A report on the Account Creation Improvement Project and the Fellowship

By Lennart Guldbrandsson
Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>Overview of the project and the Fellowship</td>
<td>4</td>
</tr>
<tr>
<td>Timeline</td>
<td>4</td>
</tr>
<tr>
<td>Before the Fellowship</td>
<td>4</td>
</tr>
<tr>
<td>The Fellowship</td>
<td>4</td>
</tr>
<tr>
<td>After the Fellowship</td>
<td>5</td>
</tr>
<tr>
<td>Participants and method</td>
<td>5</td>
</tr>
<tr>
<td>The lay of the land</td>
<td>6</td>
</tr>
<tr>
<td>The original account creation process</td>
<td>6</td>
</tr>
<tr>
<td>The technical aspect</td>
<td>6</td>
</tr>
<tr>
<td>The content aspect</td>
<td>7</td>
</tr>
<tr>
<td>The &quot;before&quot; statistics</td>
<td>8</td>
</tr>
<tr>
<td>The surveys</td>
<td>9</td>
</tr>
<tr>
<td>Results</td>
<td>11</td>
</tr>
<tr>
<td>Comments from people taking the survey</td>
<td>11</td>
</tr>
<tr>
<td>Testing it live</td>
<td>12</td>
</tr>
<tr>
<td>The low-quality tests</td>
<td>12</td>
</tr>
<tr>
<td>Samples of versions that were tested</td>
<td>13</td>
</tr>
<tr>
<td>Results of the first tests</td>
<td>14</td>
</tr>
<tr>
<td>The backlash with version 2</td>
<td>14</td>
</tr>
<tr>
<td>CustomUserSignup</td>
<td>15</td>
</tr>
<tr>
<td>The high-quality tests</td>
<td>16</td>
</tr>
<tr>
<td>The user page creator</td>
<td>16</td>
</tr>
<tr>
<td>Step 0 and 1:</td>
<td>17</td>
</tr>
<tr>
<td>Step 2:</td>
<td>17</td>
</tr>
<tr>
<td>Step 3:</td>
<td>18</td>
</tr>
<tr>
<td>Step 4:</td>
<td>18</td>
</tr>
<tr>
<td>The new user bar template</td>
<td>19</td>
</tr>
<tr>
<td>Category: New Wikipedians</td>
<td>21</td>
</tr>
<tr>
<td>The Sparked.com model</td>
<td>21</td>
</tr>
<tr>
<td>Step 0 and 1:</td>
<td>22</td>
</tr>
<tr>
<td>Step 2:</td>
<td>22</td>
</tr>
<tr>
<td>Step 3:</td>
<td>23</td>
</tr>
<tr>
<td>Step 4:</td>
<td>23</td>
</tr>
<tr>
<td>Results</td>
<td>24</td>
</tr>
<tr>
<td>What we measure and why</td>
<td>24</td>
</tr>
<tr>
<td>The raw data</td>
<td>24</td>
</tr>
<tr>
<td>Crunching the numbers</td>
<td>25</td>
</tr>
<tr>
<td>All account creation processes</td>
<td>25</td>
</tr>
<tr>
<td>ACP1</td>
<td>25</td>
</tr>
<tr>
<td>ACP2</td>
<td>25</td>
</tr>
<tr>
<td>ACP3</td>
<td>26</td>
</tr>
<tr>
<td>Total scores</td>
<td>26</td>
</tr>
<tr>
<td>Which account creation process is better?</td>
<td>26</td>
</tr>
<tr>
<td>Other results and aspects</td>
<td>27</td>
</tr>
<tr>
<td>Final thoughts</td>
<td>28</td>
</tr>
<tr>
<td>Further research</td>
<td>28</td>
</tr>
<tr>
<td>Appendices</td>
<td>29</td>
</tr>
<tr>
<td>Survey 1</td>
<td>29</td>
</tr>
<tr>
<td>Survey 2</td>
<td>30</td>
</tr>
</tbody>
</table>
Interpreting the results.......................................................................................................................... 30
List of MediaWiki messages involved in the account creation process.................................................... 32
Log In / Create Account page.................................................................................................................. 32
User Create Page.................................................................................................................................... 32
Confirmation page................................................................................................................................... 33
Function that checks for available usernames.......................................................................................... 34
The code.................................................................................................................................................. 34
Implementation....................................................................................................................................... 36
Thanks.................................................................................................................................................... 37

This report in a nutshell: More people would start editing if we changed the account creation process.

(Image source)
Introduction

We made a bet after a meeting in a room on Floor 6 in Wikimedia Foundation's offices in San Francisco. It was me and Frank Schulenburg, plus Winifred Olliff. The bet was for two bottles of good champagne. The question was how much we would manage to increase the number of people who started to edit after we had changed the account creation process. For the life of me, I can't remember what Winifred said before we left our notes for safekeeping with Bryony Jones, but I remember distinctly that Frank surprised us by saying that we would go up to 42%. Since the current number was under 30% on most days, that was a bold guess.

But in my hand I held a note with another number on it. It was 47%...

This is the story on who won that bet. Or if you will, a report on the Account Creation Improvement Project and my, i.e. Lennart Guldbrandsson’s, Wikimedia Foundation community Fellowship during the first half of 2011. After reading it, you'll know what the recommendations for the future are about this. You'll also know how to implement the suggested changes on your wiki.

The Account Creation Improvement Project was started in the Wikimedia Foundation community department, by Frank Schulenberg. Lennart's Fellowship first consisted of working on two simultaneous projects: the Bookshelf project and the Account Creation Improvement Project. In the end of May, though, the Bookshelf part ended, allowing Lennart to focus on the Account Creation Improvement Project. During all this period, Lennart has reported regularly to Frank Schulenburg and Zack Exley, as well as to the Wiki/pm/edia community.

During the six months of the Fellowship, Lennart successfully conducted two surveys, led a series of low-quality tests, oversaw the design of a tracking system, and then the development of two high-quality test versions, after which we deployed them, and finally, worked on gathering and analyzing the data after the tests.

Overview of the project and the Fellowship

Timeline

Before the Fellowship
September 18, 2010 – Frank starts Account Creation Improvement Project (Lennart signed on as interested to participate on same day)

The Fellowship
December 16, 2010 – Lennart's Fellowship starts
January 2-9, 2011 – Lennart to San Francisco
January 25-February 4, 2011 - Survey rolled out on six languages
February 2-March 30 – Low-quality tests live on English Wikipedia
March 28 - April 1, 2011 – Lennart to San Francisco, again
April 24, 2011 – Tracking system up and running
June 16, 2011 – Lennart's Fellowship ends
After the Fellowship
July 1, 2011 – Statistics on how many make 5 edits after the account creation tests
July 20-30, 2011 – Report being compiled
August 5, 2011 – Lennart presents Account Creation Improvement Project at Wikimania

Participants and method

For the most part, Lennart has been working with volunteers during the Account Creation Improvement Project. The main workspace has been the Outreach wiki at http://outreach.wikimedia.org/wiki/Account_Creation_Improvement_Project, but some work has been conducted on different language versions of Wikipedia, the Test wiki, and the Mediawiki wiki.

