

*These transcriptions may contain errors, especially in spelling of names. These are unfortunate, and we regret that we do not have the resources to fix these errors. Still we believe these transcripts will be valuable to many users.*

>> David MacDonald: Okay. So again, I recognize I'm the last talk of the day and you've been sitting here all day. No? Oh there's --

>> Hours.

>> David MacDonald: Hours more. Oh. Well, as you get near the end, the challenge is to try to help kind of sum up of a little bit of what you've seen during the day. And there's some things I want to try to sum up. And that is, you know, I hope when you go down through your career you continue to have this excitement for what you do and what you study and that's just really important. And I think one of the things that you've seen with the people who have been talking today is how excited they are about this. They want to talk to you about this stuff. They want to interact with you. Don't be a shy doffle [phonetic] student and not come to these senior people and talk to them. You have to do that. But that also is not just in your research but in the way in which you build like the research community. And I appreciate the comments that Jenny made about the service to the NSF. That's a really important work that we have to have good people that go to the NSF that really are interested in building research communities and building interesting topics. It's probably not going to be TMSP say 50 years from now. It's going to be something else. Right. So research and everything progresses. And we have to think about setting up that. What's the future of those kinds of things? And so contributing to the research community is really important. And the last one is you can tell how excited people are by the number of words spoken per minute so far, right? People speak fast when there is so much they have to tell you. They're excited about this thing. And so I was paying attention to a community that has lots of words spoken per minute so far, right. Wow. So another one that I had been thinking about is you become what you measure, right. Terra-contris and they get thank you's and how that starts to relates to things like the real world impact. We want to understand and shape or improve the situation. I think just understanding isn't really moving things forward. I do want you to think about taking action with that understanding and I hope maybe a little of my talk will show you like some ways that you can start doing that so I'm going to try and deal with a little of that. And then this other piece is you are what you measure. So you know, I started thinking about things that were harder to measure in Wikipedia. And one of our early studies of Barnstars. How many of you have ever heard of Barnstars? Okay. So they are a certificate of appreciation. Thank you for doing this really valuable work. Right. How many of you get certificates of appreciation for -- you did nothing. Okay. How many of you get certificates of appreciation when your employer, you know that your employer could not care at all? [Background Sound] Okay. So there's something about an acknowledgment that's important. Community acknowledges -- somehow acknowledge things that they value. And so in this work, we were trying to understand what are the -- what is the range of work that was being valued by Wikipedia? And just to kind of fill that out, there's actually a lot of different work in Wikipedia that's valued

and a lot of this work is very difficult to just simply count. All right. And so -- because actually on Wikipedia everything is actually just an edit. All right. And so if you fall back to edit, you made a change to some text and then a lot of this gets lost. So this is the charge to not just think about the things that you measured but try to make those things that matter. All right. I think that's what Ben was actually saying in the beginning of the day and so yeah. So that's kind of where I think things are heading in the space around TMSP. But what I want to talk a little bit more about today is some work that we've been doing. It's ongoing actually. It's a set of design explorations to realize social translucence in Wikipedia and there were multiple students who worked on that and I'll acknowledge them in a minute but the longest running collaborator on this is Mark Zachry. And he works out of Human Design and Engineering, that department at the University of Washington. And so just a little bit about what it is. I'm going to talk about the motivation is to give you a brief definition of social translucence, talk about reflex system and I'm going to show you a live demo. And those of you who love live demos should be thinking I hope it crashes because there's nothing like a crash during a live demo. [Background Sound] Yeah. And what does -- how does that presenter respond? And the future directions and that is, the future directions is actually a bit of work that was just recently funded about these things called voluntary virtual teams also kind of characterized as self-managed work teams. I think you're going to start seeing a whole set of connections. I mean ironically, a whole lot of connections. Yeah. So what's the motivation here? Well, if you think about mediated collaborations and like online space, often that attenuates our ability to understand what the behaviors are. It's difficult to interpret and understand what's happening. So the example of that in Wikipedia is that editors can interpret every edit. They can sit down and look at every edit that someone else has made. But actually, validating and understanding each of those edits is a little hard. I mean, just because you can see every edit doesn't mean you've really formed a good interpretation of it. As well in these systems, of course, as you scale, as you move, as Wikipedia gets larger, right. This year, the absolute numbers of people in these large scale systems and the way in which they collaborate makes it difficult to comprehend the massive numbers of interactions. So you've got massive numbers of interactions and people have a hard time keeping all that in their heads. It's really hard to interpret. So yes, editors and Wikipedia can actually look at every edit but they really don't have the time to truly inspect the edit and every edit of everyone that's edited with someone else, right. Because it becomes this sort of geometric explosion of the number of people and number of edits you have to look at. So what we might want to do is help them understanding, comprehension, and interpretation. That is, we might want to help -- we might want to build systems that help people who are using those systems be productive. They want to form some understanding. So social translucence is one of these design ideas coined by Erickson and Kellogg in 1990s. Social translucence is three key design characteristics. An idea of mutual awareness of activities. That you have the contextual propagation of socially salient cues in the way that Kellogg and Erickson talk about this is they call it visibility. Right. But I think you want to think about how you propagate these cues. And then the last one is the accountability for your actions. So that is, that the community or another individual, when they see your action and

