

Digital Studies / Le champ numérique, Vol 3, No 2 (2012)

Ukrainian Folklore Audio Project

Geoffrey Rockwell, Professor of Philosophy and Humanities Computing, University of Alberta:
geoffrey.rockwell@ualberta.ca

Megan Sellmer, Humanities Computing, University of Alberta: sellmer@ualberta.ca

Natalie Kononenko, Professor and Kule Chair, University of Alberta: natalie.kononenko@ualberta.ca

Maryna Chernyavska, Researcher in the Kule Folklore Center, Archival Assistant in the Bohdan Medwidsky Ukrainian Folklore Archives, University of Alberta: maryna.chernyavska@ualberta.ca

Karl Anvik, University of Alberta: karl.anvik@ualberta.ca^[1]

Abstract / Résumé

How can crowdsourcing be used in the humanities? This paper describes a crowdsourcing project at the University of Alberta that developed a tool for participants to sign out audio clips for transcribing and translating from Ukrainian. The audio clips were from fieldwork on Ukrainian folklore and include stories, songs, and beliefs. We describe the design process and provide Crowdsourcing only works if you can motivate sufficient participants and different participant communities call for different strategies. an environmental comparison with other crowdsourcing projects. We discuss the challenge of motivating community members with the language skills to volunteer to help with the project. Crowdsourcing only works if you can motivate sufficient participants, and different participant communities call for different strategies.

Comment l'externalisation ouverte peut-elle être utilisée en lettres et sciences humaines? Le présent article décrit un projet d'externalisation ouverte à l'université de l'Alberta, dans le cadre duquel un outil a été mis au point afin que les participants puissent emprunter des audioclips pour les transcrire et les traduire à partir de l'Ukrainien. Les audioclips proviennent de travaux effectués sur le terrain sur le folklore ukrainien et incluent des histoires, des chansons et des croyances. Nous décrivons le processus d'élaboration et fournissons une comparaison environnementale avec d'autres projets d'externalisation ouverte. Nous traitons du défi à motiver les membres de la communauté qui parlent la langue à apporter bénévolement leur aide au projet. L'externalisation ouverte n'est efficace que si l'on peut motiver un nombre suffisant de participants, et différentes communautés de participants requièrent différentes stratégies.

KEYWORDS / MOTS-CLÉS

Ukainian Folklore, crowdsourcing, digital audio

Contents

- [Introduction to the Ukrainian Folklore Audio Project](#)
- [Developing the UFAP web application](#)
- [The developed web site](#)
- [Crowdsourcing in humanities research](#)
- [Environmental comparison and early results](#)
- [Early usage results](#)
- [Conclusion](#)
- [Notes](#)
- [Works cited](#)

The Ukrainian Folklore Audio Project (UFAP) developed a custom crowdsourcing tool for community research at the University of Alberta in 2010-11. The goal of this research project was to involve participants in translating and transcribing audio-recordings of Ukrainian folklore stories, songs and beliefs. The tool was developed both as a way of enhancing the existing online Ukrainian folklore resources and as a way of looking at a novel use of crowdsourcing in the digital humanities. The major challenge we faced was that the pool of potential volunteers able to translate Ukrainian was small and many Ukrainians in Canada felt insecure about their language knowledge. While crowdsourcing works well when the potential volunteers are digital natives, many humanities crowdsourcing projects need to tap into groups with specific expertise that don't necessarily feel comfortable using the Web. This paper describes how we tackled this challenge and the first mixed results. In this paper we will:

- Explain the background of the Ukrainian Folklore Audio Project.
- Describe the audio folklore groupsourcing tool we developed and how we developed it.
- Discuss crowdsourcing in the humanities generally and how our site compares to other projects.
- Conclude by talking about the preliminary usage of the project and how we plan to engage more participants.

Introduction to the Ukrainian Folklore Audio Project

Figure 1: The Ukrainian Folklore Sound Recordings home page.



The Ukrainian Folklore Audio Project experiments with crowdsourcing, or as we call it, groupsourcing, for the tagging, translating and transcribing of audio passages. We call it groupsourcing because while the site is open to the public, we wanted to control the uploading of materials and thus required that participants who would actually work on the materials be approved to get a password. Thus only a smaller group of those accessing the site would be able to process the files.

