575 | VN     |  0.44 |          | MP, GF, MW, MR, KI, AC, |
|   561 | EG     |  0.43 |          | PN, VC, TO, TC, SZ, SO, |
|   530 | IE     |  0.41 |          | GS, KG, UZ, CG, MS, CK, |
|   499 | IL     |  0.38 |          | SM, NF, TK, GD, BV, SC, |
|   477 | PT     |  0.37 |          | LR, NA, SB, SL, TD, NE, |
|   440 | SA     |  0.34 |          | IO, GA, NU, WF, ER, MF, |
|   437 | FI     |  0.34 |          | LS, PW, TG,           |
|   379 | RO     |  0.29 |          | +-------+------------+ |
|   368 | NO     |  0.28 |            N = 130576
|   343 | SG     |  0.26 |
|   331 | GR     |  0.25 |
|   312 | UY     |  0.24 |
|   295 | DK     |  0.23 |
|   284 | ID     |  0.22 |
|   275 | CR     |  0.21 |
|   265 | ZA     |  0.20 |
|   260 | PH     |  0.20 |
|   257 | MY     |  0.20 |
|   217 | EC     |  0.17 |
|   211 | DO     |  0.16 |
|-----------------------|
```

Table 2: Countries

---

1 Refer to ISO 3166 for country codes http://www.iso.org/iso/english_country_names_and_code_elements

Activity Types

Since Wikipedia can be used in many different ways, the key feature to differentiate respondents – besides language and demographic characteristics – is provided by the activities the respondents perform. The first question of the questionnaire was obligatory to answer and allowed to differentiate between users and different types of contributors.

Type	N	%
Reader	81677	62.55
Occasional Contributor	31981	24.49
Regular Contributor	11992	9.18
Author	[8485]	[6.50]
Editor	[7791]	[6.97]
Administrator	[1278]	[0.98]
Ex Contributor	3625	2.78
No more activity	[1585]	[1.21]
Other	1301	1.00
	130576	100.00%

Table 3: Detailed categories of respondents

Not surprisingly, pure readers provide the majority of the respondents (62.5%). However, the share of almost 34% of the respondents that contribute actively to Wikipedia appears considerably large. Since it is impossible to quantify the universe of people using Wikipedia and how many of them are contributors it will be hard to evaluate whether the shares we found in our survey are representative.

The largest share of contributors (24.5%) is provided by respondents that consider themselves as readers but occasionally contribute as editors or authors. Among respondents occasional contributions are the typical way of Wikipedia content creation. However, this does not imply that most of the content is created this way. This question will be examined in later parts of the analysis.

Regular contributors (authors, editors, and administrators – overlap possible) represent about 9% of the respondents. Approximately 1.0% of the respondents are Administrators.

It is noteworthy to mention that among the resopndents there are more than 3,600 persons that contributed to Wikipedia in the past but meanwhile stopped participating actively. Since this group has a unique insight and relationship to Wikipedia, it is analyzed separately. A remarkable share of this group (44%) checked no other activity, which seem to imply that these ex contributors do not even consider themselves anymore as readers of Wikipedia (though this raises the question how this group has heard of the survey). As a conclusion, our profile of Wikipedia activity groups takes account of a group that explicitly pursues no activity at all (share in overall response: 1.2%).

The Other category combines respondents who specified they were either "Other non-contributors" or "Other contributors), but did not check any of the other options. Further analysis of the explanation text that these respondents submitted is required to either group them within existing categories or analyze separately.

It is evident that such a detailed and uneven break-down of activities is not useful for the further analysis of demographics and other characteristics that are presented in the following sections. Therefore, for the following parts of the analysis we will use a break-down that subsumes these activities in only two categories, readers (65%) and contributors (35%) excluding Ex Contributors and Others.