they see something that you've done, they can interpret it, try to hold you accountable for it. There might be still some debate about whether it was good or bad but at least you could be called out for saying I think you did this and maybe we should talk about that. All right. So those are three characteristics. A lot of way -- social translucence skits that are sort of enacted are in these sort of pictures. So this picture on your left is of like electric hall. Like, there's sort of someone that's kind of like the focal person in the scene and then there are these other people who are sort of sitting out there and when someone asks a question there's some you know activity by the audience. Do you guys know anything like that? Did you ever [Background Sounds] is that memorable? No. All right. Good. Another one might be a chat room and then you might think of an online line waiting for help. Right. You know, technical support. You are call number 400. We will get to you in an hour and a half.

>> You are very important.

>> David MacDonald: You are so important to us. Yes. Exactly. Okay. So if we're going to think about translucent design and maybe some of the questions you want to ask, so this idea of translucence is it really a system? Do we have translucence as a system of translucence or is it more like a feature? Is it something you just have a little bit over here? Right. Is it something that you really have to have that you built from sort of the bottom up? That is, if you don't think about translucence, could you never have it? That is, do you have to imbue your system with it? I mean I'm willing to accept that maybe a designer wasn't really thinking about translucence but wanted some kind of property, wanted someone to see some of someone else's behaviors and so maybe you could have a system where it wasn't necessarily intended to be social translucence but they had opened up certain aspects of what you might want to project. So bottom up or can you just like slap something on the top and get translucence, right. And then do you just show the raw data? Do you just show like exactly what happens or do you somehow compile and distill this data down to make it somewhat kind of interpretable in some way? Because I mean raw data is interpretable. You can see that with Wikipedia. People do use those edit box. So it's not obvious that you have to have some kind of distillation. So in asking those questions, what you're really beginning to ask is so what's the shape and extent of a socially translucent design space? What does that design space look like? What are the things that are inside? What's outside? What are some of the useful dimensions? How should we speak about a socially translucent design? What are the dimensions that I might lay things out on? And how can we analyze and understand a socially translucent design? What does it mean to have a design that exhibits some properties of social translucence? And again, some of those questions upfront would have different indications for this. So we wanted to understand this a little bit more. And there really isn't a lot of work on exactly how to do this. There are a bare number of papers that claim that the system has used social translucence or illustrates social translucence but no one had actually set out to decide, well what does the social translucence design space actually look like? So what we tried to do with this thing called a domain analysis. How many of you have heard of a domain analysis? One, two. Okay. This comes -- this term comes from software engineering and it's a way of looking at all of the implementations of

something. Like say, all of the different ways we might implement spreadsheets and as a software engineer you analyze that and you figure out, well, there are certain types of data structures you need, certain types of algorithms underlying it, certain ways you present the data, right. And you basically kind of articulate all of the different systems and all of these different features and you start trying to map it out. And so if we were going to do that for social translucence, I would get a great picture of systems. What would I be missing?

