

Purdue University
Purdue e-Pubs

Charleston Library Conference

Crowd Sourcing of Reference and User Services

John G. Dove
Paloma & Associates

Tim Spalding
LibraryThing

Scott Johnson
ChiliFresh

Ilana Stonebraker
Purdue University, stonebraker@purdue.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>

 Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

John G. Dove, Tim Spalding, Scott Johnson, and Ilana Stonebraker, "Crowd Sourcing of Reference and User Services" (2014). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284315627>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

Crowd Sourcing of Reference and User Services

John G. Dove, Former CEO

Tim Spalding, Founder, LibraryThing

Scott Johnson, CEO, ChiliFresh

Ilana Stonebreaker, Assistant Professor of Library Science, Purdue University

The following is a lightly edited transcript of a live presentation at the 2014 Charleston Conference.

Slides and videos are available at

<http://2014charlestonconference.sched.org/>

John Dove: Hello, everybody. Welcome to the session on Crowd Sourcing of Reference and User Services. Some of this content is actually based on work that is reported in case studies in the book that Dave Tyckoson and I coedited for the Charleston Insight Series, but we actually have speakers for you today not just from there. So, now, first, my name is John Dove. I'm formerly with Credo Reference and who knows where I'll be next? My father would be really proud of the fact that I was involved with this, but also that as an industry are talking about crowd sourcing.

Now, I'm used to the fact that now many of you have a reference source right with you all the time, so that you can go and Google "Dove" and "Unicorn" and you'll see what my father's most famous work is, but if you went to Credo and looked up his name and his entry, you would find that one of the things that he thought was his life's work, he was a geneticist who then became a social scientist involved in logistics in World War II. In fact, one of his logistic innovations was something that we at Credo applied for a patent for, it's called Diversity Preference Ranking, and he defined this with something he called Agra Descendants. So, Agra Descendants is basically a way to properly manage with the diversity of group. It creates a better quality than any individual element of that group and that's sort of the core basis of what crowd sourcing is all about.

Now, crowd sourcing has a long history and a distinguished history in libraries and in reference. The National Union Catalog, 758 miles, I describe it as the only reference work that you can actually see from the moon. Come to the Boston Public Library, you can see the room, that's about as big

as this room that is lined with [inaudible]. I even said that, actually, if all the printed works before 1956 in libraries in Canada and the United States, World, as of 2008, only had 75% of those books. So, there's another 25% that, only if you went into those dusty old volumes, would you find, but clearly, the main thing about crowd sourcing these days is the example of Wikipedia. I've got two comments later about Wikipedia, but what I mean, the Oxford English Dictionary is an example of a crowd sourced reference work. For decades, they collected, basically, from amateur readers, who would be reading old volumes, examples of uses of words and there's a wonderful book, if you haven't seen it, called *The Professor and the Ant Man*, where the editor of the Oxford English Dictionary decided he should go visit the doctor who was his most prolific contributor, who was a doctor at the insane asylum for the criminally insane and what he discovered when he got there is that the doctor wasn't really a, he wasn't a doctor, but he was actually a patient in that insane asylum. So, that sort of brings up the old questions about madness and crowds and I'm going to get into that a little bit, but, indeed, the central challenges in terms of effective crowd sourcing, and you're going to hear some different answers on this one from our panelists and I hope that we generate some debate, but I'd like you to just hold questions until the end of all three speakers, and then we'll have time for some questions.

So, major challenges are some of the curation that you get when you just let everybody contribute, how to avoid systematic biases that might be there, and very important is how you get the crowd into the crowd sourcing, because we can come up with many ideas that if all if only we all contributed, we could create something of great value, but if you don't make the contribute

process fun and meaningful in its own right, you'll end up with a ghost town.

