

# User Study Report V

Yuwei Lin, Meik Poschen and Rob Procter  
28-09-2008

## 1. Overview

We are now in the third round of interviews. In this round, we found sampling and identifying respondents the most challenging thus far.

Currently, respondents are recruited either via myExperiment or by snowball sampling (suggested by people who have been interviewed). In the former, we visited the site from time to time and approached people who had been recently active on myExperiment (when the researchers and the potential respondents were co-present on myExperiment site, even if not in real time). We then contacted these potential respondents by sending them a message directly on myExperiment. Most of the messages were not responded to. Half of the responses we approached agreed to be interviewed, while the other half rejected our request mainly because of their inexperience with the site (for example, one got back us saying “I haven’t used any of the workflow features of my experiment since I’m not a Taverna user. Pretty much all I’ve done so far is to create an account, so I’m afraid I wouldn’t be much use for your survey.”).

Selecting respondents in this way is not optimal as it makes it more difficult to follow a sampling strategy, and we thought it would be better if we can more systematically profile the user community so as to better understand users' behaviours and roles in the community. We have been through all groups and would like to get in touch with all group moderators. However, we found that many groups only have one member (which is the group moderator) and many members have only been to myExperiment once (the day they created their account).

Facing this methodological problem, we thought quantitative methods such as social network analysis or data mining of the user logs might be able to shed light on user profiling and categorisation (provided the user log files are available). For example, we can categorise users against how often they log-in to the site, how many files they have uploaded, how many files they have downloaded, how many packs they have created, how many groups they moderate and/or join, if they have reused workflows on myExperiment (e.g., with attributions in their workflows). Based on this, we would then be able to approach users that are representative of these categories. However, to do so, we are yet to get access to these data.

Alternatively, it would help if the list of users appearing under the ‘users’ tab could be filtered on, e.g., recent use, numbers of log-ons, numbers of workflows up loaded and other attributes of their profiles.

From June 2008 to date, we have done four more interviews (user 12, 13, 14, 15 in table I below).

| Interviewee (times interviewed) | Location    | Disciplines                              | Sector   | Joined        | Workflows (credited) | Friends | Groups: admin / member | Packs / Files | Interview date       |
|---------------------------------|-------------|--|----------|---------------|----------------------|---------|------------------------|---------------|----------------------|
| User 1 (1x)                     | UK          | Bioinformatics                           | academic | 26/07/08      | 22 (86x)             | 35      | 2 / 10                 | 2 / 4         | 23/10/07             |
| User 2 (2x)                     | Germany     | Bioinformatics                           | academic | 13/08/07      | 1 (1x)               | 6       | 2 / 0                  | 0 / 0         | 07/11/07<br>23/04/08 |
| User 3 (2x)                     | UK          | Bioinformatics<br>Bioanalytical Sciences | academic | 13/11/07      | 0 (0x)               | 0       | 0 / 0                  | 0 / 0         | 08/11/07<br>21/05/08 |
| User 4 (2x)                     | UK          | Bioinformatics<br>Bioanalytical Sciences | academic | 13/11/07      | 5 (73x)              | 14      | 1 / 4                  | 1 / 0         | 08/11/07<br>21/05/08 |
| User 5 (2x)                     | Netherlands | Bioinformatics<br>Computer Science       | academic | before Nov'07 | 7 (6x)               | 5       | 1 / 5                  | 0 / 0         | 08/11/07<br>21/04/08 |
| User 6 (1x)                     | UK          | Bioinformatics                           | academic | 15/10/07      | 0 (0x)               | 0       | 0 / 1                  | 0 / 0         | 09/11/07             |
| User 7 (1x)                     | UK          | Computer Science                         | academic | 20/07/07      | 9 (20x)              | 27      | 5 / 7                  | 1 / 4         | 13/11/07             |
| User 8 (1x)                     | UK          | Bio-/Health-Informatics                  | academic | 12/12/07      | 6 (4x)               | 4       | 2 / 1                  | 0 / 0         | 08/02/08             |
| User 9 (1x)                     | Netherlands | Informatics, Medicine                    | academic | 04/03/08      | 1 (1x)               | 5       | 0 / 2                  | 0 / 0         | 28/04/08             |
| User 10 (1x)                    | UK          | Bioinformatics                           | academic | 04/01/08      | 1 (1x)               | 6       | 2 / 1                  | 0 / 0         | 30/04/08             |
| User 11 (1x)                    | Netherlands | Biology, Bioinformatics                  | academic | 21/07/07      | 28 (29x)             | 42      | 10 / 6                 | 4 / 2         | 06/05/08             |
| User 12 (1x)                    | UK          | Text Mining                              | academic | 19/05/08      | 1 (1x)               | 2       | 1 / 1                  | 0 / 0         | 03/09/08             |
| User 13 (1x)                    | US          | Social Sciences                          | academic | 17/12/07      | 4 (5x)               | 11      | 1 / 3                  | 0 / 1         | 03/09/08             |
| User 14 (1x)                    | Germany     | Neuro-/Bio-Informatics                   | academic | 06/08/08      | 5 (5x)               | 3       | 1 / 2                  | 0 / 0         | 24/09/08             |
| User 15 (1x)                    | Netherlands | Bioinformatics                           | academic | 16/09/08      | 3 (4)                | 0       | 0 / 0                  | 1 / 1         | 25/09/08             |