>> Users.

>> David MacDonald: Users. Right. I mean it's social translucence. It's not technical translucence. So what we tried to do then was push into this idea of socio-technical domain analysis. We're trying to account for not just what the systems think but how do these systems try to interpret or show behaviors for people? What are they trying to do in that space? So yeah. Account for both the technical and interpretations as well as social environment. And so in our attempt to do this and thinking about where we were heading, we came up with two particular dimensions and a dimension around action and a dimension that is around the levels of system interpretations. What are the systems trying to show in terms of interpretation? And I'm not going to go through all of these because there's a lot there. But we kind of marked it up into a framework and this particular paper which is buildings for social translucence, domain analysis and prototype system at CCW this last year and so February. Kind of marks out how do we start thinking about this, what are the spaces? And so we kind of have these two dimensions and so yeah. That's thinking about social translucence. But then what you might want to do is not just have those domains but you might want to take those domains and try to put those dimensions and try to put them into actions and try to do something with them. And so we pushed this into two things. One is idealized architecture which we call re-arch. So an architecture for translucence. It attempts to sort of push translucence down into the system. Where -- what would you have to implement down below the system to be able to do translucence well? And you know so that's one of the things. And this reflex which is a visualization in UI Platform which basically tries to say so if you were going to try to visualize something like this how would you sort of illustrate translucence built on top of something like a reflex architecture? So reflex and re-arch. So now I told you initially what the motivations were in trying to help people make interpretations of the behavior around them. And one of the things that's fun about being engaged in a research community and working on systems is that you find people who then identify problems even though we had a problem, but you find people who identify problems and you go that's a good example of what you might want to do. And so it turns out that one of the people here who worked with Brett, who happened to work with Brian Keegan, talked a little about this as well. And so this is based on Hot Off the Wiki dynamics, practices, and structures of Wikipedia's coverage of the Tohoku catastrophe. The thing that's kind of interesting is you might want to do and this is a little bit of what that work does well are there some kind of editing styles here? Are there -- could you look at people and see something about what it is that they are doing and potentially, who they are doing it with. Right. So that's one of the examples that reflex does relatively well. So let's see if we can

get it to pop up. Oh that's not very nice. So I'm going to get -- maybe I can do something else here. Let's see. [Silence]

>> Working at MSF, it has a good background for taking risks in your talk you said you developed lots of risk taking?

>> David MacDonald: Yeah. Willing to take risks. Here we go. And so this is -- this is a reflex. And what it does is this is a live Wikipedia page. So we're proxying that page. We inject our code into the page and setup the set of things that then sort of interact with stuff that's on the page, interact with servers that are remote. And so the reflex toolbar sort of sits kind of at the top. And we show you two different types of relationships. We'll visualize two different types of relationships. Person-to-person or editor-to-editor relationships and editor-to-artifact relationships. We'll let you pick you four different contexts from which that -- that information will be pulled. So Wikipedia has all of these name spaces and this is just to narrow the complexity of the implementation. We're just going to use four of those name spaces. But we pick four of the most important name spaces. Things that are article-article, top page user and user top page. We'll let you pick a time window. So you get to say do you want to show up for like all time? Do you want to show up just the last few years? And this is actually kind of useful because a lot of systems don't have a good -- especially systems that compile lots of data, they don't have a good idea of forgiveness. How do you go with an early person really kind of messed up but, you know, the last few years they've been really good at this. They've been a good citizen. So you might want to have a time frame. We'll show you some numbers of interactions then we have two different styles of visualizations. Sort of this kind of very simple egocentric graph and this sunburst kind of graph. Now, this was really the whole system was really oriented around the idea of looking at people. And if you know Wikipedia, one of the things that is interesting about Wikipedia is article pages generally don't have anchors of people on them. And this is where faculty should listen to doctoral students and I finally did. And that is the doctoral students would keep coming to us and would go, I don't know what would show anything, right. Because there's no anchors on this page. And so he said, let's just have a part of the toolbar, let's just list the people who most edited this page. And so, the Flooded [phonetic], the Goodvac [phonetic], right. These are the people who most edited this page. And you can see the number of edits for that time frame and that then forms a very simple anchor. So if we're going to look at the articles that this person edits -- I'm going to turn off these two. What we can potentially do is hover over Flooded and see -- that person in the middle of the graph and articles that they edit. And when you start looking at the articles that this person edits, what might you say about this person? [Silence] [Multiple Speakers] They kind of like disasters. I'm so glad. That's perfect. We're actually trying to test that. Because once you build a system, you kind of want to evaluate it. And so that I have then a population that can kind of see that. I'm hopeful then that maybe some of our evaluation might work out. So yeah, you can look at this -- and you can. He does a fair amount of disasters but then you might also ask for the question of like, so given that he works on those disaster pages, who does he collaborate with? Right. So now that you have seen the disaster pages, now what we have to do is see