If you haven't read either of these two books . . . I'm going to introduce these two books. In this case, hopefully it's a contrast. *The Wisdom of Crowds*. This title comes as a reference to the madness of crowds, a book about how crowd behavior diminishes moral responsibility. So, you end up with mob behavior and what James Surowiecki did was building on that phrase, talked about what circumstances are there where the crowd doesn't become diminished intelligence, but actually ends up with enhanced intelligence and Surowiecki actually identifies multiple, a couple of key factors of what distinguishes between wisdom and madness. If you've got a group process, contributing groups actually independent of each other, because if they're not independent, then you won't get the benefit of independent outlook and you will end up with, eventually, systematic biases and you can look at any crowdsourced effort and say "Gee, this has some blind spots that are automatically inherent in the way that it's set up." Introducing a fair way of summing up the inputs of the individual elements that go into the crowd. So, it can be crowd source decisions, so to speak, and Cass Sunstein has written a book called *Infotopia, How Many Minds Produce Knowledge*, and he goes into a number of things that come from actually decision theory, one of which is the Condorcet Jury Theorem and the Condorcet Jury Theorem is very powerful. If you go by its assumptions, and that's the nub, you get 100% correct answers. So, if it turns out one of those assumptions is that if everybody in a jury has a 51% chance of getting it right and if you fairly sum up the inputs of that jury, you very quickly, it converges on 100% right. The dark side is if it's only 49%, it converges on zero. So, you have to build very carefully about whether you're assuming that all kinds of inputs are good or not, and I know if you put together a jury to discuss the question of whether Paris Hilton's jail time did damage to her career, if you put my wife and I on that jury, we don't know what career she had before, since, or after, so, I'm not sure you'd come up with the right answer to that question.

Now, with Wikipedia being the most famous example of crowd sourcing today, if you look at this from sort of these description and the requirements of, these are, it needs either, you know, systematic biases or independent inputs. What is this, the summing up mechanism? One of the formal aspects of Wikipedia actually leads you to madness. I mean, last in, first out is probably not the best judicial description of what constitutes quality, but you have to look also at all of the other apparatus that exists in terms of Wikipedia, the nonformal parts, the Wiki-manias, the organization of the community and that community's principles, and some of the adherence to those principles, because those are the things that actually take what otherwise, mechanically, would be madness, and turn it into something of real value.

You have to be careful as you do this evaluation. Two good examples of curated crowd sourcing is actually what the editor of *Birds of North America* calls "Wiki with gates," had a 19-volume *Encyclopedia of Birds of North America* and he came up with an outline version and what he did was describe how to actually build a whole community of bird watchers and ornithologists and the gates are the ornithologists. So, bird watchers can contribute observational data. They can't define a new species and, similarly, the *Encyclopedia of Life*, which is based on Wilson's TED Talk from a number of years ago, solicits and credits curators who will curate the inputs that then go into *The Encyclopedia of Life*. So, as you look at these panelists, think to whether, "Gee, are they, um . . ." Let me, back up a second.

I want to talk about the crowd when you're crowd sourcing. How many of you have ever actually written a Yelp or a Zagat review at a restaurant? Oh, a good number. Well, you can see that most of your colleagues have never done it and yet, I'll have, how many of you have used a Yelp or Zagat review? Yeah, it's just about everybody. It needs to be enough to write these reviews or otherwise there wouldn't be value for the rest of us. So, that's this key question about how you develop a proper crowd.

When you look at each of these panelists, think through these questions about have they

creatively solved this update problem, by making it valuable to the updater to actually do the contributions that we've come to value and have they had an effective, found an effective way to curate an input. You mention the Zagat reviews; you've got restaurants here who want to write great reviews and you've got his competitor down the street who wants to write really negative reviews. So, you've got to be able to have some mediating mechanism that allows you to have that negative and avoid various individual biases and systematic biases.

So, now, I want to hear this Tim Spalding, and Tim is going to talk about LibraryThing. I tried this experiment. Go to a cocktail party and go up to a stranger and say, "hey, I just found a great website where you can catalog your own books," and see if they don't roll their eyes.

Unknown Speaker: You go to the wrong parties, John.

John Dove: Indeed, then. I'll tell you what he said, when he started to do it, I put in my professional libraries that I maintained at Credo up on LibraryThing, so that people could see what's in it. It's like, you know, it's books I've collected over the last decade and a half, about reference and weird, murky reference books and I was putting in my first hand, I was using this little device you can get that reads an ISBN, you can get from LibraryThing, and in the first half dozen books I put in, I decided to look at one of the features they have, you can look and see who else has this book, because everybody puts their, not everybody but, you know, lots, most people put their library collections online, and one of the books I put in was only in two libraries in the world. One of them was from South Africa, under the name Jenny B. Walker, which many of you know, because she's been a frequent attendant at Charleston, so, it's just the kind of quirky thing that LibraryThing does. So, Tim? Shake my hand.

Tim Spalding: I'm shaking. All right. All right. Let me see what I can do in eleven minutes. Okay. I want to talk about LibraryThing. So, LibraryThing is personal cataloging, as John mentioned. If everyone catalogs together on a site that everyone shares, it becomes something called

social cataloging. If everyone does it, then it becomes social networking. Now, LibraryThing also makes a product for libraries, which I'm not going to be talking about at all, for risk of seeming commercial, but you can Google LibraryThing for libraries. The ladder of engagement is about how people climb the site, about how members for the site and I hope to show you these and other things, like quality engagement and about how you might go about adding crowdsourcing to whatever you want to do.