Table I: myExperiment interviewee summary statistics (28/09/2008).

In these interviews, we continued to focus on users' experience with workflow sharing, workflow reuse, and how myExperiment can help them. We also asked them about their experiences with other social networking sites and how myExperiment differs from them.

## 2. Summary of Findings

### 2.1. Workflow reuse and modularised ways of building workflows

User 15 was the only respondent from our recent interviews who said unequivocally that he has reused workflows available on myExperiment. User 14 said he hadn't reused workflows very much. He reported some malfunction of myExperiment when browsing the technical structures of some workflows. We looked up the workflows he uploaded and found that some of them contain attributions, which is a sign that he has reused others' workflows on myExperiment. But for some reason he didn't consider this as reuse. All other recent interviewees mentioned that they haven't reused any workflows on myExperiment. This is either because "we haven't got any needs for doing that so far" or due to irrelevant workflows (for different tasks or domains). Both User 14 and User 15 are relatively new myExperiment users. The workflows seeded by other champions have saved them time to build a

workflow from scratch. User 15 also reported that the workflows on myExperiment sometimes inspired him to find solutions to his problems.

But some of them (User 12 and 13) also mentioned that they would normally build workflows in a modularised way. User 13 said they “spend the effort of building a workflow that makes them as transparent and reusable as possible”. They try to increase the flexibility of what others can do with the workflows so that additional value can be added. User 12 distinguished his strategies of building modularised workflows from the usual “putting components and combining them” way:

“The kind of modularisation I was talking about was... rather than separating individual items itself and processing and get outputs ... it allows easier ways to create interactions between different interactions between different workflows, just by changing input and output components. So it allows you to create different files conversion or input translation rather than completely redesign the workflow. You just have to mark something and you can mark out something completely different.”

However, as noted in the previous report, we suspected that workflow developer/creator and workflow user may prefer different kinds of workflow structures and sizes: end-users prefer larger workflows ready to be down-loaded and executed while developers prefer modularised workflows (small components) which can be assembled/customised. Our interviewees seem to fall into the category of the workflow developers/creators rather than end users. So their strategies of building workflows and the kind of workflows they upload may be influenced by their roles.

## *2.2. Motivations for using myExperiment*

All interviewees mentioned that they have used myExperiment for publishing and disseminating their workflows.

User 14 expressed a more altruistic view about sharing workflows: “When my solutions work, it would be great to share with others.”

However, beyond the usual publication and dissemination purpose, User 12 said she used myExperiment strategically to “build an actual community on myExperiment using those kind of approaches, e.g., Group system etc.”

## *2.3. Searching workflows*

Despite the recognition of myExperiment’s role in facilitating the publication and dissemination of workflows, many have expressed difficulties in searching and finding workflows. This is due to either the insufficient descriptions of the workflows, or inaccurate tags.