Folklorist Dr. Kononenko has recorded over 200 hours of audio related to Ukrainian folklore since 1998. Many of these recorded Ukrainian folksongs, narratives, and beliefs are not written down anywhere. Ukraine was newly independent at the time of Kononenko's fieldwork and the upheaval experienced by the country at that time meant that folklorists were doing little, if any, local collection work. That Kononenko's material reflects this period of change makes it especially interesting, as Kononenko's recordings preserve oral history taken during a period of change. Obviously these stories are also of interest to folklorists.

As a prior project, digitised versions of the audio materials were made available through the Ukrainian Folklore Sound Recordings website where users can use a Ukrainian or English index to find and listen to passages (see [Figure 1](#) above. Also see the website: <http://projects.tapor.ualberta.ca/UkraineAudio/>). Users navigate through a hierarchical index until they get to the metadata (see example below) for a long recording that includes information about all the topics (subjects) covered. The system then gives users controls to listen to the portions of the audio where the subject they want is located. Below is some metadata for one of the Ukrainian Folklore Sound Recordings where there is folklore information about Spring Festivals:

Tape/disc #/name: Berlozy2005G

Interviewer: Dr. Natalie Kononenko

Project Title: Calendar Rites/Central Ukraine

Date: July 15, 2005

Medium: CD (digital recording)

Overall interview time: 26.49

Place/village name: selo Berlozy, Kozelets'kyi raion, Chernihivs'ka oblast'

Interviewee (last name, first name): Hnida, Hanna Fedorivna, DOB 1950

TIME (min:sec) SUBJECT

04.45-15.13 Наратив-клуб

24.40-26.20 Наратив-сучасна молодь

15.19-15.56 Хрестини-звездини

21.20-23.53 Вечорниці/досвітки-загальна інформація

01.35-04.40 обряд/свято-Івана Купала...

The limitation of the Sound Recordings site is that it doesn't contain transcriptions or translations, just recordings and a topical navigation index, meaning that only people who know Ukrainian can understand the material presented. Transcribing and translating 200 hours worth of materials would be expensive and time-consuming, especially as more audio materials continue to be gathered, a more effective way of communicating design ideas. So we decided to experiment with groupsourcing as an alternative way to get transcriptions and translation. A groupsourcing site would allow Kononenko to receive volunteer help transcribing and translating on an ongoing basis, even after the grant funding runs out. We also hoped that groupsourcing would strengthen the connections between researchers at the University of Alberta and the Ukrainian community. All along, however, we knew that our greatest challenge would be developing a groupsourcing application and process that would suit the likely participants, many of whom are modest about their language skills and are older, and therefore have less experience with the Internet.

Developing the UFAP web application

To make it possible for volunteers to contribute transcriptions and translations, we had to design and build a custom groupsourcing web application that would use the already digitised audio from the Sound Recordings site. The new web application was programmed by Karl Anvik based on wireframes and other design documents prepared by Megan Sellmer after extensive design discussions following a persona/scenario user interface design process (we subscribe to an open research philosophy of sharing our relevant documents. You can see all our design documents on a wiki at: http://circa.cs.ualberta.ca/index.php/CIRCA:Ukrainian_Folklore_Audio_Project). The design process was important given the anticipated challenges of involving older participants and the need to husband our programming resources. The persona/scenario process helped us structure the discussions. The team had to negotiate common expectations and possibilities between humanities computing researchers and folklore researchers. What follows is a brief description of the process.

Personas: The persona/scenario process we used has been adapted from Cooper's *The Inmates are Running the Asylum* (2004), which argues that when designing, one should develop profiles of specific and believable anticipated users and then design for them. These anticipated users or personas are not real customers nor are they general descriptions of user needs, but are fictionalised people who are given specific names and histories in negotiation with stakeholders. Using personas allows a design team to talk through stories as if we were talking about real people, as in "What if Elena forgets her password?" In other projects this has proven a more effective way of communicating design ideas. Here is the description of Elena, our primary persona:

Elena is 72-years-old; she and her parents left Ukraine when Elena was 9. She has 5 children and 11 grandchildren. Sadly, Elena lost her husband three years prior (he was an Ukrainian immigrant as well). She lives by herself in Callingwood North. She does not know how to work the Internet, but her 17-year-old granddaughter (Rachel) can help her after school on Wednesdays. Elena can both read and write in Ukrainian, she learned this from her mother at a young age. She knows Natalie from the local Edmonton community and was invited to attend the website workshop for local Ukrainian community members. Elena wants to do this work and feels community pride at being acknowledged on the website. She focuses on the stories her mama told her, like "The Flowering Fern." (From the Personas and Scenarios page of the open research wiki: http://circa.cs.ualberta.ca/index.php/CIRCA:Personas_and_Scenarios) ^[2]