Personal cataloging is the basis of everything on LibraryThing. We started with the idea that it would only be personal cataloging, go out, you could catalog your books. Not only do people do it, as something else emerged. Here's my catalog. I add books. I use a scanner. I use the scanner to add tags to my books. Tags are the best way to catalog your books, to categorize your books, better than shelves. Members do this by the millions and LibraryThing has added over 112 million tags from users, which means that we have tens of thousands of people using the tag "romance." We even have tens of thousands of people using the tag "paranormal romance." You can say "romance zombies," "romance YA Greece." When you have this enormous combination of tags, something can emerge which nobody intended in the first place. People tag for themselves, and something emerges out of that. There are 393 covers that members have added for Harry Potter and the Sorcerer's Stone, four million covers overall. And here's the trick. They're not adding it to help other people, primarily. You can't add a cover unless it's on your own volume. You're adding for yourself, so you have all these covers and when you look at your catalog, it looks pretty. It looks like your books. Right? Well, the result of that is an enormous collection of covers. Exhibitionism and voyeurism. This was a photo that was submitted for one of our contests. Exhibitionism is, look at my library. Okay? There's lots of different ways of doing this, posts on Facebook, etc. Voyeurism is, "Hey, let's look at that person's library." Right? It's pretty low level with social interaction, but it is social interaction. Self-expression is the next step on the ladder and the idea here is, I'm not just going to share what I have. I'm going to talk about it. So, here's, for

example, is *The Hunger Games*. Members have written 2,600 odd reviews on *The Hunger Games*, categorized by language and *Mockingjay* in general has 2.3 million reviews. Now, this wouldn't happen if it wasn't something people wanted to do on the site. Okay? There have been many attempts to do reviews in library catalogs. My opinion is that most of them have failed quite miserably, because standing at a terminal is not the time you want to be entering a review. Not to mention it's a waste of time when you haven't read the book yet. You're looking for the book, okay? So, if you look at the ChiliFresh thing, you'll see that ChiliFresh has nine reviews for *Gold Finch* that he ordered in '96 and it's because people in LibraryThing actually want to do it, whereas in library catalogs, including the products that we make for library catalogs, they're just not that interested in doing it.

Social cataloging. Now, social cataloging is, in some ways, where things get interesting. It is cataloging on shared data. LibraryThing members want to have series rights. Series is one of the things that libraries generally don't do well. So, the library data we get and the Amazon data we get is not good enough for most users. This is the series page for *Star Wars* that members have settled by cataloging the series on their own books. Now, it's pretty awesome that there's 946 books in the *Star Wars* series but what's really awesome is the thing on the right, which is that there's more than 180 subseries of the *Star Wars* series, *Star Wars: Republic*, *Star Wars: Tales of the Jedi*, right? This is more information about how *Star Wars* books link to each other than any library in the world has or should have, okay? This is the power of people in their underpants at 2:00 in the morning, who care more about *Star Wars* than anyone in this room and that's great, okay?

Here's the example. We have a system called Common Knowledge. You can see some of the things you can enter: series, canonical title, original publication date, characters, races. You can do, for authors, you can say when they were. Okay? And this system has more than six part formulated edits. It's a fielded wiki. So, imagine Wikipedia, but every single field is a Wikipedia page, okay, which is acute data abundance, but it

produces this amazing abundance of cool data. Okay. LibraryThing members do all of the authority control on the subject. That may scare the hell out of people, but it's true. All of the work-edition dissemination, what's called Ferberization, right, are these two editions part of the same world, is done by members. They've done more than four million acts of dissemination. All of the combination of separation of authors, right, Stephen King, Richard Bachmann, all of the homonymous author division. There are 39 John Smiths. LibraryThing members have figured it out. Okay?

Tag disambiguation. So, on LibraryThing, members combine tags. It's one of the weaknesses of tags, is there's all these different parts, and on LibraryThing, on some level, the tag World War II is also WW2, WWII, because members have made that come together. This is, well, does it work? It works really well. It, we put our information in OCLC in 2008. It's even better now. Here's some organizing all the John Smiths in the world by works. Okay.