Type	N	%
Reader	81677	65.00
Contributor	43973	35.00
	125,650	100

Table 4: Main categories

Demographics

Age

Valid responses were received from respondents between 10 – 85 years. Overall, the average age of the Wikipedians that have participated in the survey is 25.8 years. Half of the respondents are are younger than 22 years. The most frequent age that can be observed within the repondents is 18 years. Splitting the respondents in four equally large age groups shows that 25% are younger than 18 years old, 25% are between 18 and 22, a further 25% are between 23 and 30 (e.g. half of the respondents are between 18 and 30 years) and the remaining 25% are between 31 and 85 years old. .

Although the age difference between readers and contributors appears marginal - readers are, on average, 25.3 years old while contributors show an average age of 26.8 years – a t-test revealed that this age difference is statistically significant.

Finally, female Wikipedians are significantly younger (24.2 years) than male community members (26.4 years).

Gender

Though both groups are dominated by men, there are significant differences in the gender composition of readers and contributors of Wikipedia. Contributors show a substantially larger share of males than readers.

Gender	Reader	Contributor	Total
Male	55648 (59.45%) (68.28%)	37958 (40.55%) (86.56%)	93606
Female	25451 (18.10%) (31.23%)	5626 (81.90%) (12.83%)	31077
Other	398	266	664
Total	81497	43850	N = 125,347

Table 5: Gender Composition of Wikipedia Activity Groups

Education

Given the young age of Wikipedians it is just natural that secondary and undergraduate tertiary education provide the prevailing highest educational degrees among the respondents. Contributors show a slightly but significantly better education than readers.

	Reader (%)	Contributor (%)	Total (%)
Primary education	14.77	12.36	13.93
Secondary education	37.69	33.45	36.20
Tertiary education – Undergraduate	29.17	28.38	28.89
Tertiary education – Masters	10.11	15.07	11.85
Tertiary education – PhD	2.30	4.59	3.10
Other	5.96	6.15	6.03
N = 124,752	100%	100%	100%

Table 6: Highest Educational Degree by Wikipedia Activity Groups

Relationship Status

The status of the respondents regarding partnerships also corresponds to the age structure of the Wikipedia community, as only 30% of the respondents say they have a partner. Correspondingly, the share of Wikipedians with children is only 14%. The difference between readers and contributors in this regard is negligible.

Motivations and Incentives

Motivation to contribute

In this section we examine some incentives and disincentives to contribute to Wikipedia.

On average, contributors spend 4.3 hours per week contributing to Wikipedia. There is a wide range of motivations behind people's active engagement in Wikipedia, but two motivations stand out clearly: the wish to share knowledge and the desire to fix errors.

Motivation (max 4 answers, ranked by importance, 1 = most important to 4 = least important)	Mean
I like the idea of sharing knowledge and want to contribute to it	1.68
I saw an error I wanted to fix	1.94
I do it for professional reasons (it belongs to my professional tasks)	2.32
I want to learn new skills / acquire new knowledge	2.46
Because I think information should be freely available to everyone	2.71
Don't know/can't remember/no real reason	2.79
I want to make topics more popular that are not widely known yet	2.80
I think the Internet provides a better medium for encyclopediae than traditional media	2.84

Motivation (max 4 answers, ranked by importance, 1 = most important to 4 = least important)	Mean
I want to demonstrate my knowledge to a wider public / community	2.86
Because I like Wikipedia's philosophy of openness and collaboration	2.88
I saw a red link so I wrote a new article	2.91
To improve my job / career opportunities	2.96
To make the world a better place	2.99
To earn money	3.03
Because friends of mine are doing it and motivated me to join	3.12
Because I like mass collaboration/cooperation	3.15
Just to see if it is really open for anyone to edit	3.18
To gain a reputation in the Wikipedia community	3.30
Total N=43973	

Table 7: Reasons why people contribute to Wikipedia

Reasons for not contributing

The attractiveness of Wikipedia is particularly visible if the non-contributors (i.e. readers) are regarded. Two thirds of the respondents that do not contribute to Wikipedia have at least considered to contribute (e.g. by creating a new article, edit an article or participate in a discussion, etc.), only 26% did not consider this, and 8% replied that they don't know.