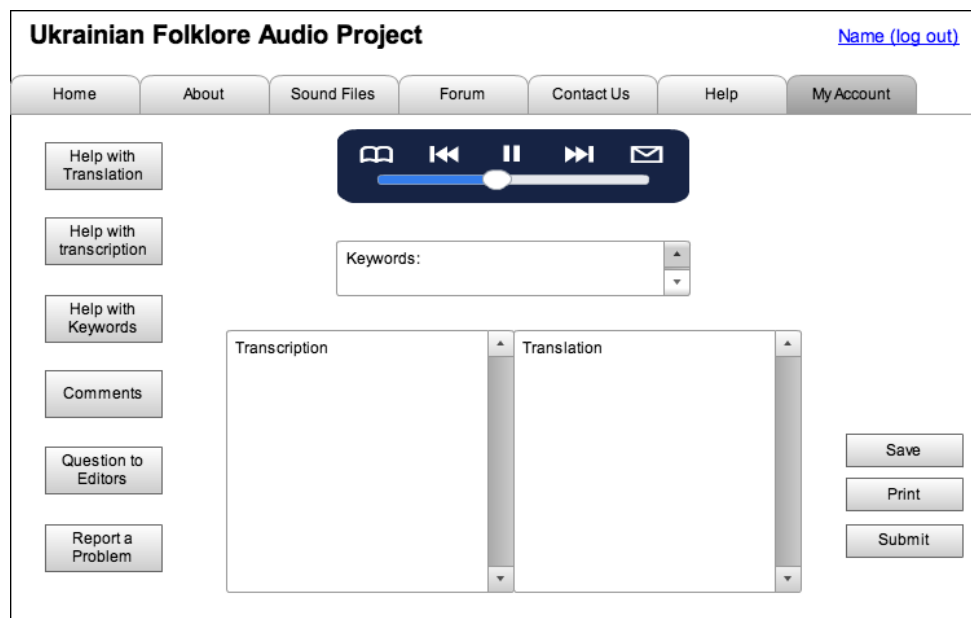
Important to this process is negotiation. Developing these personas was a way for the team to discuss and agree on who this website is designed for by talking about people, albeit invented people. It was also a way for the computing team responsible for development to understand the participant community that the folklorist is accustomed to. We concluded that one type of user might be an elderly Ukrainian community member who hopefully might receive help from family members. The persona of Elena represents this anticipated type of volunteer. Other personas represent other types of users. You can see the other two primary personas at http://circa.cs.ualberta.ca/index.php/CIRCA:Personas_and_Scenarios. It is worth noting that there were differences within the team as to what the most likely user would be, which is another reason for trying to describe the hypotheses with personas so as to be able to balance priorities. It should be noted that currently the volunteers are not like Elena. They are younger than Elena and reasonably familiar with the Internet. Developing and prioritising personas is not a science, but a way of anticipating users in design – one has to start somewhere. And, while Elena may not have turned out to be our typical user, we felt it was important that we design for someone like her to make the web site as accessible as possible.

Scenarios: After developing personas we prioritised them and developed multiple usage scenarios for each. The scenarios tell a story of usage, and go through, step-by-step, what the users would do on the website. Here are the opening steps of the first scenario for Elena:

1. Her granddaughter helps her log in with the password and account Maryna set up for her at the workshop.
2. They sort through and pick the story that Elena would like to transcribe.
3. Rachel reminds Elena how to select the audio clip and download it.
4. She shows her Baba where they saved it (the desktop).
5. They log out of the website; she doesn't need to be on the website because she downloaded the audio and is using her own word processor.^[3]

The scenarios are a way for the team to negotiate what features are important in anticipated use and to tease out expectations as to how the system will work. For example, from this scenario we knew that the website would have to allow users to download the short audio clip so that transcription could be done outside the site. Instead of imagining what features it would be more beneficial to have, scenarios also keep the design focused on what we anticipate the personas should do. Once scenarios are developed and negotiated, the scenarios then provide useful documentation to the programmers and a way to test/audit the application. A programmer can start developing what is needed just to support the scenarios and not worry about undocumented expectations on the part of the design team.