Policing and helping. At some point, people get invested enough that they want to police and help. That's an essential part of the LibraryThing experience, is a group devoted exclusively to fighting spam. There's a group devoted exclusively to work combination issues, which is mostly shutting down people who are doing it wrong. And there's people who really spend their lives on LibraryThing, improving the data and, but God bless them. I don't think of them like, "hey, you work for us," is just, it's interesting for them, so they do it and that's how we do it.

The top level of the ladder of engagement is collaborative cataloging. So, this is LibraryThing members that catalog the libraries of more than 250 dead people. The rule is they have to be dead and they can't be [inaudible]. So, Tupac Shakur to Thomas Jefferson, Jackie Gleason, Marilyn Monroe, books that Darwin had aboard the Beagle, right? Huge number of presidents and all kinds of random celebrities, too, and this really has no personal value to anyone, but once you get into LibraryThing, once you start enjoying it and you run out of books, right, you still want to do it. You start getting into projects

like this. Another time, members cataloged all the books that were in the movie *Dr. Horrible's Sing Along Blog* by freeze framing and cataloging in LibraryThing all the fictional mastermind's of the singing superhero.

The ladder of engagement moves in different ways. It moves from the personal to social, it moves from the love of the thing, and don't underestimate the love of thing. Love of thing is everything. Use your map to mark the other. The love of self to altruism.

LibraryThing is a social network with the primary connector between people is the stuff that their interested in. Right? That's not how everyone works. Right? That's not what makes the world go 'round in general, but, there are a lot of people who want to talk about books they have, even with people they don't know, a lot of people who want to work together on books, because they love them. Low interest people or high interest people. Uh, many, few, right? So, every time you step up the ladder, people drop off. There's a lot of people, about half the users of LibraryThing only use it to personally catalog and don't do anything else, right? Now, in, this is bad in some ways, but it's also good in other ways, because as you step up the ladder, you tend to get geekier, more committed, smarter people, okay? And, you know, the reviews on LibraryThing are just written really well because the people on LibraryThing are people who love to write reviews, the cataloging and so forth on LibraryThing, and somebody who will really, really care about it.

Here are the lessons. Last slide. Secure the bottom of the ladder. The bottom of the ladder is the most critical thing. If you want to do something in crowdsourcing, you need to start with what you're going to get and all the wonderful engagement that people are going to have with each other, you will not get to the top of that ladder. You can build it rung by rung, but when you build it rung by rung, you should think about it rung by rung. Each rung needs to make sense. Most of all, in library context, crowdsourcing is not a feature. If there's one line I want you to take away from this, is that it's not a feature. Over and over again, things like tagging are added to library catalogs. Look, we have

tagging! No one's using it. It's worthless. That tagging is worthless if you've got 20,000 tags. Okay? So, above all, it's not what you get, it's not about what you get, it's about what you're giving to the people who do. Thank you very much.

John Dove: Thank you, Tim. I'm going do something that's really scary, because I made a mistake this morning and actually got the wrong slides for Scott. So, I'm now going to pull up the right slides for Scott. And Scott is somebody who actually, no matter what slides you'd have in here, he'd be able to hold all of our attention. So, but I mean, he should have the benefit of the slides that he wanted, so . . .

There we go. Now, now that that's done, I'm going to introduce you. So, ChiliFresh. A lot of people haven't heard about ChiliFresh, and yet it's in 4,000 libraries and it's actually done some very creative things in terms of how to deal with the curation. If you're going to allow input from various patrons, you know you're going to have problems in the sense of, this is Nazi trash and somebody else will say, you know, so, you've got to be able to say, is this a commentary that you really want to continue to have in your social media or is it something you want to discourage or, more likely, it actually does something that is really good and then the local librarians, they would promote that to the community of the 4,000 libraries that have ChiliFresh, so that everybody, every library benefits. And one of the things that ChiliFresh has had to do in order to make that possible is that they, a lot of people talk about user experience, and mainly they're thinking about features, which are clearly very important in terms of the user experience, but if you're going to ask librarians to do some additional work, then you better really think through how the user experience is for those librarians. And not only that, but you probably want to think about how, in fact, the work that they might be doing for you might actually augment work that they're already doing for the general purpose that they have.

So, ChiliFresh has done a really good job in terms of thinking about how their involvement in a library can enhance the very activities that

librarians are already doing. I think it's been a real secret to its success, so, Scott.

Scott Johnson: As he said, my name is Scott Johnson. We're actually in Kansas City. How many of you have not heard about ChiliFresh before today? I like that. You can't answer the question that way tomorrow. Let me tell you why that's interesting to me. Because we don't brand what we do. Our brand doesn't exist in your catalog or in your library anywhere. Our brand only exists in the concept, meaning that your users, your students, your patrons, they never see the word ChiliFresh. They never see who we are. They only see the functionality and they only see the platform now.