who he collaborates with and this is the jeopardy. [Singing] Oh, okay. It came back. And so what you can see is the set of people that he most collaborates with the shorter distance here in this ego center graph represents closer tie. That's what we're doing here. And so you can kind of start looking at, well, this person WWGB. That's a person that he collaborates with a fair amount. Almathea [phonetic], Geramo [phonetic], T.B., so you can kind of just see who he collaborates with. And if you wanted to chase that down and you said, well. All right. You know, it's interesting that he collaborates with this guy. What I might want to see is whether it's kind of a reciprocal relationship. So I hover over this person and that person becomes the center of the graph. And I can then begin to see jeopardy. These are long running queries and so we're using a big fat server at the back to try to make this happen but it still takes a long time. Technically, this is where the graduate students should be listening to me. And I keep saying, how come we're not cashing the result? We do this over and over again. It should get faster and so what you can then see whether WWGB, whether they collaborate with Flooded on these pages? And well let's see, is he on the list? No, he's not. So one kind of edits more with the other but that person's editing activity is probably on other pages. Where Flooded doesn't engage so you could then also chase that down. So it's a very simple way of starting to understand who collaborates with whom and what kinds of things do they collaborate on and you can kind of look at those in that context. Okay. Let's go back. So there is another example. And that is the example that we originally started -- originally motivated our work and that is this idea of selecting Admins, new Admins for Wikipedia. So Wikipedia is an online community where the new administrators of the online community are picked from within the community. And they have this rather rigorous review process and we did a little bit of work on looking at how do Admins get selected. When people are reviewing Admins, what are some of the features? What are the things that the Admin review process looks at? One of the things that of course that popped up is who do they hang out with because people know that if they edit with these other people and those other people are trustworthy and I know those trustworthy people because I trust them. And then maybe this guy is trustworthy. And so they ask a lot of questions about that. And so we can kind of do the same thing with the RFA pages. And the RFA pages -- how many of you have actually seen an RFA page? No. All right. Let me at least just show you one of these so that you'll have an idea how one might use a tool like this on pages like that. And yet again, that's disappointing. [Silence] Does this one work? Yeah. Okay. So an RFA page basically -- basically there's a lot of header stuff and then the way an RFA starts is the name of the person and then the nomination and the nomination is signed so Bigumba [phonetic] is being nominated by Automatic Strikeout. And he says, why? And then there's a set of questions that the candidate is expected to reply to and I think the first five questions are fairly standard and then -- oh well. Three questions are standard. And then you have a set of additional questions that individuals can ask and many of these questions are specifically about the policies of Wikipedia. And they aren't in the space of tell me about policy x. They are if you see an x and then you see y and you see z, what would you do? And often x, y, and z represent a kind of confounding problem with regard to the policy. So I forget who said people have not read all of policy. I said, turns out there are a lot of Admins who have read an awful lot of the policy

and they really know it very well. And so this is their way of testing it. And then what happens is down after -- and look all of these additional questions. Then down below that are people who either support nominations or people who are opposed to the nomination. Turns out in this case, there's no opposes and one neutral. And so what you might do on a page like this you might want to inspect who are these people who are actually doing support? So I've seen this user on quite a bit. Calm, collected evenhanded, I believe they'd make an excellent Admin. All right. So who is this guy? C'mon. No.