Figure 2: Transcription and translation wireframe



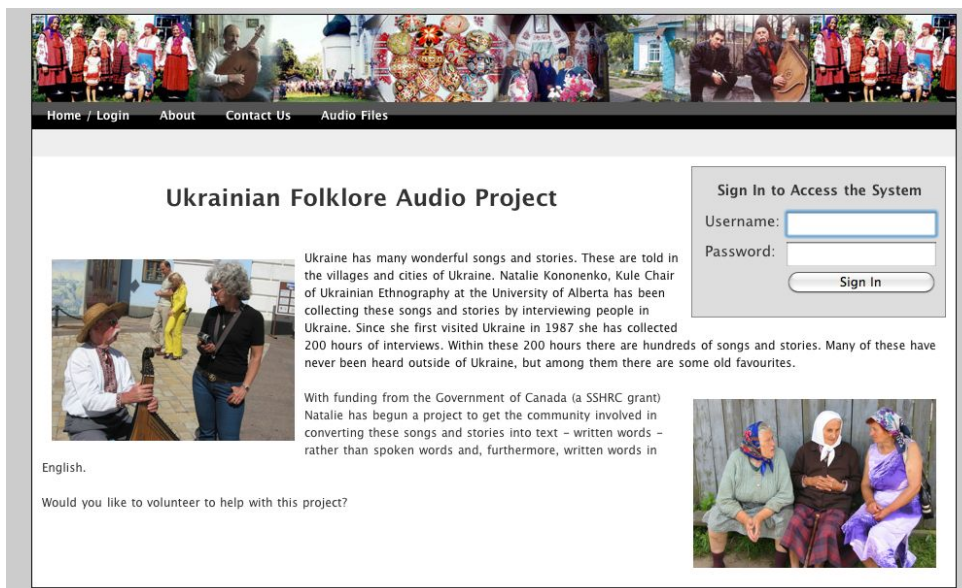
Wireframes: The next step, once we had agreed on the scenarios and their priority, was to create wireframes of the major screens from the scenarios. For example, the wireframe in [Figure 2](#) is for the transcription and translation of the site. The wireframes were built with Cacao, a free site that allows users to create a limited number of diagrams (<https://cacao.com/>) (the free version limits the number of diagrams you can create; you can pay for unlimited diagrams). The wireframes were not designed to show the graphic design or even a definitive arrangement of features. The purpose of wireframes is to show the functionality that has to be on each web page following from the scenarios and to suggest possible arrangements of functionality. Wireframes can also serve to separate discussion of functionality from discussion of graphic design features like logo, background, colours, and fonts. One of the advantages of this process is that a team can discuss issues and come to a decision step by step rather than struggling with feelings about graphics when working out anticipated functionality. The down side is that it takes a long time to go through the process. As a team, we again went over these to make sure that the website included what everyone thought was necessary and what was found in the scenarios. The wireframes, like the personas and scenarios, were redone until the team was satisfied.

Once completed, the wireframes and the scenarios were used by the programmer to program the site and then used to test the site. The programming went quickly given the thorough design process we had gone through, which helped as the project had a limited programming budget and we could not afford to change our minds. To some extent this process allows a large amount of the design and programming documentation to be managed by graduate research assistants rather than professional designers and programmers. This then has the advantage that the GRAs get trained and the budget is not dominated by professional salaries.

The developed web site

How does the groupsourcing web site actually work? As mentioned above, the Ukrainian Folklore Audio Project uses previously digitised audio, but these audio files were too long for our purposes. We hypothesised that volunteers would be more likely to participate if they were given smaller tasks that they could do at their leisure. Transcribing or translating a 45-minute audio recording would frighten away even the most enthusiastic volunteer. For this reason, the longer archival recordings were edited into shorter clips each with only a single story, song, or belief to be translated or transcribed. An example of a song that has been recorded and posted to our website is "And I look to find my Marusia," a story about a bumbling thief searching for his love (to hear the song and see the transcription and translation see http://research.artsrn.ualberta.ca/ukrfolklore/submissions_view.html?clip_id=4&filter=published). Once the clip was edited down, it was transferred from the Ukrainian Folklore Sound Recordings site (see [Figure 1](#) above) to the UFAP where it could be listed as a clip to be translated by a volunteer in the group.