So, before I get into it, I want to talk a little bit about crowd source and the importance of that, that collaborative data. As I look around the room, I think I'm probably the oldest one here. One question that I have is how many of you remember the 1970s? Not how many were there, how many remember it? I remember a little bit about them. Most of you were probably in diapers. I was riding my bicycle. How many of you had a waterbed in your house in the 1970s or 80s? How many of you knew people who had a waterbed? What was that all about? At the time, waterbeds made it to 20% of all the households in the United States. Today, how many of you know somebody with a waterbed in their house? Oh, there's two people. That's funny to me. If the internet existed in the 70s and 80s, if there was a review system allowing people to read and write their comments about waterbeds, you think waterbeds would have reached a 20% penetration? I don't know the answer to the question, but I also know that during the 70s, everybody on my street had a station wagon. Do you remember that? Everybody had a station wagon. How did this happen? I believe, and the question was asked, I believe that the wisdom of crowd sourced information is also the madness of crowd sourced information.

I also believe that each one of your users are seeking to connect with other people, with data, and with this crowd sourced message when they go to your catalog and this is how I know. We exist, ChiliFresh exists because I was standing

behind a reference desk at a public library. As I was standing behind the reference desk, the reference librarian was sitting there. There was nobody in front of her, she wasn't helping anyone. She had her computer on and there were two windows open on her screen. One on the left was a Searcy Dynex catalog. The one on the right was Amazon. And I watched curiously from behind her, because I really didn't understand what she was doing. And I walked around front, I knew her, her name's Sarah. I said, "Sarah, what are you doing?" She said "I'm just looking for my next read." And I looked at her a little weird. I said, "I don't get it." And I won't, granted, I'm a simple man. You work at the library, but you're going to buy this book at Amazon? And she said "No, no. No, no, no, no. I look through the catalog here and then I read the reviews about what people say about the book and that's how I decide if I want to take a book home." My next question was, I think it was fairly obvious, "Well, why aren't your reviews in the catalog?" She looked up at me, like I had spiders on my face, and she said, "That's not possible." And you know, I took that message home and I learned very quickly that it wasn't possible, and the reason I saw that is because of what Tim discussed earlier about this critical mass and the same thing that John talked about. How do you get enough people to do this?

Just after we started ChiliFresh, one of the largest library systems, not only in North America, the world, by circulation, King County, called me up and says "We need your product. We need to put patron reviews in our catalog." And I opened up their catalog while I was talking to her and, and I said "You've already got reviews in there." She said "Yeah. It's a contained system. It's just our patrons writing reviews." And we did some math. We learned that, if keep the system going like it is, it'll be decades before they reach any kind of critical mass. And so what we did is we created a collaborative database of reviews that allows a library to participate in this database of reviews that connect to an ISP or any unique identifier to catalog. So, there's somebody in Seattle, in King County writes a review, that review's visible in a library in Miami or Salt Lake County, or in Australia or in the UK. And all of a sudden, it made all the sense in the world. It solved a lot of our

problems, like whose going to be posting reviews? Well, the truth is, everybody wants to read 'em, and this many people will write them. I mean, you proved that by raising your hands earlier.

So, what we've done is we've kept all of this in catalog. Every review in the ChiliFresh system has been written by a library patron in a library catalog and it's been moderated by a librarian. I'm not sure there's many libraries that want to have a Viagra commercial in their catalog or a school librarian who wants the F word showing up in the catalog, but every public school in the state of Ohio has the ability to allow their students, empower their students, to read and write reviews. The system can be as open or as closed as they want. That means that they only want to show reviews from their schools. A library may want to show just reviews from the US. Maybe they're just like that. So, anyway, we've moved past all of this, trying to figure out what is community, because community, I think, is key in this crowd sourcing.

Back to the 1970s, when I was riding my bike, community was about as far as I could ride my bike. There was a grocery store, there was, there was a police station and there was the library. That was my community. My parent's community was their workplace. That's as far as they stretched. Every once in a while, they'd go on vacation. They'd read books to try and create community. Well, community today is completely different. Because of the internet, our communities are interspaced. They have to do with our hobbies, our reading interests. It has to do with where we travel. It has to do with so much more. I swim every day. I like to try to talk to other people that like to swim. Well, there aren't that many of them, but I can find them online. You might like to do other things too that you connect with people.