Besides the volunteers, the project has had some Wikimedia Foundation staff support, primarily in the form of ideas and technical support.

For brevity’s sake, this report uses the pronoun “we” to signify the work of either of these three groups of people (Lennart, volunteers, staff), instead of constantly naming who did what - the focus being on the results.

Note: most of the links in this report lead to English Wikipedia. This is to keep things brief. If you change the ”en” in the URL to the ISO code for your languages, all MediaWiki pages, for instance, go to their counterparts on the Wikipedia in that language.
The lay of the land

Before we started to change anything in the account creation process, we wanted to understand what we had. That meant that we had to do a bit of documentation and research. There wasn't much before, at least not collected anywhere.

The original account creation process

The account creation process as it was before the Account Creation Improvement Project started has been created by the a) tech staff of the Wikimedia Foundation, and b) Wikipedia volunteers.

The technical aspect

Since the project had very limited support in terms of full-time tech staff help, the existing account creation process would have to remain relatively unchanged. That was not a major problem. Compared to many other websites, Wikipedia has a rather simple and clear account creation process, with only a few fields to fill in. In contrast to websites such as Facebook, Wikipedia will of course let you use the service (both read and contribute) without an account.

The process to create an account in Wikipedia looks like this:

Steps for account creation:

0: User clicks on the Log in/ create account link in the navigation
1: User is taken to the Login page, which asks the user to either log in or to create an account
2: User is taken to the Account Creation page, which asks the user to create an account by filling in a form
3: User is taken to a Confirmation page
Each of these pages consist of a number of different MediaWiki messages that can be edited by administrators only, and some system messages that only programmers can change. For a complete list of MediaWiki messages that administrators can edit, see the Appendix.

The content aspect

Whereas the technical aspect of the account creation process is controlled by the Wikimedia Foundation staff, the content of the pages in the process is decided by each Wikipedia language community. That kind of high-traffic pages are normally built up based on experiences, such as mistakes made by newcomers and frequently asked questions. On the account creation process for English Wikipedia, this arguably led to instruction creep, and even a process that was unwelcoming and daunting for the newcomers. This is how one of the pages in the original account creation process looked before the project started:

(See Source)
A comparison between Wikipedia and Facebook's account creation processes makes this instruction creep very clear. Facebook's 54 words are compared with Wikipedia's 681 words, including a few warning signs and phrases marked in bolded red!

The "before" statistics

We considered that the existing account creation process was working fine, but needed to have some statistics to back that up. During September 2010, Frank Schulenburg wrote several queries to the Toolserver. The goal was to understand how many created accounts per day, and how many of them started to edit. Here's the result for August 2010:

<table>
<thead>
<tr>
<th>Project</th>
<th>Number of new user accounts overall</th>
<th>New user accounts: average per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Wikipedia</td>
<td>169,706</td>
<td>5,474</td>
</tr>
<tr>
<td>Spanish Wikipedia</td>
<td>27,102</td>
<td>874</td>
</tr>
<tr>
<td>German Wikipedia</td>
<td>17,963</td>
<td>579</td>
</tr>
<tr>
<td>French Wikipedia</td>
<td>16,712</td>
<td>539</td>
</tr>
<tr>
<td>Russian Wikipedia</td>
<td>12,653</td>
<td>408</td>
</tr>
<tr>
<td>Dutch Wikipedia</td>
<td>5,743</td>
<td>185</td>
</tr>
<tr>
<td>Polish Wikipedia</td>
<td>5,393</td>
<td>174</td>
</tr>
<tr>
<td>Swedish Wikipedia</td>
<td>3,992</td>
<td>129</td>
</tr>
<tr>
<td>English Wikinews</td>
<td>2,801</td>
<td>90</td>
</tr>
<tr>
<td>Czech Wikipedia</td>
<td>2,686</td>
<td>87</td>
</tr>
<tr>
<td>Hungarian Wikipedia</td>
<td>2,602</td>
<td>84</td>
</tr>
<tr>
<td>Hindi Wikipedia</td>
<td>1,293</td>
<td>42</td>
</tr>
<tr>
<td>Tamil Wikipedia</td>
<td>807</td>
<td>26</td>
</tr>
<tr>
<td>Bengali Wikipedia</td>
<td>677</td>
<td>22</td>
</tr>
<tr>
<td>Malayalam Wikipedia</td>
<td>616</td>
<td>20</td>
</tr>
</tbody>
</table>

Number of new accounts per month.
<table>
<thead>
<tr>
<th>Project</th>
<th>Number of new user accounts</th>
<th>New accounts with at least one edit within the first 10 days after creation</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Wikipedia</td>
<td>169,706</td>
<td>51,864</td>
<td>31%</td>
</tr>
<tr>
<td>Spanish Wikipedia</td>
<td>27,102</td>
<td>6,769</td>
<td>25%</td>
</tr>
<tr>
<td>German Wikipedia</td>
<td>17,963</td>
<td>6,221</td>
<td>34%</td>
</tr>
<tr>
<td>Russian Wikipedia</td>
<td>12,653</td>
<td>3,335</td>
<td>26%</td>
</tr>
<tr>
<td>Dutch Wikipedia</td>
<td>5,743</td>
<td>1,413</td>
<td>24%</td>
</tr>
<tr>
<td>Polish Wikipedia</td>
<td>5,393</td>
<td>1,255</td>
<td>23%</td>
</tr>
<tr>
<td>Swedish Wikipedia</td>
<td>3,992</td>
<td>1,190</td>
<td>30%</td>
</tr>
<tr>
<td>English Wikinews</td>
<td>2,801</td>
<td>2,798</td>
<td>100%</td>
</tr>
<tr>
<td>Hindi Wikipedia</td>
<td>1,293</td>
<td>121</td>
<td>9%</td>
</tr>
<tr>
<td>Tamil Wikipedia</td>
<td>807</td>
<td>82</td>
<td>10%</td>
</tr>
<tr>
<td>Bengali Wikipedia</td>
<td>677</td>
<td>74</td>
<td>11%</td>
</tr>
<tr>
<td>Malayalam Wikipedia</td>
<td>616</td>
<td>70</td>
<td>11%</td>
</tr>
</tbody>
</table>

Number of new accounts that are used in the first 10 days.

(The rest of the results of the "before" statistics are [here](#).)

Is this result amazing or bad? According to a paper on Facebook, 44.6% of the new users drop off in the first 3 months, but since Facebook's system is very different, and most other websites don't publish data such as this, we concluded that the only data we would compare ourselves to were earlier data, i.e. if we could increase the numbers in the table above.

**The surveys**

Around January 25, we asked administrators on 7 language versions of Wikipedia (English, German, French, Spanish, Swedish, Bengali and Bahasa Melayu) to put a link to a very short survey on the page that new users "land" on after they have created their new accounts. That survey was the next step to understanding what we needed to change. The goal was not to have a statistically perfect survey, but some inkling on what new editors wanted out of their account and what we needed to explain to them. The survey asked them about what they expected when they created the account, or in essence, "why create an account on Wikipedia?".