Figure 3: Home page of UFAP site



On the UFAP site, one can learn about the project, contact the editor, Dr. Kononenko, and hear the audio clips. For those clips that have been transcribed and proofed, the text in Ukrainian may also be seen. For those clips that have been translated and proofed, the text in English also appears (for the web site, see <http://research.artsrn.ualberta.ca/ukrfolklore/>). Volunteers are recruited through community events and contacts, and given accounts by the editor Dr. Kononenko, who works closely with the Ukrainian community in Alberta, nationally, and internationally. Once a participant has an account, he or she can go to the home page and sign into the system. We felt it was necessary to have users log in to ensure the quality of the transcriptions and the translations; Dr. Kononenko didn't want frivolous contributions. The login process also allows Dr. Kononenko to correspond with each volunteer to help them learn about the site and project.

Figure 4: Screen with audio clips that users can sign out and work on

Audio Clips

Comments Question to Editors Report a Problem

<< first < prev 1 2 3 4 5 6 7 next > last >> 10

Name	Speaker	Description	Play Clip	Action	View
Talkative Feska - Язиката Хвеська		Magic tale / Казка	00:00 04:31	Edit	View
How an old woman found some money - Баба гроші знайшла			00:00 01:45	Sign Out	View
The little ox - Бичок			00:00 02:52	Sign Out	View
How children found some money - Діти знайшли гроші			00:00 03:03	Sign Out	View
Ivan the servant and the tongue - Іван-челядник і язик			00:00 01:55	Sign Out	View
Ivasyk Telesyk - Івасик Телесик			00:00 06:14	Sign Out	View
The half-skinned goat - Коза-дереза			00:00 05:33	Sign Out	View

Once a participant signs in, they can see a list of the available short clips and listen to them. If they want to transcribe or translate a clip they can sign it out, which then locks others out and prevents them from working on it (though others can still listen).

Figure 5: Screen with audio clips that users can sign out and work on

Edit Clip

Clip Name: Talkative Feska
Языката Хвеска

Description: Magic tale / Казка



Download File: 79-talkative_feska.mp3

Save This will save all of your work so you can continue working at a later date. This includes any text entered into the keywords, transcription and translation fields below.

Submit This will submit all of your work to be reviewed by the editorial staff.

- [Help With Keywords](#)
- [Help With Transcription](#)
- [Help With Translation](#)
- [Comments](#)
- [Question to Editors](#)
- [Report a Problem](#)

Keywords:

Transcription

Translation

To transcribe or translate, users go to "Sound Files" and click on their reserved recordings in the "My Clip" table. There the volunteers can listen to the clip while typing in the text boxes supplied. Along the left side are links to send a comment, report a problem, or ask a question. If at any time the volunteers need assistance, they can use one of these options. Participants can also add keywords; this is located in the area above the transcription and translation boxes. When a volunteer is finished, they are able to save their work and then submit it. Once a volunteer submits his or her work, it goes to the editors to be edited. After that step is complete, the work is published on the website for others to view. Volunteers can choose to remain anonymous but still have their work published or be recognised as the transcriber or translator.

We decided to let volunteers choose to complete either a transcription or translation in order to encourage them to do what they were comfortable with, as we expected that some volunteers would be shy about their rusty language skills. The volunteers' language level is yet another reason why they may choose to remain anonymous. We wanted to minimise the fear that volunteering could lead to embarrassment in the community.

Figure 6: Editor's view

Edit Clip

Clip Name: Talkative Feska
Языката Хвеска

Description: Magic tale / Казка



Download File: 79-talkative_feska.mp3

Save

This will save all of your work so you can continue working at a later date. This includes any text entered into the keywords, transcription and translation fields below.

Submit

This will submit all of your work to be reviewed by the editorial staff.

- Help With Keywords
- Help With Transcription
- Help With Translation
- Comments
- Question to Editors
- Report a Problem

Keywords:

Transcription

Translation

As for administrators, who have control over the audio clips, the categories (story, song, belief), and the submissions, they log in on the main page as well. There are pages that the administrators use to monitor the "comments," "problems," and "questions" sent by volunteers. These options were included in the design to ensure the quality and comfort of the volunteers' experience. Specific information about volunteer contributions is also available to administrators. Examples of information logged include completed submissions by participants, and the user and clip activity shown here.