So, let's talk about this for just a minute. What is the power of the opinion? And we're talking about a group opinion here. The power of the opinion, in my humble opinion, is as strong as any. Every revolution in the world took place because of the crowd. Every change in the world took place because of the crowd. So, we have to share his collaborative data, but how do we know that it's

good? How can we tell if it's good? I think that, as somebody looks through this, they can see very simply. It's why, when we buy a new computer, we're not taking it home with us until we read about somebody who already has it under their arm. How many of you ever bought a washer and dryer without reading a review on it? And it's like every one of your users, if you don't have the ability to read and write reviews in your catalog, everyone of your users, students, it doesn't matter, researchers, it can be patrons, they are all seeking peer comments of what they want to do. And they will open up Amazon or decide to go open up other sources. These activities should be taking place in the catalog. So, we move this a little bit further. This is what the review content looks like here. The way we have done this and the way we have been able to put ChiliFresh content and functionality in libraries is by integrating into the library, the IOS software. And this is really key. And when you talk about personal cataloging, imagine your catalog is the place that not only houses your connection and your information, but it empowers your patrons to catalog their stuff right beside it and share their stuff right beside your collection and then communicate with other users on a global scale, based on common literary interests. That's the power of what we're doing. We never drive your patron out of your catalog. We only build functionality into your catalog that empowers them to communicate and connect on a global scale. This is an example of how it's integrated into the [inaudible] catalog. A person can put in their profile information, they can, they've got bookshelves, which is cataloging. They've got friends and followers.

I can just go through these again. When somebody logs into the library catalog, it automatically logs them into the ChiliFresh network. They don't have to go to ChiliFresh.com to do anything, ever. They only see a connection inside your catalog. I'll just click to a user kind of quick. There's ways to add things to their catalog and to their bookshelf. There's ways to tag things and comments they can share and here's what happens. When people start communicating and making recommendations back and forth and it's a global communication that's going on, we

capture that data and make it available to you that you can use in your collection development. You'll also find that our review engine appears in Brodart's catalog, when you're buying your stuff. It appears in Ingram's Catalog when you're buying stuff. It appears in . . . some of these, so basically, we also have functionality that interacts with your users where they live. This all lives in their pocket with mobile apps. It lives on their Facebook page. Your catalog alongside their catalog in all of these places. You want to keep them in your ecosphere, because that's what's important. Thank you for your time.

John Dove: Thank you, Scott. Now, I've got to switch back to the other presentation. Now, what about the idea of crowd sourcing reference or health questions itself? This has got to be kind of personal, given the questions about what is the competency of people who can provide answers to people in the library systems. I've been giving a number of talks over the last few years about user centered designed and online reference systems. A common question I will ask early in the talk is where do people turn first when they have a reference question?

Unknown Speaker: Their friends.

John Dove: Who else? Very few people ever say their friends. People will immediately say "Oh, well, Google" or "Wikipedia" or "the library's website" or, but in fact, if you look at this anthropologically, people first turn to whoever's within shouting distance that they haven't annoyed yet. "Hey, who won the World Cup in 2008?" You know, so, or "How do I make this thing work?" It just happened to me this morning. So, Ilana Stonebreaker's going to describe a very interesting approach to crowd sourcing of questions in the library, too.

Ilana Stonebraker: Before I get started, it's Saturday morning, and you guys are here and I really appreciate it, so we're going to watch a short video. Have, do, are any of you people familiar with the Vlog Brothers? Yay! Okay, that's it. It's an online community out there. And we're going to watch a short clip of jokes.

Video: It's been over a year since the last time I told you jokes, which means that it's time for me to try to beat my record again. This time, 54 jokes, and I'll do it in four minutes. What's the difference between a cat and a compound sentence? One has claws at the end of his paws. One has a pause at the end of its clause. What's the difference between a tuna and a piano? You can tune the piano, but you cannot piano a tuna. The difference between the moon and Julius Caesar? The moon is rocky and full of craters and Julius Caesar is dead. Why do you think was "Civil Disobedience" such a fantastic essay? Thorough editing. Thoreau. What cell phones do traveling nuns use? Virgin Mobile. And how come her cell phone bill was so high? She was a Roman Catholic. Why did they kick Cinderella off the baseball team? She kept running away from the ball. And the mermaid. That was weird. What was she wearing in Math class? Oh yeah, an algae-bra. Why was the sand wet? Because the sea weed. The sea, it weed. What happened when the butcher backed up into his meat grinder?