This is the English version:

**Why did you create your account?**

Answer choices:

- *I wanted to be part of Wikipedia*
- *I was just curious about what would happen if I created an account*
- *I thought an account is required to edit articles*
- *I thought I would get some extra features*
• I usually register for websites I like¹
• I tried to start an article, but I couldn't do that without a user account
• I wanted to upload images
• Other: …

(Source and other language versions)

After a few days, we started a new survey on English, German and Spanish Wikipedia, with the following question:

Congratulations! You just created your own Wikipedia user account. If you had to create a user account again, which of the following statements would convince you more?

Answer choices:

• Having a user account under my real name means that everybody will know how I helped improve Wikipedia.
• Having a user account will enable me to customize the appearance and behavior of Wikipedia to my own preference.
• Having a user account means that I can follow all my favorite topics on Wikipedia.
• Having a user account means that I will get my own user page on Wikipedia. Other users will be able to send me messages.
• Having a user account means that my contributions to Wikipedia are not connected to an IP-address, but to a nickname that can't be tracked back to my physical address.
• Other (please specify)

(Source)

We got roughly 9545 persons in all answering our surveys (the French and the Bahasa Melaya was not started despite being translated). In order from the fewest to the most responses:

• Bengali - 112
• Swedish - 155
• German version 1 - 255
• Polish - 359
• Spanish version 1 - 881
• German version 2 - 932
• Spanish version 2 - 1050
• English version 1 - 1412
• English version 2 - 1200 (roughly²)
• English version 3 - 3189

(Source)

¹ This idea is humorously sketched out as a graph here. This is probably the main reason why people with accounts don't start to edit Wikipedia.
² Due to a mistake, the specific result is swimming around somewhere on the LimeSurvey, but we have a screenshot of the results:
http://upload.wikimedia.org/wikipedia/commons/6/65/Account_creation_survey_2_en.png
Results

The most common answer across all language versions we surveyed was "I want to be part of Wikipedia". In fact, the stable results on that point was the reason why we did another survey to understand what "part of Wikipedia" meant.

“Following favorite articles on Wikipedia” was the most common answer in the second survey, with the notable exceptions of German Wikipedia where anonymity is a bigger reason, and Spanish Wikipedia where getting a user page and being able to send/receive messages are the most important features. When forced to choose one alternative, newcomers to English Wikipedia think that being able to follow articles is the most important reason to create an account. That is also important when they can choose.

Full results of the surveys can be found in the Appendix.

Comments from people taking the survey

In the first and second surveys, we gave people a free text area. People used that option a lot. The most common responses were along the lines of:

- "I don't want to edit articles anonymously."
- "Class assignment"
- "want to create a book I can same and print" [or other function related answers]
- "I LOVE WIKIPEDIA"

Some gave responses that showed that we still have some outreach work to do:

- "i thought i would be able to reazd alot more arlicals”
- "chat with online friends”
- "To give my forum a page”

And some gave cryptic responses:

- "I wanted to vandalize!”
- "Am an artist”
- "Dunno. I was bored."
Testing it live

During Lennart's trip to San Francisco in January, we decided that surveys and statistics are but the first step, and that we needed to test out different things live on Wikipedia.

**The low-quality tests**

Step 1 in the testing process was to gather as many different ideas as possible without too much thought into design, deploy them and measure how many new users who started editing. We called the “low-quality tests”, in contrast to the more polished versions that we would work on later, once we knew more about what worked and what didn't. We were certainly aware that the versions in many cases were crude and simplistic, but the tests could show which we should work more on.

Starting February 10, the community was invited to contribute their designs of the three pages in the account creation process: the login page, the account creation page and the confirmation page. To help people out, we created a list of assumptions that they could use when designing their own versions. For example, we believed that showing the newcomers images of people may inspire more confidence. Read [the full list here](#). Among the suggestions were one that was used on Polish Wikipedia, where the newcomer gets the “rulebook” on their user page. Other designs had videos and links to reading materials.

On February 23, the first test were performed on English Wikipedia. In all, 10 low-quality designs were tested, the last test ending on March 30. Most of the tests were focused on the last page in the account creation process, the confirmation page. The original version looked like this:

![Welcome to Wikipedia, $1!](Image source)

This version became known in the project as “Commons sallad”, because of all the images and extra ingredients.

---

3 This version became known in the project as “Commons sallad”, because of all the images and extra ingredients.
Samples of versions that were tested:

(version 2)

(version 3)

(version 4)

(version 7)

(version 9)
## Results of the first tests

<table>
<thead>
<tr>
<th>Test version</th>
<th>Number of new accounts (best day)</th>
<th>New accounts with at least one edit on the day of account creation (best day)</th>
<th>Best percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>7,180</td>
<td>2,177</td>
<td>30%(^4)</td>
</tr>
<tr>
<td>1</td>
<td>7,341</td>
<td>2,151</td>
<td>33%</td>
</tr>
<tr>
<td>2</td>
<td>6,574</td>
<td>2,328</td>
<td>35%</td>
</tr>
<tr>
<td>3</td>
<td>7,266</td>
<td>2,364</td>
<td>32%</td>
</tr>
<tr>
<td>4</td>
<td>7,267</td>
<td>2,433</td>
<td>33%</td>
</tr>
<tr>
<td>5</td>
<td>6,935</td>
<td>2,447</td>
<td>35%</td>
</tr>
<tr>
<td>6</td>
<td>6,080</td>
<td>1,917</td>
<td>32%</td>
</tr>
<tr>
<td>7</td>
<td>7,151</td>
<td>2,183</td>
<td>31%</td>
</tr>
<tr>
<td>8</td>
<td>6,066</td>
<td>1,879</td>
<td>31%</td>
</tr>
<tr>
<td>9</td>
<td>6,697</td>
<td>2,124</td>
<td>32%</td>
</tr>
<tr>
<td>10</td>
<td>6448</td>
<td>1978</td>
<td>31%</td>
</tr>
</tbody>
</table>

(Source and all results)

Version 2 and 5 had a few percentages more people who started to edit than the others, but the remarkable thing is that all versions were better than the original one. Version 2 and 5 gave 5 percent units more newcomers than the standard one, which on a day with 6000 new accounts means 300 more potential Wikipedia contributors – per day! These versions were certainly worth exploring more thoroughly.

### The backlash with version 2

We announced the changes before we implemented them, on the Village pump of English Wikipedia, as well as the Administrators’ Noticeboard. Most who responded were interested in the tests, but the activity was not very high.

However, when we rolled out version 2 (see picture above) there was outcries in several places, and the page was reverted. To understand the reason why, let's look at how version 2 works, since this will become important later on. Never mind that the design lacks a certain style. The function was proposed by editors from Polish Wikipedia, based on the system that they had there. Once you have created an account, you're given one clear choice: to create a user page. When you click on the button, a template preloads information on your user page and all you have to do is click the “save” button. The Polish version had a list of around 30 links to guides, policies, and other pages on Wikipedia. The English version looked like this.

What the community reacted to was that so many new users created user pages that looked exactly the same. This function made the newcomers’ user pages blue links instead of red links on for instance Recent changes, also made it more difficult to distinguish between newcomers (who according to Wikipedia wisdom are more prone to vandalism) and Wikipedians (who are less prone to vandalism).