Crowdsourcing in humanities research

At this point we will turn to reflect on the project by considering the use of crowdsourcing in the humanities. Crowdsourcing is an emerging digital method for getting a large project done by using a "crowd" of volunteer participants. Scholars are using crowdsourcing to complete large-scale projects that can be broken into smaller tasks, and as a way of involving the larger community of the humanities. Most uses of crowdsourcing in the humanities have been focused on textual materials, as in the Suda On Line project, which applies the power of the crowd to translating a Byzantine Encyclopedia (for more on the Suda On Line see <http://www.stoa.org/sol/>).

Involving volunteer participants in research is not a twenty-first century invention. *The Oxford English Dictionary* can be considered an early example of crowdsourcing in the humanities. The editors of the dictionary sent out letters to collect quotations. In *The Meaning of Everything: The Story of the Oxford Dictionary*, Simon Winchester explains:

There were ... no fewer than 1,827,306 illustrative quotations listed - selected from five million offered by thousands of volunteer readers and literary woolgatherers ... These were essential: the millions of words from these quotations offer up countless examples of exactly how the language worked ... (2003, xxv)

The Internet, however, provides us with a communications channel that facilitates the distribution of small research tasks and the automatic integration of volunteer contributions. There have, therefore, been a number of digital humanities projects that use crowdsourcing including The Dictionary of Words in the Wild (see <http://lexigraphi.ca>), The Day of Digital Humanities (see http://tapor.ualberta.ca/taporwiki/index.php/Day_in_the_Life_of_the_Digital_Humanities), Suda On Line (see <http://www.stoa.org/sol/> and Anne Mahoney in *DHQ*, "Tachypaedia Byzantina: The Suda On Line as Collaborative Encyclopedia" [2009]), and Transcribe Bentham (see <http://www.ucl.ac.uk/transcribe-bentham/> and the recent article by Causer et. al. [2012] on the project). The Internet allows researchers to better communicate and distribute knowledge tasks to volunteers, which is why experimenting with this approach to research tasks is gaining popularity (see Patricia Cohen's *New York Times* article [2010] "[For Bentham and Others, Scholars Enlist Public to Transcribe Papers](#)").

Crowdsourcing clearly can be an effective form of getting a large project done in a short amount of time. This does not mean that it is perfect; Piotr Organisciak in his thesis, *Why Bother? Examining the Motivations of Users in Large-Scale Crowd-Powered Online Initiatives* (2010), deals with one of the main challenges of crowdsourcing - that there is a

difficulty in motivating people to participate. For every project that succeeds there are doubtless many others that don't get enough volunteers to make headway.

Another problem is the ethics of crowdsourcing – is it ethical to ask others to do the work under all circumstances? Jonathan Zittran proposes that some crowdsourcing projects take advantage of those who need money (you can see Zittran talk on "Minds for Sale" on YouTube at <http://www.youtube.com/watch?v=Dw3h-rae3uq>). Jeffrey Young, in an article in the *Chronicle of Higher Education*, quotes Zittrain to the effect that a crowdsourcing site like Amazon's Mechanical Turk is a "digital sweatshop" (2011). Zittran specifically targets Mechanical Turk because it encourages people to work on other people's problems for pennies. While it is beyond the scope of this paper, it is worth asking what ethical considerations should govern volunteer (as opposed to paid) crowdsourcing projects like those run to engage a community in humanities research.

Environmental comparison and early results

This raises the question of how UFAP compares to other crowdsourcing sites. One easy way to assess a project is to compare it to similar sites. To that end, we conducted an environmental scan and comparison. First, a "check list" of characteristics we wanted to look for on each site was assembled; then successful crowdsourcing sites were identified with which UFAP could be compared. We settled on ten websites and six characteristics, based on the challenges we anticipated having with the UFAP site, especially that of motivating volunteers who weren't heavy Internet users or familiar with crowdsourcing. The websites we surveyed included Galaxy Zoo, Google Image Labeler, Kickstarter, Transcribe Bentham, Suda On Line, Buzzillions, Dictionary of Words in the Wild, Day of Digital Humanities, Herdict Web, and Foldit. The characteristics we looked at included:

1. *How does the user log in?* What is involved in getting an account?
2. *How many clicks* to get to the work? How difficult is it to get to where you can contribute?
3. *How does it recognise you?* How are users presented back to others on the system? Can you enter a profile? An image?
4. *How does it motivate?* How does the project encourage participation?
5. *What is the project applying crowdsourcing to?* What sorts of research problems are they using crowdsourcing to solve?
6. *What is the audience?* Who do they expect to have as volunteers? (Our notes on the scan of the crowdsourcing sites can be found at, http://circa.cs.ualberta.ca/index.php/CIRCA:Crowdsourcing_Website_Analysis)

What we discovered is that the Ukrainian Folklore Audio project fell in the middle when examined according to these characteristics. For instance, it took six "clicks" of the mouse to start crowdsourcing for our project. In the environmental scan, the highest number of clicks was ten, which requires too much effort, and the lowest four. Though we did not have the lowest number of clicks, the interface design was intentionally left plain so that volunteers will not be bogged down by too much information presented on the webpage.