Ilana Stonebraker: Okay, all right. So, that's, that's the video, and you may be asking yourself, how is my library like the video "54 Jokes in Four Minutes?" The answer is that the Vlog Brothers is a very active and vibrant online community, does a lot of crowdsourcing, a lot of crowd funding and things that they do like "54 Jokes" video, which may seem trivial, are part of supporting a larger community. So, the things that you do above the water, when you answer reference questions, when you, you update your catalog, are supporting a large community of learners that are bound to the library. So, I would say that each and every one of you supports online community and I would say online community, because a huge amount of what you do supports communities which may have never spoken to each other or you may have never seen in person but, nevertheless, is very vibrant. Some examples of some online communities that you may support. So, you are probably physically located somewhere and, your library is, and you support online community about your city, your school or organization, there are many reference questions you may probably answer about your school or organization. Your alumnae or retirement base.

You may support them. The fan, people who are just fans of libraries, like Neil Gaiman, who's always talking about his love of libraries. People who are just fans of reading who want to contribute to these systems that you gentlemen have talked about. And also your collection strengths, the things that you collect, creating community as well. And so you support these online communities. So, what I'm going to be talking about is some specific ways that we've looked at some specific problems involving reference in a R1 large academic library.

I'm going to start with four not shocking facts. So, everyone uses the internet, our patrons are part of a community, and they are, we are supporting, and then students don't read manuals, and the majority of reference questions are lower level. Where do I find this? Where are the printers? Those sorts of questions, but there's also lower level online questions, right? This doesn't work. Why doesn't this work? Where should I check for this person? So, what's supposed to happen at a reference desk is that students develop questions about which they ask the librarian about. The librarian, at a one-to-one level, answers that question dazzlingly well and then the student rocks that, so the next point is, thanks the libraries forever and gives the library a million dollars. What actually happens, and you can start this at any level, is the student starts their path, they find that they can't do whatever it is, they find at least one resource that works for them and then they try and like make that work for all their projects, that's the cat trying to get into the box. And the cat in the box is very important, because I think that, when they look at what sorts of resources they're using, they may not be aware of what even their other students have found.

How many of you guys have had the experience at a reference desk, where you have a long line and the students start helping each other? You know, they're like, "Oh, I had that English 106 class last week." "Yeah, what'd you do?" You know? And that's what we want to happen. When we create our online reference transactions, we're not allowing the line help to happen. So, also, reference service models only assume that the librarian can get to that answer, which is not

always true. People turn to their friends, to other students within the class, to the professor, to give them all sorts of different types of answers. Our questions are also all treated alike. The majority of reference questions, like I said, are lower level, but they're also context-based. People don't generate questions about nothing. They question, they generate questions, especially in academic environment, because they're all part of some sort of online community, be it a class, be it a group of projects, and the process of reference decontextualizing those questions from the environment in which they were asked. If you don't then put that context back in and the use, they don't utilize graduate students, instructors who may have additional information or may be able to answer the question more specifically for the user. And it neglects to think about that we live in an information ecosystem, where we give excellent answers but we don't exist, the library doesn't exist just to answer questions. We're trying to provide help; we're trying to help our users get things done faster. So, we want to use all the information to help us that we can.

And then there's a really excellent book out on MIT Press from 2013 called *Crowdsourcing* and it has a really excellent definition of crowdsourcing, which is an online distributed problem-solving and production model that leverages the collective intelligence of online communities to serve specific organizational goals. So, you need a specific organizational goal and you need a specific community. The crowd is not just the crowd. The crowd is your community. Additionally, the locus of control regarding the creative production of goods exists between the organization and the public. So, you don't own your crowdsourcing and neither do your users. It's rather collaboration between you and this is an excellent place for libraries because libraries are deeply collaborative. Projects of collaborative cataloging have been going on for as long as I've been, well, I'm very young, but you can, you know, if you've been, it's been, the important elemental part of our mission is being collaborative. So, I think that we're a really great location for crowdsourcing.