\(^4\) This percentage here deviates a little bit from the percentage in the original statistics.
However, we reasoned that it could be a good thing if Wikipedians who patrolled Recent changes didn’t automatically scrutinize new users harder. A test on English Wikipedia using “under cover” Wikipedians posing as newcomers noticed that edits they made were more often reverted than if they made the same type of edits with their usual accounts. This made us think harder about how we could have the best of both worlds: tools for the vandalism patrol on the one hand, and an easier time getting accepted as a Wikipedian on the other.

**CustomUserSignup**

To get a better measuring rod than statistics from the Toolserver accounts – which needed to be updated manually for each update – the tech department were given the task of creating a tracking system with A/B testing capabilities as described here.

In essence, what we wanted was a method to randomly send those who wanted to create an account one of several alternate routes, like this:

![Diagram showing alternate routes for account creation](Source)

Also, we wanted track how many went through the different routes – and most importantly, how many of them began to edit after creating their account. This basically meant, which route was the most effective way of getting people to edit?

That task was done by an extension to the MediaWiki software, which placed a tracking cookie in the web browsers of the new users. The tech team wrote:
What data we are storing
We are storing a new cookie upon visiting the “Log in/create account” page, with a lifetime of three months. This cookie will be used to track the following information:

- Which account creation messaging group the user was placed in (identified as ACP1, ACP2 or ACP3 for now)
- What version of the account creation campaign they received
- Whether the particular user made it to the end of the account creation process, or whether they dropped off after reaching the login screen or the account creation screen
- If (and only if) the user creates a new account, the number of edits or previews during the course of the trial

The information is associated with browser sessions (each of which has an individual unique identifier), not with an individual user or user account.

(Source)

The extension was deployed on April 27, which is when the more advanced statistics begin.

The tech team theorized that more outreach activities could use the CustomUserSignup extension to gauge how successful the activity was in getting more people to sign up as editors.\(^5\)

The high-quality tests

After the second trip to San Francisco, Lennart and Frank set out to create two high-quality account creation processes. They were different from the low-quality tests in two fundamental ways:

1) they were created with more style in mind, over a longer period of time than the low-quality designs, which among other things meant that we designed whole processes\(^6\) than individual pages
2) they were deployed with the CustomUserSignup extension in play, so that we could measure what happened to the people who went through the account creation process

The two new high-quality designs were based on the two most successful versions in the low-quality tests. Internally we called those versions “the user page creator” and “the Sparked.com model” (more details below). Concurrent with them, we would run a third process. The third one was the existing process, which we wanted to use as a baseline, in case we suddenly saw a spike or dive in the statistics of the new version.

The work to design and code the two high-quality versions proved to be harder than anticipated. There were technical challenges as well as content issues that needed to be addressed before they could be rolled out on English Wikipedia. Fortunately, even though Lennart was responsible for the user page creator and Frank took charge of the Sparked.com model, we had plenty of help from the communities of the Outreach wiki and other Wiki/pm/edians.

The user page creator

The user page creator was based on version 2 of the low-quality tests — the controversial one. However, we felt that once we could provide the vandalism patrollers some tools to deal with the

5 This can be made easier by using such services as QRpedia, a website which turns URL:s into codes that can be read by mobile cameras.
6 For instance, we placed a progress bar at the top of both account creation processes.
problems that the model created, this model would have a good chance to get many newcomers to start editing, which was the goal. It would also present the newcomer to the rest of the community, which we hypothesized would make the Wikipedia community more inclined to welcome them into the fold. The various tools and features will be explained as we go through the steps in the process.

This process was designated `campaign=ACP1` in the CustomUserSignup extension, which meant that you can go directly to this account creation process (instead of getting randomly assigned to one of the three account creation processes) by clicking this link:

http://en.wikipedia.org/w/index.php?
title=Special:UserLogin&action=submitlogin&type=signup&campaign=ACP1

This is how it looks:

**Step 0 and 1:**
Step 0 and 1 in the account creation process are the same as the normal one. This is also true for the Sparked.com model.

**Step 2:**

You'll note the difference to the original one: shorter text, the softened frame with shadows and the colorful progress bar at the top.

We should also mention that much of the page consists of a form that we could not change, for technical reasons. It would take too much time for a six-month project, and with very unsure results. The original design called for an inspirational “You can also become a part of Wikipedia” sidebar, but we were unable to get this to work with the form part without changing the software.
Here, there are two main new features, beside the progress bar:

- the blue box to the right, which showcases an example of a user page\(^7\) (including a photo of a Wikipedian)
- the “tell us about yourself” form\(^8\)

The “tell us about yourself” form is a JavaScript-created function that is prefilled with the following text:

Replace this example text below with information about you:

Hello, my background is in biology, with a main interest in snakes.

I speak English and French. In my off-time, I listen to a lot of music, and I have discovered that Wikipedia is a very good source for information in that department. Hopefully I can help make it even better.\(^9\)

This text is one of the tools we gave the vandalism patrollers: they can now check for user pages that include this exact wording.

**Step 4:**

Normally, step 3 is where we let the newcomer find his or her own way. In this version, we added a fourth step. Once you click “Create my user page for me now”, you will be taken to your new user page, with the text you added pasted into the edit area, so that all you have to do is to scroll down and click “Save”. This looks like this:

---

\(^7\) The user page belongs to Tim Vickers, after a thorough search among user pages for a simple, informational user page that a newcomer could imitate. It was a happy coincidence that the user turned out to be a scientist.

\(^8\) My Swedish web browser underlines the words it doesn't understand, even after I switched it to English. In other words, it shouldn't look like that.

\(^9\) This text is generated by one part of the code on this page: [http://en.wikipedia.org/wiki/MediaWiki:Common.js](http://en.wikipedia.org/wiki/MediaWiki:Common.js), which you also have to change if you are to import this version of the account creation process to other wikis.
There are a couple of things to note here:\n\n- there are two “hidden” comments, reminding the user to click “Save” and welcoming the user to Wikipedia
- the edit summary automatically adds the text “New user page through Outreach:ACIP” to this edit in Recent changes, which makes it very easy to spot for vandalism controllers
- this is not the first thing the user sees after coming to this page. Above this is something quite interesting, see the next section.

The new user bar template
One thing that we knew that the new users needed, and we wanted to give them, was information about how Wikipedia works. But we needed to do it without giving them a list of 30 guidelines and manuals on their user page, as in the Polish version. Our idea was to use tabs, i.e. place the information they needed close at hand. The method we chose was to use a template that could be pasted onto every new user page during the account creation process.