For example, there are plenty of workflows in the text mining area. “One of the issues we found is everyone kinds of puts things down as a text mining tag which is not very specific in order to get the different bits and pieces that we need. So users need to put in or people who actually put things there need to start considering how to label things in order to have them being found.”

User 13 said “I always think of tags as supplements for navigation because you can’t really rely on people conforming to tagging standard ways that other people doing it. So if it’s a low frequency tag you are not necessarily going to get everything even related to it.”

User 14 said at the moment when the number of workflows remains small, finding and searching workflows is not yet problematic.

Nevertheless, if the search function is not sufficient, myExperiment's goal of being a "social curation site" (rather than just a social networking site) may not be fully realised. As boyd<sup>1</sup> (2006) argues, 'searchability' is one of the four properties that are typically present in online social networking sites (the other three are persistence, exact copyability, and invisible audiences).

#### 2.4. Usability

We have noticed that the developers have been doing a good job responding to users' comments. For example, a new feature of 'Invite' has been created to invite people outside myExperiment to participate. And people can now delete messages in their message box.

Other new features such as packs and enactment may not be immediately useful for others because they either don't have a lot of workflows or files to be bundled together (which is the functionality of packs) or they don't have their own servers to process and start a workflow on myExperiment site (which is enactment).

#### 2.5. How does myExperiment differ from other social networking websites?

Asked how myExperiment differs from other social networking websites, User 12 said:

"[Facebook or LinkedIn are] more of a personal contact with colleagues and that's why I think myExperiment has been so useful for us, in terms of it allows that extra level you don't usually get from the resources ... It puts the contacts on the kind of work you discuss, whereas it is general chatters on Facebook ... You're actually talking around the actual work itself. One of the other benefits of using that kind of thing is like you also got your documentation with that kind of interaction there as well. Again, with the context of the actual object you discuss. So that's the major benefit of it. Rather than using your system or email or SMS, you constantly get different sets of discussion. Now although you've got the archive of those, it's often useful to have those in a safe place, as an object that you discuss. Just for simple things like archive and retrieval that kind of information. And it's often easier to say that, when you have looked at a particular workflow, you can see the benefit of this, rather than just say that I'll just send you the latest version in an email start discussing it. There's a significant sense of gravity around which you can discuss different items. And everyone seems to be there at the same point."

In a similar way, User 13 distinguished myExperiment from other "profiles-based and contacts-based" social networking sites such as Facebook, and characterised it as a "functionality-oriented" and "object-oriented" site (like Flickr). And we think it is also for this reason that myExperiment is better defined as a 'social curation site' rather than as a 'social networking site'.

Feature requests to improve the communication process:

- To be able to send messages to the whole group
- Mark all replied messages in the message box
- More APIs and plugins

#### 2.6. Longitudinal Analysis

Table 1 above depicts the latest myExperiment statistics of all 15 interviewees so far. As packs are a relatively new feature and also may incorporate files (which might enhance their importance and number; before users rarely used them according to our interviews) we have now included these as categories. The more frequent use of packs first had to be established and our findings indicate they are

---

<sup>1</sup> danah boyd insists that her name can only be spelt in lower case.

known and used regularly by now, although the recent numbers of uploaded packs among the interviewees are still relatively low: in our current round of interviews (Users 12-15), three out of four users stated they know what packs are, two have used them and User 15 had even uploaded a pack. We started raising these statistics with the third interim user report (06/06/08), but at that time only for the seven interviews we conducted in the second round, as can be seen in Table II.

It is evident that the statistics for Users 2-5, which represent those users interviewed twice, have not change at all within the last four months – with the exception of User 4. This user added one more workflow and one pack, but the most significant change refers to his credits: being credited 5 times beginning of June the number of credits has risen to 73. This is one issue we will look into further in future fieldwork and data analysis, as well as into checking the activity/inactivity of the other three users, whose numbers did not change. User 11, a very active power user of myExperiment, shows a slight increase in some numbers, which is not surprising given her high totals and her general visible activity within and cooperation with the myExperiment family. Users 9 and 10 statistics also show some activity which we will drill deeper into in later interviews.