I'm going to talk about a specific example, which is an IMLS-funded project over the last year over at Purdue. Our community was the students, staff, and faculty at Purdue University. Our specific goal that we're trying to do with crowdsourcing was provide contextual answers for students and also alumni questions. So, when they wanted to know when the school mascot got from this big to this big, or where they can buy a bobble head, or when the next football game is, and where people are tailgating, that's the sort of questions that we're interacting with as well, to strengthen alumni networks, which are also part of our community. So, this is CrowdAsk. It's very similar to, if anyone uses Fact Overflow, it has similar gamification and packaging. You log in using a variety of different types of providers, and this is also, we really, this is open source on GitHub, so you can download this as well and we've branded it with all of our information, so this is a lot of Purdue gold and silver. You can ask a question on the system and you can ask a question. You can also assign a bounty to that question. So, it's a point system. So, you can give out more of your points to get an answer faster. You can assign a category, be it, we launched it in a group of lower level English classes and information literacy classes, so students could work together, as well as we had launched it through the alumni networks for Purdue History, in collaboration with the special elections. This is an example of a question within the system. It's citing references using APA format. You see that users can vote on answers and questions. You can see that one of their, the number one voted answer, which goes to the top, is a link to Purdue Owl, which is a very well-known Purdue source and they, and the students may say that that is probably the best, and other people have also given other sources they use for citing in APA format as well. So, this is also an example of what a user page looks like. You can see this student has a number of badges. There are badges that are implemented, such as good question, knowledgeable, good answer, type-thing. These all have to do with different cases of it. Additionally, this is a meritocracy, so, the more points you get the higher badges you get, the more power you have within the system. So, this is encouraging students to be really good at answering each other's questions.

So, next, I want to talk about some, some stats, because usage is important. So far, since we've launched it, we've had 184 users post questions and 129 users post answers. This is all within the last year. 257 people voted. There are additionally more questions, so people are repeat users. The most views on a question is 182. I think it's still the MLA question. The most answers on a question is 16, the most votes is 48. Additionally, something interesting, and if you do any studying of online communities, there's a whole group of people that are lurkers. So, a lot of you guys would have that on Yelp as lurkers. So, you use the system, but you don't contribute to the system. But, you can see that the average amount of time, from Google Analytics, is six minutes and seven seconds, which means that people just hang out on the page and we had a lot of users, of people who were just getting answers from the system, in addition to contributing, which, lurkers are a very important part of online communities. They build knowledge and they lurk to market that community.

We did usability tests of four students, two novices, who'd never used it before, and two expert users, who were really into it. What was really interesting to us when we asked the expert users why they contributed is their motivation wasn't the points, though the points were cool. What they wanted was reciprocity. They said, and I quote one of the users, "Someone helped me and I wanted to help someone else." So, I think that use, I think there's a lot of optimism. There's the, there's some very earnest users out there who have been helped who want to help other people. Reciprocity seems to be a much more powerful tool than gamification when it comes to crowdsourcing. They want to help each other. They want to be part of a community. So, this is part of kind of a larger goal, and within this is our, once again, our first speaker talked about it, crowdsourcing not being a feature. I totally agree with that. Our goal is to develop sustainable user engagement and community involvement as part of the Purdue University's Library website, not to just answer the questions or to load off some of the late night reference to students who are all working at that time, but it's also just to cultivate a community of learners

who help each other, who can answer each other's questions and ask better questions, because they have engaged in a reference interaction from both sides. Yeah, thank you.

So, as I'm kind of finishing up, I just wanted to reiterate that crowds are, you get this idea, when people talk about crowdsourcing, like they're this mass of kind of infectious zombies that just come out of nowhere. I think that, really, crowds are probably people you know. Crowds are helpful students who really love the library. They're people that you know that want to help and this is merely an invitation for them to do so, to be part of the system, as well as strengthen your relationship with your community, to move some of that iceberg above the water mark. Some keys that we found, overall for crowdsourcing reference help, the stronger the online community for the update problem, the stronger the user base. The stronger that community is, so classes are a very strong community but for a very short amount of time. Alumni bases are very strong communities, so finding those communities and then cultivating those resources. Crowdsourcing can also work to strengthen an online community, by bringing people together who did not necessarily know

that they had common interests. And I think the library, as a conduit for that interaction, is a fabulous contribution to our mission. And then, once again, reciprocity is important to these communities. They want to give as well as to take. So, what this means is students, even students, even alumni, they want to feel as if they can feel the impact of what they do. They want to help people. So, we need systems that can make that more clear. I think it's an important part for those users. So, this is an ongoing project, so if you're excited about crowdsourcing and you want to help or you want to help develop further, once again, it's an open source code, so if you want to develop it or just do interesting things with it, if you have a community which is, you want to try it out on, or even if you want to come up here and tell me that your users would never use this. I'm interested, and I love to talk to you guys afterwards and talk about possible weaknesses, threats, opportunities, any of those things. Please, I'd love to have this be much more a conversation, so we can talk about better ways to kind of strengthen our online communities. So, that's it. Here's a link to our code and a short video on CrowdAsk. Thank you.