That template is expanded on the user page to give the new user information and guidance, without taking up too much space on the page. You can actually see the code that produces the tabs template in the image above, right after the first line of hidden text: \n\n```
{{New user bar}}
```

If you didn't notice it the first time you saw the picture – good. That's the point.

Let's take a look at the template that sits on every new user page created through this account creation process:

\n\n10 This page cannot be easily sourced, since it's session-based. So you'll have to try it yourself.
11 Besides the red-link problem I told you about before.
One feature we have in the New user bar template is the prominent tab “Getting started”, which links to a page we created, called Starting editing. This page is a very short introduction to editing Wikipedia – much shorter, less technical and more focused on inspiration than almost any other manual on Wikipedia:

---

12 The links in the tab are: About me, Getting started, My test area, Get help and Inbox.
But back to the New user bar template: Note also the placeholder image to the right, with a link to a special page about [uploading images of yourself](#). We reasoned that it would be easier for new users to enter the community if their user pages had pictures – and perhaps increase their chance of avoiding frivolous reverting by vandalism patrollers.

At this stage, you can still edit the text in the edit area, since the first edit is when you click “Save” at the bottom of the page – and this is one of the weak spots in this approach. We assume that many newcomers believe that the process is finished by this stage, that they shouldn't have to “create their user page twice”, so to speak, although we have no direct data, such as interviews, to support this hypothesis.

*Category:*New Wikipedians

One other tool that we wanted to give vandalism patrollers is a quick way to find as many new Wikipedians as possible, so in the code for the New user bar template, we added code that automatically places anyone who has the New user bar template on their user page in a special category for new Wikipedians.

### The Sparked.com model

The Sparked.com was based on version 5 of the low-quality tests – essentially a box divided into four sections with the text “We need your help in making Wikipedia better – [here are some tasks where you can help out](#). It's easy!”. But the name came from a website for microvolunteering, unsurprisingly called Sparked.com, especially starting on page 2: “Which causes really get you fired up?” Once the new user's chosen at least one cause he or she's on to a page where another question is asked: “Which skills do you have to offer?” The last step matches a cause with the chosen skill to produce a list of tasks where the new user could volunteer. That list should ideally contain at least some articles that the new user finds interesting, much like Amazon.com's Recommendations include some articles that the person would perhaps want to buy.

This idea was conceived for new users who didn't know what tasks still needed to be done on Wikipedia (some readers believe Wikipedia to be finished, partly because they go to a highly visible page like World war II, and never see the stubs and the other problematic pages). Once we can direct new users to articles in need, we can perhaps inspire the new users to start editing.

We have similar lists on Wikipedia as Sparked.com have, but most of them are too unsorted and long to be practical for newcomers. That's why Frank first had to create a list of problematic areas that we could reasonably expect newcomers to contribute - and a set of topics that could inspire people. We decided on these four skills and these six topics:

- History, Biology, Technology, Geography, Mathematics, and Arts
- Copyediting, Research & Writing, Searching the web, and Organizing

---

13 And indeed in Wikipedia itself: a more prominent “Save” button than below the edit area could potentially make it easier for some people to find it.

14 One difference between our version and Sparked.com's version is that you can only pick one skill and one topic. Part of the reason that our version don't have that is technical, as it would take more software work than we had time for. A difference between our version and Amazon.com's version is that their version are sent out regularly via email too, based on previous purchases. On Wikipedia we have [SuggestBot](#) that functions in much the same way, albeit not by sending out emails – yet.

15 Asking newcomers for instance to [wikify](#) an article is probably an exercise in futility.

16 We got a suggestion to change these to Arts & entertainment, People, Places, Science & technology, History, and Current events, and we would have liked to tested that version as well. However, there was no time.
These lists have been manually created and would have to be manually updated, as the tasks are being dealt with, but with crowdsourcing, this should be manageable.

This process was designated `campaign=ACP2` in the CustomUserSignup extension, which meant that you can go directly to this account creation process (instead of getting randomly assigned to one of the three account creation processes) by clicking this link:  

This is how the process looks:

**Step 0 and 1:**  
Step 0 and 1 in the [account creation process](#) are the same as the normal one.

**Step 2:**  
**Tell us what you are interested in**

The header was consciously chosen to project the feeling that we want to know more about the new user's interests.

By making the newcomer choose a topic during the account creation process, we hoped to focus them on contributing to Wikipedia, instead of reading only. We also deliberately used pictures as links on this page, to give the image of Wikipedia as “more than just text”.\(^{17}\)

Finally, note the blue box with instructions on the right.

---

\(^{17}\) One comment we have received about these versions is that they are not 100% compatible with the CC-by-sa license, since we haven't provided information about the sources for the images. We chose not to, since it would detract from the “straight” flow from account creation to starting to edit. The annual fundraiser have had the same discussions.
Step 3:
Which skills do you have to offer? (History)

1. Get a new user account
2. Learn what you can do
3. Be a part of Wikipedia

Copyediting
Research & Writing
Searching the web
Organizing

Let us know which skills you have to offer so we can suggest how you can get involved right now.
Click on one of the skills to learn how you can help.

(Source)

Here we have tried to focus the new user into one of four tasks that we need volunteers to help out with. The blue box is present again with instructions.

Step 4:
Be a part of Wikipedia (History, Copyediting)

Get a new user account
Learn what you can do
Be a part of Wikipedia

List of history articles that need copyediting:
- Aftermath of World War II
- Alexander's Conflict with the Kombjas
- Ancient History of Cyprus
- Asim ibn Abdullah and Muhammad ibn Abd Allah
- BBC History of World War II
- Colonialism
- Contents of the United States diplomatic cables leak (India)
- Contents of the United States diplomatic cables leak (Pakistan)
- Geomatics
- Kontoks
- Lund Kimor
- Maiden Abduction from Vietta
- Mannin dynasty
- Michael Joe Costello
- Migration Period
- Mikhail Miloradovitch
- Mongol Invasion of Europe
- Muslim conquest of Egypt
- Nguyen Khoth
- Paolo Selen Amphitheater
- Pedro Fernández de Castro (twelfth century)
- Prehistory of Taroyvanlia
- Quantitative revolution
- Rua Kenana Hesetip
- Running in Ancient Greece
- Sarmatia
- Timeline of Cherokee removal
- Union of Brittany and France

Here is how you can help:
Click on one of the links in the list on the left hand side ("List of history articles that need copyediting"). This will bring you directly to the article that needs attention.
Then, click on "Edit" at the top of the screen (next to the search field):
Read
Edit
View history
Copyedit the article.
Finally, enter a description of what you have done in the "Edit summary" field and click on the "Save" button. Wasn't that easy? Thanks for helping to improve Wikipedia!

(Source)
This is one of the 30 pages with similar lists based on skill and interest. All of them have been manually created, which is one of the drawbacks in this method. We have been in preliminary contact with WikiProjects that could help out with updating these lists.

**Results**

We deployed the three account creation processes (ACP1 – the user page creator, ACP2 – the Sparked.com model, and ACP3 – the standard one) on June 9, 2011. The first results came the next day, and they were very encouraging in some respects.

**What we measure and why**

The data we sought was how many start to edit after going through a particular account creation process. If for instance, ACP2 – the Sparked.com model - manages to attract more editors in the long run than any of the other processes, and the numbers are conclusive, we should implement that model on English Wikipedia. An important angle we looked at was how many make 1 edit versus how many make 5 edits. This distinction is crucial, because we want the newcomers to stay as editors, and not just make a random edit to test out Wikipedia.