[ Silence ]

C'mon. All right. I'm not sure why -- this is the oops, it broke. Yeah, who knows. It could be some crazy thing with what's going on in the page. Is this one even working?

>> You can't have and you can't have that because [Inaudible] I think that's worked before.

>> David MacDonald: Yeah. All right.

>> You told us to hope that it failed.

>> David MacDonald: [Laughing] Then it did. That's asking for the wrong question.

>> We were all really psyched.

>> David MacDonald: Well it half worked and then half broke. It's like what does that mean? I got both. So I'll let you know. This is -- it is actually proxy restricted. Because it is sort of an inhouse tool that we're working with. But I opened up on the wireless here, so if you want to just sort of play with it now. You could go and go to our proxy and play with this so I'll show you that at the end so you can

>> So then we can have your demo work ten times slower.

>> David MacDonald: Yeah. Well, but I'm going to stop demoing right now and finish the talk. [Laughing] We started being motivated by this set of work in management and organizational behavior. This goes way back to the 1990s. That was like last Century. And there was this idea that, you know, the new form of the organization was everyone was just going to kind of come to the organization and they were going to self manage. We were going to have these teams that they just form together and it's going to be magic. And so actually a lot of management with literature were sort of interested in this and so there were some good studies and some models. And we kind of lifted this model about these high performing teams from Yates and Highton to say, well, even if this model isn't exactly the right model, there are some factors in this model that we might need to account for if we were going to visualize what a group is or what a team is. All right. So this starts to form that empirical part and then what would you apply this empirical part to if you were a Wikipedia person? Any idea? Turns out that Wikipedia has teams? How many of you know Wikipedia teams? Yeah. Awesome. That's good. Yeah. Great.

Okay. So Wikipedia has these things called Wiki Projects. A Wiki project is a group of editors that want to work together as a team to improve Wikipedia. So it's got both. It's not just a group, it's a team. That's awesome. And that's their definition of Wiki Projects. And so we're going to leverage it. And there are hundreds of Wiki Projects. All right. There's hundreds of them. And actually, there are whole bunches of them that are active and successful and there are whole bunches of them that are not active and that somehow died off. And so you've got this wonderful ecosystem of teams. And another thing that's interesting [Silence] but also they can declare membership by behavior. So it turns out that Wiki Projects claim pages as their domain. So Wiki Projects sociology actually has a number of pages that they claim as part of sociology. Would that be surprising to you? No, no, no. They probably should. Right. And the issue is, if you edit that page and more of your editing is around pages that are claimed, you're sort of by your behavior declaring that you are more affiliated with Wikipedia sociology than perhaps some other project. So this is a great ecosystem for looking at the membership and co-membership and the distributed membership among these. Because, of course, you know that not everyone edits one thing if they're really an editor. They're going to edit a few things, you kind of saw that in those graphs. So you have this wonderful way of starting to explore this. So on the design side, we need to build a few things, so one of the things we're working on is actually making our reflex toolbar a pluggable tool. So that you could take different visualizations that you might want and plug them in. And that's kind of interesting because that means we could potentially collaborate with people who have cool plug-ins that they want to visualize along the top of an active Wikipedia page. So we're working on that. And we're working on simple tools that are Wikipedia specific quality of coding. So we have this little tool that we're working with this summer called Indicoder that allows you to annotate what the Wikipedia page is and this is kind of nice because you know that Wikipedia pages change. They have multiple versions and they change over time. So some behavior, something that you qualitatively code in one page, it may not be there in the next version or it may or may not have been there in prior versions. And you may want to track that. So special tools to think about how you qualitatively code these kinds of systems are an important piece of this so this is one of the things we're doing on the design side. And then I'll just kind of give you a quick view of what the reflex redesign is. It's going to be moving to the top of the page. We're going to have this kind of ribbony look. How many of you love ribbons? No? Okay. Good. Some people love ribbons. But you know -- well, we'll try that. And you can select different plug-in types. And as you add plug-in types the toolbar adjusts to show the different amount of these different tools. Settings that are a part of the tools that are hidden that you can show. When you turn a tool, then there's stuff -- there's the control part of the UI and then there's the visualization part of the UI. So you're going to try to keep those things separate so all the controls are in one place. And of course, if you go to a regular page you're still going to get the kinds of radio or sunburst graphs we have as well as top editors. So with that, I'm going to say thank you. Acknowledge support from the NSF and all the students who have worked with us. Michael is currently working with us. Stephanie moved to Cornell. Dylan is an undergraduate, Scott is an undergraduate, Casey is an undergraduate, Cheng is an undergraduate. Alena was an undergraduate,

