

Using the crowd: curated vs. unknown



BY BEN CHRISTENSEN



You've heard from colleagues, from industry news, maybe even from personal experience of the success of crowdsourcing. With nearly half the world's population online, it makes sense to tap into this tremendous resource to collect large quantities of data by breaking the collection down into micro-tasks that the enormous crowd of Internet users will complete for you at pennies per task. But you've probably also heard of cases of crowdsourcing gone wrong, cases like the British

government agency that put the task of naming its new research vessel out to the crowd and ended up with Boaty McBoatface [1]. So you have a human annotation task – a data refinement or evaluation task that requires human input – and you're trying to decide whether to take your chances with crowdsourcing. Where do you start? The first thing you need to know is that you have options; it's not Boaty McBoatface or nothing.

"When should I use crowdsourcing, and when should I use a curated crowd?" This is the question anyone interested in



The power of crowdsourcing is in its numbers. Many hands make light work.

Photo Courtesy of 123rf.com | Kheng Ho Toh

staffing human annotation tasks should be asking, but many don't because they don't even know there are two different options. So let's start there – defining the options. Assuming you need human annotation, for example for search relevance evaluation, there are two ways you can gather the necessary humans to do that work: 1) crowdsourcing, where the task is made available to a large crowd without any management beyond a very limited set of task instructions and possibly a simple screening test; or 2) curated crowds, where a smaller group is selected to complete the task accurately according to quality guidelines.

WHEN TO USE A TRADITIONAL CROWDSOURCING MODEL

The power of crowdsourcing is in its numbers. You can accomplish a lot quickly

because many hands make light work. A hundred thousand people can do quite a bit more than a hundred can. The cost is less because crowdsourcing typically pays only a few pennies per task. Most members of the crowd aren't trying to make a living – they're just trying to make a few extra bucks in their spare time. There's usually little overhead involved in crowdsourcing because the crowd looks after itself. You put the task out there, and if it's interesting enough and pays enough, the crowd will get it done.

This model works well for simple tasks that require little explanation and even less expertise. For example, you can ask the crowd to choose which of two images contains a dog, to tell you whether a business listing is or isn't a restaurant, or to transcribe words from

images [2] and have a reasonable amount of success. Opinion-based tasks where you are looking for a wide variety of responses are also well-suited for crowdsourcing: which image do you like better, what's the best Italian restaurant in your town, or how would you word this request for a voice-activated personal assistant.

CROWDSOURCING CHALLENGES

With the advantages of traditional crowdsourcing come a few limitations. First, quality control is minimal. Without this, you must rely on clear instructions, automated understanding checks and high overlap to get data you can trust. Overlap is important because there will always be noise in the crowd – bad data that you have to identify and sift out – so you'll likely pay for at least five members of the crowd to review each result. Some members of the crowd will try to game the system, using bots to do their work for them, so you'll need to account for this with screening tests on top of your high overlap.

The second limitation is your lack of control over task completion. The crowd can get a lot done quickly, but it will only get your task done quickly if it wants to. This means if your task is more difficult or less exciting than the shiny new task another team is offering,

then you're going to have to incentivize the crowd with higher payment to work on your task. The crowd has made no commitment to you and may not be motivated to make a deadline.

Finally, if you're looking for data in smaller markets, you may be out of luck – crowdsourcing is huge in the United States and a few other countries, but the same doesn't hold true globally.

THE ALTERNATIVE: USING A CURATED CROWD

Curated crowds, on the other hand, are all about quality. With this solution, offered by a small number of specialized crowd data solution providers, you have a group of people who are dedicated, if not specifically to your task, then to similar tasks. People in curated crowds become experts in search relevance evaluation, social media evaluation or whatever type of human annotation tasks they work on. This is not simply a case of counting on their accumulated experience to ensure quality, although that experience does play a large role. The key to quality is constant checks and balances. They are held to quality metrics, receive quality feedback, and are removed from your task if they don't deliver the required quality.

This means that you can use very little overlap, paying for each judgment

only one to three times instead of five or more times, because you can trust the data each person delivers. Curated crowd providers also monitor productivity and throughput, ensuring that the crowd meets their weekly, daily and hourly commitments so that you have the data you need when you need it. And if you've chosen a good vendor, then the manager will also be an invaluable resource, leveraging years of experience to partner with you in building out tasks and guidelines based on your needs.

With this higher level of quality and productivity management comes a cost. Curated crowds cost more than crowdsourcing because this work is typically a primary source of income. You also pay for the quality oversight that you don't have in crowdsourcing. Keep in mind, though, that lower overlap mitigates these costs because you aren't paying for each collected data point multiple times. Apart from the financial cost, curated crowds also require more of a commitment from you in exchange for the greater commitment you get. The curated crowd will be happiest and will keep their skills sharpest when you provide work for them consistently.

That said, if the natural ebb and flow of your need for human annotation necessitates more flexibility, there are

alternatives, such as sharing a flexible curated crowd with other teams running similar tasks.

USE CASES: SEARCH AND SOCIAL

Major worldwide search engine providers have been using curated crowd solutions for years. Curated crowds are used for search relevance evaluation, local search result validation, query classification, spam identification and countless other tasks that require more attention than what traditional crowdsourcing provides. By using this model,

**Defense Strategies Institute's
4th Big Data for Intelligence
Symposium**

Leveraging the Big Data Opportunity

November 29-30, 2016
AUSA Conference & Events Center, Arlington, VA

Free for Military & Government

To Register, Download Agenda, & Explore Exhibit Opportunities:
BIGDATASYMPOSIUM.DSIGROUP.ORG

these search engine providers gather high-quality data they can trust to accurately measure the success of their current algorithms, compare their search engine against competitors, and test out new iterations before launching.

Social media network providers have more recently come to appreciate the value of the curated crowd. While it's common for these providers to poll their users regarding their experience with the site, the social feed, the ads and the search functionality, the data gathered from these traditional crowd-based methods is limited and uncontrolled. By contrast, engaging a curated crowd that is able to provide targeted feedback on specific aspects of the social feed, filtering their subjective experience through a set of objective criteria, has produced much more useful data. Social media providers who take advantage of this model are able to leverage the resulting data to improve their social feed algorithms, their ads and their search features in order to create a user experience that stands out above their competitors.

CHOOSING THE RIGHT OPTION

So when should you use crowdsourcing and when should you use a curated crowd? Crowdsourcing is great for simple tasks that can be adequately

explained in two or three sentences. You'll get a lot done quickly, but be prepared to raise the pay rate if you have a tight deadline and the crowd doesn't find your task sexy enough. On the other hand, if you have a more complex task, particularly if it's a longer-term or ongoing task that dedicated people can build expertise on over time, then a curated crowd is for you.

Either way, be sure you fully understand your options so that you can make the best choice for your business. And if you're trying to name a boat, you might want to limit the vote to names you won't be embarrassed to paint on the side of that brand new vessel. ■

Ben Christensen, director of content relevance operations at [Appen](#), has been managing crowd-based search, eCommerce and social evaluation work since 2008. He has a master's degree in library and information science from the University of Washington. Appen is a global language technology solutions provider with capability in more than 180 languages and 130 countries, serving companies, automakers and government agencies.

REFERENCES

1. http://www.nytimes.com/2016/03/22/world/europe/boaty-mcboatface-what-you-get-when-you-let-the-internet-decide.html?_r=0
2. This is exactly what the Gutenberg Project does: <https://digitalciv.wordpress.com/2010/10/26/crowdsourcing-captchas-and-the-gutenberg-project/>