One area where UFAP is different from most of the other projects is that we require human approval from the editor before you can contribute. Most crowdsourcing sites, though not all, require accounts, but they give accounts automatically. While this may be a disincentive to participation in UFAP, we suspect that it would also reassure some users.

Of particular importance to us was the fourth characteristic, the source of motivation of each site and how UFAP could motivate potential participants. As mentioned above, motivation is a vital part of crowdsourcing. Hars and Ou in "[Working for Free? Motivations for Participating in Open-Source Projects](#)" (2002) identify two types of motivation in crowdsourcing projects. *Intrinsic* motivation is the motivation of doing something to make yourself feel good and to contribute to society, and *external* motivation is a monetary or recognition reward. Many crowdsourcing websites used both intrinsic and external motivations such as the cultural and historical significance of the work (intrinsic) and gamifying the task (external). Gamifying a task refers to taking what participants are crowdsourcing and twisting it into a playful game, giving points for each correctly completed task as in the website Fold it. Another example is Google Image Labeler where users compete with other players to match tagging words (for more on gamification see Jane McGonigal's *Reality is Broken* [2011]).

The UFAP project also uses both types of motivation, but does not exploit external motivation as much as other projects. The primary motivation for volunteers is the intrinsic motivation of contributing to the preservation and exploration of Ukrainian stories, songs, and beliefs. Given that the volunteers come from the Ukrainian community, participating lets them explore their folklore heritage even if their language skills are not excellent. The external motivating factor comes in the recognition volunteers receive when their work is published on the website, though they can choose to remain anonymous. We do not gamify the act of contributing by maintaining a leaderboard or otherwise broadcasting participation on, for example, the home page. As part of the design process, we decided to keep participation information discreet and to minimise the comparison between volunteers which gamification encourages. This was based on our guess as to the type of volunteers we would get. We didn't think older volunteers would appreciate a gamified interface that compared them to others and possibly embarrassed them. We do, however, hope that within the Ukrainian community participation might reinforce community and vice-versa.

Early usage results

As for levels of participation, it is still early in the project, but our concerns about motivation seem to have been well founded (this paper reports on the results gathered in the first few months of operation). As for writing, about fifty one clips have been fully transcribed, translated, and published, close to 30 percent of the 176 clips mounted, but the audio

clips mounted are just a fraction of all the audio we have. Another nineteen clips are completed but still have to be proofed and published. Currently we have the "long-tail" effect where a few participants who have heard about it have contributed a lot to the site, and many have contributed little. In this case, one participant has completed thirty nine transcriptions and/or translations and another has completed twenty five; the rest have done one or none. This is normal behaviour on the Internet and it is clear that we now have to concentrate on either engaging more participants or helping the less active participants feel comfortable contributing more (see Anderson, "The Long Tail" [2004]).

We always expected that recruiting volunteers would be difficult – that is the challenge of any crowdsourcing project. Those in the community who have the language skills are older and not as computer-literate as youth. For that reason, we designed the site to make it possible for a volunteer like Elena to participate. Now comes the difficult work of iteratively reaching out to the community to involve more active translators. If the use of crowdsourcing is to be considered effective in this type of situation we must involve a substantial number of volunteers and generate continued interest in the site by updating it frequently with new submissions. Some of our hypotheses about how to increase participation include:

- We cannot count on people discovering the web site and volunteering without knowing about the project. We, therefore, need to find ways to explain the project more widely in order to recruit more participants.
- It may be that we need to rethink who the volunteers are going to be. One of the active participants is from Ukraine. Perhaps Ukrainians who are comfortable with the Internet and who know English are a better source of active volunteers.
- We now need to experiment with activities like workshops that help potential participants understand the project and develop habits of contributing.
- Perhaps we have to revisit the way we provide external motivation and find other ways to recognise activity. We need to do this, however, without alienating the community members already contributing.