As we stated above, the data is built upon a browser-based cookie. That means that if a person creates an account on her own computer and then starts editing with her new account on another computer, she would not be counted here. Currently, we have no method of getting data for those persons, but on the other hand, they will be present in approximately the same numbers in the three different account creation processes, so in the end it will still be fair.

**The raw data**

The tech department made an automatic script on the Toolserver which updated daily on the statistics of the three account creation processes. It deposits the raw data numbers in this text file. The data there is a little hard to understand, so here is a short legend:

---

18 We haven't run these tests on any other Wikipedia. Our intention at the beginning was to have these kinds of statistics for several language versions of Wikipedia, but time constraints became a big factor. That's why we hope that volunteers make these experiments on their own wikis in the future.
What is says in the text file – what it means
/home/catrope/acp/perday/lennart-log-20110427-ACP1-final – date and ACP (in this case it's ACP1, the user page creator)\textsuperscript{19}
WITH EDITS: 45 – of everyone who went through this ACP, how many made 1 edit?
WITH 5 EDITS: 9 – how many made 5 edits?
TOTAL: 110 – how many can we match to a specific user name? (ignore this)
TOTAL WITH UNMATCHED: 119 – how many created an account through this process?
Percentage with 1 edit: 37.8151 – how big share of everyone who went through this ACP made 1 edit?\textsuperscript{20}
Percentage with 5 edits: 7.56303 – how big share of everyone who went through this ACP made 5 edits?

Crunching the numbers

Sorting through the raw data, we can find these numbers:

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
 & With 1 edit & With 5 edits & Accounts created & Percentage with 1 edit & Percentage with 5 edits \\
\hline
Average day & 337.87 & 63.96 & 967.38 & 34.83 & 6.90 \\
Best day & 533 & 103 & 1166 & 48.93 & 10.61 \\
Worst day & 216 & 28 & 733 & 25.58 & 3.58 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
 & With 1 edit & With 5 edits & Accounts created & Percentage with 1 edit & Percentage with 5 edits \\
\hline
Average day & 437.66 & 69.90 & 907.49 & 44.58 & 7.02 \\
Best day & 533 & 103 & 1059 & 48.93 & 10.61 \\
Worst day & 336 & 36 & 733 & 40.97 & 4.33 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
 & With 1 edit & With 5 edits & Accounts created & Percentage with 1 edit & Percentage with 5 edits \\
\hline
Average day & 287.61 & 63.07 & 881.27 & 30.96 & 6.66 \\
Best day & 382 & 96 & 1062 & 34.98 & 9.06 \\
Worst day & 216 & 28 & 703 & 25.58 & 3.58 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{19} In the text file, the first day is April 27, which is when the extension was deployed. But we didn't roll out the three different account creation process until June 9. The data before that is based on three exactly alike account creation processes. We have therefore excluded any data from before June 10 – the first full day of testing. The last day included is July 20.

\textsuperscript{20} Remember, the original statistics said that 30% (on the best days) create accounts and then make one edit. This method of measuring is a little bit different still, so any numbers will deviate from the original statistics. The data in the legend is much higher but it seems that it was a peak day. The normal level is much closer to the original statistics.
Total scores
In total there were **118988 accounts created** during the test period. Of these **41558 made at least 1 edit**, and **7867 made at least 5 edits**.

Which account creation process is better?
To determine which account creation process is better, let's look at the important aspects:

- it has the highest percentage of people that make at least 1 edit on the best day
- it has the highest percentage of people that make at least 5 edits on the best day
- it performs well even on the “worst day” on how many percent make at least 1 edit
- it performs well even on the “worst day” on how many percent make at least 5 edits
- it has a high percent average of people that make at least 1 edit
- it has a high percent average of people that make at least 5 edits

The one that's best in the most of these is clearly better at the job. Here are the scores for the three account creation process compared on these five aspects. The winner of each category is bolded.

<table>
<thead>
<tr>
<th></th>
<th>ACP1</th>
<th>ACP2</th>
<th>ACP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest 1 edit</td>
<td><strong>48.93</strong></td>
<td>34.98</td>
<td>33.49</td>
</tr>
<tr>
<td>Highest 5 edits</td>
<td><strong>10.61</strong></td>
<td>9.06</td>
<td>7.84</td>
</tr>
<tr>
<td>Worst day 1 edit</td>
<td><strong>40.97</strong></td>
<td>25.58</td>
<td>26.79</td>
</tr>
<tr>
<td>Worst day 5 edits</td>
<td><strong>4.33</strong></td>
<td>3.58</td>
<td>3.91</td>
</tr>
<tr>
<td>Average 1 edit</td>
<td><strong>44.58</strong></td>
<td>30.96</td>
<td>29.54</td>
</tr>
<tr>
<td>Average 5 edits</td>
<td><strong>7.02</strong></td>
<td>6.66</td>
<td>6.12</td>
</tr>
</tbody>
</table>

These figures may seem very abstract, but by increasing the inflow of new Wikipedians who make 5 edits from 6.12% to 7.02 on an average day, we gain 9 persons per day! (From 59.90 with the present model) to 68.90 persons (with ACP1).

To sum it up. **There is no other account creation process of the ones that we have tested that can measure up – in any category – to that of the user page creator. This is the account creation process that we recommend.**

And, should be added – that version also scored above **47%**. So unless Winifred's note said 48%, 49% or 50%, those two champagne bottles are mine.
Other results and aspects

However, while the statistics are an important part of choosing a new account creation process, there are other concerns and aspects to keep in mind. To those six categories above, we should add that the account creation process we choose should be effective in terms of community effort. The ACP2/the Sparked.com model has an important drawback in that the community would have to keep the lists of tasks constantly updated.

The ACP1 and ACP2 has both been complimented on looking nice, which makes it easier to convince the community to accept them as the default account creation process.

That does not mean that ACP1 is without its challenges. Here are some of the things we need to take note of, should we decide to go with ACP1:

- we should ask the community to keep an eye on the category for new Wikipedians, and use it as a place to recruit new members of the WikiProjects based on their descriptions of themselves
- some Wikipedians have expressed concern that the ACP1 increases the number of pictures of people on Wikimedia Commons – and since Commons is not a photo album, but the pictures should have educational value, they are likely to be deleted unless some thinking is done on how to approach this
- we would need to discuss with the community about the issue of users with formulaic user pages, many of whom are inactive.
- there have been some talk about browser-compatibility issues, especially when it comes to the JavaScript-based “tell us about yourself”-form. Those would have to be resolved
- this is a big one: when we ask the newcomers to add information about themselves, we need to be very careful on what types of information we ask them for. Several Wikipedians have extremely hard objections to including data such as name, where you live, what age you are and other identifying information. Wiki-stalking is one reason. Another is that if a new user adds a university degree or another expertise to their profile, this may lead that person to expect special treatment
- the deployment should be announced on both the Village pump (proposal) and the Administrators' noticeboard, plus the Wikimedia Foundation blog, with enough numbers to convince people
- other language versions of Wikipedia may discover that their numbers differ from these. Before they deploy ACP1 uncritically, they should test this for a month and measure the difference. Of note is that some Wikipedias already have other account creation processes, that could be fine-tuned as well

As some Wikimedia Foundation staffers have noted, the ACP1 is an amalgamation of several pieces – the user page creator, the new user bar, the new Wikipedians category, etc. If there is time, it would probably be a good idea to test the various part further. We may yet increase the numbers to above 50%!
Final thoughts

Although we didn't intend for it, the new user bar have already started to make its way around Wikipedia. On Chinese Wikipedia, some have begun to use it on their user pages, see for example: http://zh.wikipedia.org/wiki/User:Tommyang.

The extension CustomUserSignup has already been used for a special outreach project: http://hciresearch2.hcii.cs.cmu.edu/~rfarzan/wikipedia/tool/wssignup.php. We should use it more.

While we ended up with two designs, 19 different designs were collected. Some of those who contributed were very skillful and their talents should be harnessed.

Wikipedia's account creation process is really only technically lacking in one aspect only: it does not use AJAX for when you choose a username, which means that if that name is taken, the new user may have to re-do the entire form, instead of getting a warning. This function exists on Swedish Wikipedia, see Appendix.