| Interviewee (times interviewed) | Location    | Disciplines                              | Sector   | Joined        | Workflows (credited) | Friends | Groups: admin / member | Packs / Files | Interview date       |
|---------------------------------|-------------|--|----------|---------------|----------------------|---------|------------------------|---------------|----------------------|
| User 2 (2x)                     | Germany     | Bioinformatics                           | academic | 13/08/07      | 1 (1x)               | 6       | 2 / 0                  | n/a           | 07/11/07<br>23/04/08 |
| User 3 (2x)                     | UK          | Bioinformatics<br>Bioanalytical Sciences | academic | 13/11/07      | 0 (0x)               | 0       | 0 / 0                  | n/a           | 08/11/07<br>21/05/08 |
| User 4 (2x)                     | UK          | Bioinformatics<br>Bioanalytical Sciences | academic | 13/11/07      | 4 (5x)               | 14      | 1 / 4                  | n/a           | 08/11/07<br>21/05/08 |
| User 5 (2x)                     | Netherlands | Bioinformatics                           | academic | before Nov'07 | 7 (6x)               | 5       | 1 / 5                  | n/a           | 08/11/07<br>21/04/08 |
| User 9 (1x)                     | Netherlands | Informatics<br>Medicine                  | academic | 04/03/08      | 1 (1x)               | 2       | 0 / 2                  | n/a           | 28/04/08             |
| User 10 (1x)                    | UK          | Bioinformatics                           | academic | 04/01/08      | 1 (1x)               | 4       | 2 / 1                  | n/a           | 30/04/08             |
| User 11 (1x)                    | Netherlands | Biology<br>Bioinformatics                | academic | 21/07/07      | 26 (27x)             | 39      | 9 / 5                  | n/a           | 06/05/08             |

Table II: myExperiment interviewee summary statistics (06/06/08).

Users 3 and 4 are working in the same research group on different activities. The first interview with both of them took place in November 2007 and was focused on their views of the potential value of myExperiment and of becoming part of the myExperiment community. At that point neither of them had really used myExperiment (they had explored it briefly a day before).

User 3 was not sure if future use would make sense (reflected by no visible activities so far, as can be seen in Table I), as she considered most of her workflows ‘very specialised’ and really using myExperiment would ‘add too much complexity right now’; she stated she would think about sharing her simpler data analysis type workflows on myExperiment, especially for students. The follow-up interview in May 2008 (again with both users) picked up on the findings of the first interview and showed that it is not feasible for User 3 to upload/share her own workflows. However, she has searched for workflows and looked at a few (not downloaded them), but in the end those have not been very useful, despite the time invested: ‘it’s quite a bit effort to see what they really did or didn’t do’. User 3 did not use any of the social networking features on myExperiment (and does not use any social networking site), mainly because those workflows have been ‘more on the fringe of my work’, so it

made no sense to invest time to contact anyone. For User 3 it seems that the initial attitude, i.e., that myExperiment is of no real use for her work, is confirmed by the follow-up interview and the non-existent activities looking at the stats. Nevertheless, feedback on issues like better descriptions of workflow in/outputs, enhancing the transparency and functionality of nested workflows and including (biological) models in myExperiment has been useful.

User 4 has a quite different approach to User 3. From the beginning (Nov 2007) he stated that he wanted to use myExperiment mainly as a 'broadcasting tool', benefiting from its network functionality, to upload workflows and link to them in his research publications (papers, journals). In the second interview (May 2008), he confirmed having done this successfully with having one paper which exactly does this accepted and the another in review. He has not found anything to download or re-use in any way in his area, but he did not really expect to – but he also mentions that more sequential descriptions could be more useful for finding suitable workflows in general, especially when they are complex. User 4 finds it nice to be credited, but has no collaboration or exchange on myExperiment in connection with his workflows or his research field (or anything else). In the first interview he stated he was interested to see if the community would evolve in that respect, but so far nothing had happened (he is the only member in his group). Enactment could be an interesting feature in the future. In publishing workflows on myExperiment (and getting credits for this, especially in the last months) User 4 has been quite active since the first interview. It remains to be investigated if he also managed to establish contacts in his field, could download useful workflows and make use of the community in that regard.