Conclusion

The Ukrainian Folklore Audio Project is a humanities crowdsourcing project that was designed to engage a small group of volunteers, some of whom might be less Internet-literate than those involved in other crowdsourcing projects. In this paper we have described the design process, the web crowdsourcing application developed, our environmental scan and comparison, and the initial usage of the system. The slow uptake illustrates a fundamental truth about crowdsourcing – that you need to consider motivation and work with the community of potential volunteers carefully. Not all crowdsourcing projects are a success and this one may prove to be a case where crowdsourcing does not work. While we believe the web site design is accessible, the challenge now is to develop ways of encouraging active participation and expanding the pool of potential participants. Ultimately, the result may be that the pool of potential volunteers we can reach is too small or that participants are not willing to contribute to an online project. That said, it is still early in the process. Successful scholarly crowdsourcing websites take nurturing, promotion and adjustment into account. The crowd can be a powerful volunteer force but having an aesthetically pleasing and user-friendly web site is not enough: one needs to reach out to the crowd in different ways to bring the right people in.

Notes

[1] The project is led by Dr. Natalie Kononenko. Dr. Geoffrey Rockwell and Megan Sellmer managed the digital aspect of the project reported in this paper. Karl Anvik of the University of Alberta Arts Resource Centre programmed the application and Maryna Chernyavska worked on the original Ukrainian materials. The project was supported by the Social Science and Humanities Research Council of Canada.

[2] On this page you can see the other anticipated users that represent our starting assumptions about the types of users.

[3] From the Personas and Scenarios page of the open research wiki:
http://circa.cs.ualberta.ca/index.php/CIRCA:Personas_and_Scenarios

Works Cited

Anderson, Chris. 2004. "The Long Tail." *Wired*. Oct. Web. <http://www.wired.com/wired/archive/12.10/tail.html>

Benkler, Yochai. 2001. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale University Press. Print.

Catone, Josh. 2007. "Crowdsourcing: A Million Heads Is Better Than One." *Read Write Web*. Mar 22. Web.

Causser, Tim, Justin Tonra, and Valerie Wallace. 2012. "Transcription maximized; expense minimized? Crowdsourcing and editing *The Collected Words of Jeremy Bentham*." *Literary and Linguistic Computing*. 27(2): 119--137. Print.

Cohen, Patricia. 2010. "For Bentham and Others, Scholars Enlist Public to Transcribe Papers." *The New York Times*. 27 Dec. Web.

Cooper, Alan. 2004. *The Inmates Are Running the Asylum*. Indianapolis, Indiana: SAMS. Print.

Hars, Alexander, and Shaosong Ou. 2002. "Working for Free? Motivations for Participating in Open-Source Projects." *International Journal of Electronic Commerce*. 6: 25--39. Print.

- Howe, Jeff. 2009. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. New York: Three Rivers Press. Print
- . 2006. "The Rise of Crowdsourcing." *Wired*. June. Web.
- McGonigal, Jane. 2011. *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. New York: Penguin. Print.
- Mackay, Charles. 1995. *Extraordinary Popular Delusions & the Madness of Crowds*. New edition. Broadway. Print.
- Mahoney, Anne. 2009. "Tachypaedia Byzantina: The Suda On Line as Collaborative Encyclopedia." *DHQ: Digital Humanities Quarterly*. Winter. Web.
- Organisciak, Piotr. 2010. *Why Bother? Examining the Motivations of Users in Large-Scale Crowd-Powered Online Initiatives*. Fall. MA Thesis, University of Alberta.
- Rich, Laura. 2010. "Tapping the Wisdom of the Crowd." *The New York Times*. 4 Aug.
- Shirky, Clay. 2008. *Here Comes Everyone: The Power of Organizing with Organizations*. London: Penguin. Print.
- Surowiecki, James. 2005. *The Wisdom of Crowds*. Anchor. Print.
- Winchester, Simon. 2003. *The Meaning of Everything: The Story of the Oxford Dictionary*. New York: Oxford University Press. Print.
- Young, Jeffrey R. 2011. "Beware Social Media's Dark Side, Scholars Warn Companies." *The Chronicle of Higher Education*. 20 Mar. Web. <http://chronicle.com/article/Beware-Social-Medias/126813/>
- Zittrain, Jonathan. 2009. "Minds for Sale." YouTube Video of talk. Web. <http://www.youtube.com/watch?v=Dw3h-rae3uo>