The reasons given by those who never considered to contribute (considering the aggregate data) show two main reasons and four less important factors. Not having enough information to contribute or being happy to just read Wikipedia are the two options that have been checked by roughly half of the respondents. Four other options were selected by about a quarter of the respondents respectively.

The most frequently selected options that would make non-contributors contribute are help with identification of content areas that need input, and some form of feedback that would show how contributions would help others.

Answer Option	%
I don't know how	25.18
I am not sufficiently comfortable with the technology	12.49
I would never interact on the internet	4.72
Others are already doing it, there is no need for me	18.88
I don't have time	31.48
I don't feel comfortable editing other peoples' work	26.43
I don't think I have enough information to contribute	51.98
I am afraid of making a mistake and getting "in trouble" for it	24.43
It's a waste of time: my edits would be reverted or overwritten	7.02
I am happy just to read it; I don't need to write it	48.53
Other	3.94
Don't know	3.99
Checked none of the options	*0.93*

Answer Option	%
Total N=21432	

Table 8: Reasons of Non-contributors for not Contributing to Wikipedia

Answer Option	%
Someone would show me how to do it	21.19
The technology was easier to use	8.83
I knew that other contributors would be welcoming and encouraging	15.69
I knew there were specific topic areas that needed my help	42.23
I was confident my contributions would be valued and kept	25.46
It was clear to me that other people would benefit from my efforts	32.95
Other / don't know / don't want to say	31.18
N=21432	

Table 9: Factors that would make contribution more likely

Motivation to donate to the Wikimedia Foundation

In this section we examine some incentives and disincentives to donate to the Wikimedia Foundation.

Regarding the reasons why members of the Wikipedia community do not donate to Wikimedia foundation age structure of the Wikipedia community is a main factor: Many community members are still going to school and have no income that would allow them to donate.

Answer Option	% of respondents
I don't know how to do that	42.03
I did not know it is supported by a non-profit organisation	19.73
I donate my time instead of money	18.88
I never donate to charities	18.75
Donations to the Wikipemedia Foundation are not tax-deductible where I live	16.34
I can't afford to make a donation	15.54
Don't know	13.18
I was never asked to/ wouldn't know how	11.98
I disagree with Wikipedia's policies and practices	9.70
It seems that enough of the other people are making donations to keep the project running	3.84
I do not trust that my donation will be used wisely	1.90
N=130576	

Table 10: Reasons not to donate money to Wikimedia

Another barrier to donate money to Wikipedia results from a lack of knowledge that this opportunity exists or how to donate money. The least important reason for not donating money to Wikipedia is disagreement with Wikipedia's policies and practices.

Lack of knowledge about the legal status of Wikipedia and how to do donations is comparably stronger pronounced among the readers than the contributors. Contributors show only one item that they emphasise much stronger than the average as a reason not to donate money to Wikipedia: they consider the time they spend in contributing to Wikipedia as their donation.

Reason (multiple responses)	Readers	Contributors	All respondents
I can't afford to make a donation	20.66	6.26	15.54
I was never asked to/ wouldn't know how	12.39	11.31	11.98
I don't know how to do that	41.85	43.04	42.03
I donate my time instead of money	9.85	34.62	18.88
It seems that enough of the other people are making donations to keep the project running	3.63	4.10	3.84
I did not know it is supported by a non-profit organisation	23.75	12.84	19.73
I do not trust that my donation will be used wisely	1.12	2.52	1.90
I never donate to charities	21.37	14.25	18.75
Don't know	14.36	10.57	13.18
Donations to the Wikipemedia Foundation are not tax-deductible where I live	15.49	18.10	16.34
I disagree with Wikipedia's policies and practices	10.77	7.99	9.70
	N=81677	N=43973	N=130576

Table 11: Reasons not to donate money to Wikimedia (by activity type)