If you are thinking about implementing any of these account creation processes, there is a guide for it here: http://outreach.wikimedia.org/wiki/User:Hannibal/guide_for_the_next_person. That guide will probably be moved somewhere else, and be continuously updated, as more challenges are discovered and new hacks are being implemented.

Further research

Of course, there is plenty left to do. Some of the more important things are these:

- "how do we increase the number of people who create accounts?"
- "how do we increase the number of people who start to edit?"
- "how do we make sure they continue to edit?"  

We should also develop the email that is sent to anyone who creates an account, and include parts that use the same types of algorithms as SuggestBot and Amazon.com. We should also think about some sort of “We miss you” email to accounts that have been inactive for a period of time.

There have been some talk about adding a flag for new users, so that newcomers are easily spotted in Recent changes. That is a complicated topic that should be brought up on Wikimania and other meetups between Wikipedians.

More about creating formulaic user pages: Structured profile

During the Summer of Code of 2011, Akshay Agarwal worked on the Account Creation code.

---

21 In my experience, some Wikipedians aren't sure they want a big influx of new Wikipedians, since it takes a lot of effort to train them to become Wikipedians. If they did want them, they would treat them better. This is an important factor in this change in direction to attract more Wikipedia editors.
## Appendices

### Survey 1

<table>
<thead>
<tr>
<th>Language version</th>
<th>Alternative 1: I wanted to be part of Wikipedia</th>
<th>Alternative 2: I was just curious about what would happen if I created an account</th>
<th>Alternative 3: I thought an account is required to edit articles</th>
<th>Alternative 4: I thought I would get some extra features</th>
<th>Alternative 5: I usually register for websites I like</th>
<th>Alternative 6: I tried to start an article, but I couldn’t do that without a user account</th>
<th>Alternative 7: I wanted to upload images</th>
<th>Alternative 8: Other</th>
<th>People responding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bengali</td>
<td>71.8% (79)</td>
<td>24.5% (27)</td>
<td>40.0% (44)</td>
<td>25.5% (28)</td>
<td>30.9% (34)</td>
<td>24.5% (27)</td>
<td>19.1 (21)</td>
<td>15.5% (17)</td>
<td>112</td>
</tr>
<tr>
<td>Swedish</td>
<td>60.4% (93)</td>
<td>17.5% (27)</td>
<td>21.4% (33)</td>
<td>9.1% (14)</td>
<td>24.7% (38)</td>
<td>22.7% (35)</td>
<td>9.1% (14)</td>
<td>11.7% (18)</td>
<td>155</td>
</tr>
<tr>
<td>German</td>
<td>66.7% (168)</td>
<td>6.7% (17)</td>
<td>28.6% (72)</td>
<td>6.7% (17)</td>
<td>17.5% (41)</td>
<td>15.9% (40)</td>
<td>3.6% (9)</td>
<td>15.9% (40)</td>
<td>255</td>
</tr>
<tr>
<td>Polish</td>
<td>45.5% (161)</td>
<td>13.8% (49)</td>
<td>17.8% (63)</td>
<td>20.1% (58)</td>
<td>16.4% (58)</td>
<td>30.2% (107)</td>
<td>7.9% (28)</td>
<td>13.0% (46)</td>
<td>359</td>
</tr>
<tr>
<td>Spanish</td>
<td>71.6% (618)</td>
<td>12.2% (105)</td>
<td>28.3% (244)</td>
<td>21.8% (188)</td>
<td>35.5% (306)</td>
<td>15.8% (136)</td>
<td>9.0% (78)</td>
<td>11.0% (95)</td>
<td>881</td>
</tr>
<tr>
<td>English</td>
<td>68.1% (952)</td>
<td>18.0% (252)</td>
<td>27.1% (379)</td>
<td>22.8% (319)</td>
<td>23.9% (334)</td>
<td>18.9% (264)</td>
<td>15.6% (218)</td>
<td>16.1% (225)</td>
<td>1412</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>64.02%</strong></td>
<td><strong>15.45%</strong></td>
<td><strong>27.77%</strong></td>
<td><strong>17.67%</strong></td>
<td><strong>24.82%</strong></td>
<td><strong>21.33%</strong></td>
<td><strong>10.72%</strong></td>
<td><strong>13.87%</strong></td>
<td><strong>Total: 3174</strong></td>
</tr>
</tbody>
</table>

22 In many language versions you can create articles without accounts, so any answers here are significant of what the new users think they need.
### Survey 2

<table>
<thead>
<tr>
<th>Language version</th>
<th>Alternative 1</th>
<th>Alternative 2</th>
<th>Alternative 3</th>
<th>Alternative 4</th>
<th>Alternative 5</th>
<th>Other/Comments</th>
<th>People responding</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>45.9% (406)</td>
<td>30.3% (268)</td>
<td>24.7% (219)</td>
<td>32.0% (283)</td>
<td>53.0% (469)</td>
<td>(101)[3]</td>
<td>932</td>
</tr>
<tr>
<td>Spanish</td>
<td>34.3% (347)</td>
<td>31.8% (322)</td>
<td>60.8% (615)</td>
<td>36.5% (369)</td>
<td>27.7% (280)</td>
<td>(129)[4]</td>
<td>1050</td>
</tr>
<tr>
<td>English variant 1</td>
<td>21% (244)</td>
<td>13% (151)</td>
<td>13% (151)</td>
<td>36% (408)</td>
<td>15% (175)</td>
<td>1% (13)</td>
<td>1200</td>
</tr>
<tr>
<td>English variant 2</td>
<td>45.16% (1459)</td>
<td>42.25% (1365)</td>
<td>41.91% (1354)</td>
<td>60.04% (1940)</td>
<td>38.13% (1232)</td>
<td>8.08% (261)</td>
<td>3189</td>
</tr>
<tr>
<td>Total</td>
<td>36.59%</td>
<td>29.34%</td>
<td>35.10%</td>
<td>41.13%</td>
<td>33.46%</td>
<td></td>
<td>6371</td>
</tr>
</tbody>
</table>

### Interpreting the results

What conclusions can we draw from these survey results? I think there are a number of conclusions that are interesting.

- The number of participants over a short period (volunteered by clicking on an extra link after creating an account) makes the numbers relatively significant. Of course a longer survey would mean more certainty, but the stable results across the survey period suggest that we would get a variation of these results. But I am very happy that this many took the time to answer the questions we put to them. Perhaps the short surveys (only one question at a time) is key here. Also, we can very soon get big groups of people to answer our surveys, which I feel we should use in the future.

- The most common answer across all language versions we surveyed was "I want to be part of Wikipedia". In fact, the stable results on that point was the reason why we did another survey to understand what "part of Wikipedia" meant.

- "I thought an account is required to edit articles" and "I usually register for websites I like" are the next most common answers. "Usually register" probably means that they do not intend to edit very much, while the other is the opposite.

- Not many create accounts because they want to upload pictures.
• A fair share of people thought they needed accounts to create articles (which is true for some but not all Wikipedias). On German and Spanish Wikipedia only about 15% thought that.

• About a sixth of the new users thought they would get some more features. Whether they know about the features or not is not certain. That's why we added the alternative in the second survey about the watchlists ("follow articles"). On Swedish and German Wikipedia, this was not a big deal.

• Following favorite articles on Wikipedia was the most common answer in the second survey, with the notable exceptions of German Wikipedia where anonymity is a bigger reason, and Spanish Wikipedia where getting a user page and being able to send/receive messages are the most important features.

• Not many Germans were curious about what would happen when they created their accounts. Maybe they had enough information already.

• When forced to choose one alternative, newcomers to English Wikipedia think that being able to follow articles is the most important reason to create an account. That is also important when they can choose.

• When not forced to choose one alternative, newcomers to English Wikipedia more than the others that we surveyed like to tinker with the look and feel of Wikipedia.
List of MediaWiki messages involved in the account creation process

This is a list of MediaWiki messages (pages) that administrators can edit, and that are involved in the account creation process:

Log In / Create Account page

- [http://en.wikipedia.org/wiki/MediaWiki:Loginerror](http://en.wikipedia.org/wiki/MediaWiki:Loginerror) Default Value: “Login error” (just the title of the error, when you fail to log in correctly)
- [http://en.wikipedia.org/wiki/MediaWiki:Loginend](http://en.wikipedia.org/wiki/MediaWiki:Loginend) Default Value: Appears to contain all the “Secure your login” text. (in the loginend div, just below the login and email new password buttons)

The page that are assembled by these messages can be seen here: [http://en.wikipedia.org/w/index.php?title=Special:UserLogin](http://en.wikipedia.org/w/index.php?title=Special:UserLogin)

User Create Page

- [http://en.wikipedia.org/wiki/MediaWiki:Prefs-help-email](http://en.wikipedia.org/wiki/MediaWiki:Prefs-help-email) Default Value: Starts with “* You do not have to provide an e-mail address, but if you forget your password”
The page that are assembled by these messages can be seen here: http://en.wikipedia.org/w/index.php?title=Special:UserLogin&type=signup.

Note: if you go to that URL while you are logged in, that page will also contain a field asking for a reason (for why you are creating another account).

Confirmation page
This page consists of: http://en.wikipedia.org/wiki/MediaWiki:Welcomecreation
Function that checks for available usernames

This JavaScript checks for available usernames, and gives an error message if you try to create a new account with a username that is already taken. It was written by User:Fluff on Swedish Wikipedia.

The code