The first interviews with User 2 and User 5 independently took place also in November 2007, with follow-ups done in April 2008. These users previously had filled out the myExperiment user survey and were chosen because of their participation, with the first interviews being based on survey findings (see User Evaluation Report I, 29/10/07).

User 2 points out that at the time of the first interview (Nov 2007) in the launch phase of myExperiment there were not very many workflows. From the beginning she was looking for people to get to interesting workflows. She has currently not put up her own workflows, but states she will do so in the future, hoping the community will grow. She also points out the importance of services and that it would be good to have an icon to recognise them quicker inside the workflows. In the second interview in April 2008 User 2 reports using myExperiment more often now as the community has grown. She also has, as announced, uploaded a workflow, which she used for a course at her University. She plans to use the network functionalities of myExperiment to make her workflows accessible in two ways: first in a closed myExperiment group for an EU project (<http://www.eu-sol.net/>) and second in another more open group for teaching/students. This is possible through the access rights functionality, which also solves concerns about IPR raised in the first interview. Enactment could be an interesting feature in the future. Re-use of workflows was no issue at the time of the interview. In the next round of interviews we will investigate further if plans to use groups and putting up workflows for two activities has been realised, as there is no activity visible right now.

User 5, interviewed in November 2007 and subsequently April 2008 started on myExperiment putting up simple, generic and recyclable workflows which he migrated from Perl to Taverna for this purpose. In the second interview he emphasised this as the main benefit. He mentioned the fear of data loss and of possible IPR issues in the first interview, but found these issues have been resolved, and the platform will be around for a while (talking to the myExperiment family also helped to gain trust). Also, the documentation/information in the meantime had improved substantially. In this context, he stressed the importance of the short lifecycle of myExperiment. For useful workflows or real re-use, he expected the amount of workflows will grow and he confirmed his view in the second interview: publishing and sharing workflows is the main benefit of myExperiment, packs are really useful; the proper re-use of

workflows remains difficult, but small bits are helpful. As with User 2, we need to check how active User 5 is nowadays.

In general it is very probable that all four users have been further engaged in exploring myExperiment stronger in some degree as a result of the additional information/motivation received from being involved in the first interview (and for Users 2 and 5 also the myExperiment survey), which fostered reflection on potential benefits of using and sharing workflows on the system (and looking into new features like Enactment). Further research will examine how sustainable their engagement will be – and what it will encompass. Two (User 4 and 5) of the four users also contributed to the community in uploading more than one workflow, while the number of their credits shows those workflows are actually regarded as useful for others. If this value lies in merely referencing the workflows or even in re-using them, has to be looked into further. Additional research also has to explore if Users 2 and 5 are still active on myExperiment and in which way – User 3 does not seem to have clung to idea from the beginning and User 4 seems to be active in terms of especially using myExperiment as a base for publication (with 73 credits); User 2 claimed the same additionally for the purposes of establishing a project and teaching workspace/group, but no recent activity is visible. In the end this makes publishing of workflows and establishing a community using the group functionality the two most important features of myExperiment coming out of these repeated interviews with four users. Besides publishing, User 5 also states the sharing and re-cycling of workflows is the main benefit. It remains an issue for future investigation how downloads of workflows actually relate to them being used/re-used.

### **3. Future Plans**

We will continue the interviews, and profile the user community. By the end of March 2009, we aim to have a sample of about 25-30 interviewees and to have completed a total of 40-50 interviews. In the mean time, we will refocus and update our interview schedule/guideline, continue to recruit new users and to re-interview previous subjects in order to develop an understanding of how users' use and attitudes to myExperiment change over time. To achieve this, the longitudinal account on myExperiment users and their online sharing and networking behaviours will be further developed. This will also help us to understand how network effect and externalities are built up over time.

Finally, we will work closely with Antoon to compare the myExperiment user community with that of the OpenWetWare project.

### **References**

boyd, d. (2006). Friends, friendsters, and top 8: Writing community into being on social network sites. *First Monday*, volume 11, number 12 (December 2006), URL: [http://firstmonday.org/issues/issue11\\_12/boyd/index.html](http://firstmonday.org/issues/issue11_12/boyd/index.html) (retrieved on 24 Sep 2008).