```javascript
/* Making AJAX calls when a user wants to register a new username ***********************
* Description: Displays a text noticing the user if the target username is available or not.
* Maintainers: [[User:Fluff]]
* Link for stats: [[User:Fluff/cfau.js]]
*/
var cfauTimer = false;

function CFAUloader () {
    if(!document.getElementById('wpName2')) {
        return;
    }

    if((wgPageName == "Special:Inloggning" || wgPageName == "Special:UserLogin")
        &&
        (document.getElementById('wpName2').className == "loginText")) {

        addHandler(document.getElementById('wpName2'), "keyup", CFAUtimerstart);
        var tdNode = document.getElementById('wpName2').parentNode;

        var CFAUspan = document.createElement('span');
        CFAUspan.setAttribute('id', 'CFAUresult');
        tdNode.appendChild(CFAUspan);
    }
}
```
function CFAUtimerStart() {
    if (cfauTimer) {
        window.clearTimeout(cfauTimer);
    }
    if (document.getElementById('wpName2').value.length > 0) {
        cfauTimer = window.setTimeout('CheckForAvaibleUsername()', 500);
    }
}

function CheckForAvaibleUsername() {
    document.getElementById('CFAUresult').innerHTML = ''; 
    injectSpinner(document.getElementById('CFAUresult'), 'cfau'); 
    var cfauUsername = document.getElementById('wpName2').value; 
    cfauUsername = cfauUsername.charAt(0).toUpperCase() + cfauUsername.substr(1);

    if (!cfauUsername) {
        alert('Enter the username you want, to see if it's available.');
        return;
    }
    var aobject = sajax_init_object();
    try {
        aobject.open('GET', 'wgServer + wgScriptPath + '/api.php?format=xml&action=query&list=allusers&aulimit=1&aufrom=' + cfauUsername, true);
    } catch (e) {
        throw e;
    }
    aobject.onreadystatechange = function() {
        if (aobject.readyState != 4) {
            return;
        }
        if (aobject.status == 200) {
            alert('The username ' + cfauUsername + ' is available.');
            return;
        }
    }
}
var res = aobject.responseXML.getElementsByTagName('u')[0].attributes[0].value;
removeSpinner('cfau');
if(res == cfauUsername) {
    document.getElementById('CFAUresult').style.color = 'red';
document.getElementById('CFAUresult').innerHTML = ' The username ' + cfauUsername + ' is busy.';
}
else {
    document.getElementById('CFAUresult').style.color = '';
document.getElementById('CFAUresult').innerHTML = ' The username ' + cfauUsername + ' is available.';
}
else {
    // if we get an http-error, we won't bother showing an error message.
    removeSpinner('cfau');
}
addOnloadHook(CFAUloader);

Implementation
To implement this function on other language versions of Wikipedia, copy the code above, paste it into the JavaScript page of the wiki, e.g. http://en.wikipedia.org/wiki/MediaWiki:Common.js and translate the phrases marked in red above.
Thanks

In many respects, I, Lennart, was not the ideal person to handle the Account Creation Improvement Project. However, I had many people to help me. Some of them were volunteers, such as Fetchcomms, Rock drum, Mono, Sertion, Ainali, and everybody who made suggestions. Special thanks to those who helped out with the surveys.

There has also been a lot of help from Wikimedia Foundation staffers, such as Frank Schulenburg, Roan Kattouw, Nimish Gautam, Katie Horn, Howie Fung, Winifred Olliff, Erik Möller, Zack Exley, Brandon Harris, Alolita Sharma, Philippe Beaudette, and Steven Walling.

This report is licensed CC-BY-SA 3.0. This means that you may copy, distribute and remix it freely, as long as you attribute the author and as long as you distribute it under the same/a similar license.

If you want to contact Lennart Gulbrandsson, try emailing him at lennart@wikimedia.se.