went to our Master's program. And then Seway [phonetic] is an undergraduate. So a lot of undergraduates worked on our project. Wikimedia Foundation. If you'd like to try it now. It's open. I'll leave a proxy open for the week so you can like inspect your favorite Wikipedia page. And I'm open for questions.

[ Applause ]

[Silence] Yeah.

>> So you were -- so far it's you're testing a primarily in terms of

>> David MacDonald: Yeah.

>> Do you have enough just like people using it even internally to have some idea about, you know, what they're using codes? What visualizations they're using?

>> David MacDonald: So no. And one of the things that's interesting about this is we did go to Wikipedia foundation to show the tool. They were really interested. They really liked it but also you can kind from the tool, it is not really meant to be like an absolutely every day user's tool. You kind of have to understand a little bit about how Wikipedia, how it works. Right. It's more a tool for someone who is really into the community and understand. So they're more interested in sort of from the groups and foundation interested in the community management pieces. They're actually very interested in that. And the one that was funny was that I did a fuller demo of the RFA for them and it was an RFA that was ongoing at the time. And I said, well, you know this is very interesting. I looked at this user x and then I kind of chased it around and I said one of the things that you see here is if you look at this. There's this very interesting hole around this guy. You can come to understand this guy is like -- he is sort of like leading this opposed. And of course there's a lot of people here and they go yep, yeah he is. That's the guy. You totally saw that. It's right there. Yeah. So you start to see them even in the very simplistic views of the network, you can kind of chase it through and just kind of browse and go, oh wait there's something going on here. This opposed isn't completely right. A ground swell. It's more astro turf. Yeah.

>> I'm trying to think how to phrase this. One of the charges you hear about Wikipedia is that it's sort of an in-crowd that does all of the writing and thinking about Nash's [phonetic] talk and the idea that if you're connecting and might be different from you then you might have more [Inaudible] I'm wondering if the kind of approach that social translucence has anything to offer in terms of -- it sounds like you're working from within the community who is best equipped to rise up to higher slots. I'm wondering does it speak at all? Are there ways in which it can speak to how can people from outside the community be brought in?

>> David MacDonald: So, I mean, just off the top of my head, I don't know exactly how I would frame it as how I could address the issue of bringing people from outside in. I haven't really thought that much about that problem. I think it does allow people to inspect activity or make it more

accessible to inspect these kinds of activities. And that's important to the community because the community claims that when someone goes up for RFA anyone in the community can have an opinion and while it's not a completely like vote system, it really is the case that sort of majority kind of rules. And so, if you make it easier to participate, maybe that would allow people to participate more. Right. Because you don't have to do all of those individual edits. You can kind of get a picture so that might be a way of talking about it but I haven't really pushed that too hard.

>> Just to add to that, I like that comment because by putting the most frequent editors there, you're stimulating the centrality of the main community but if you put most recent editors or newcomers who started editing, you would encourage the arrival of newcomers because they would instantly get visibility.

>> David MacDonald: So one of the features in this new version is that top editors is going to become like a plug-in tool, and we have what are the hot pages? What are the new editors? So in essence, there will be multiple ways of having that kind of thing as like a tool that you could pick at the entry point.

>> But her point is to support that Jenny and I argue is that the design features encourage people to move from readers to contributors to collaborators to readers.

>> David MacDonald: Right.

>> Based on what you think of the design.

>> David MacDonald: Right. So in this case, what we would try and do is help, allow people to kind of pick which ways they would like to see it. Right. Give them the control to see it that way. That way then -- yeah, yeah, yeah. So. There was another -- yeah. In the back.

>> So I have a host of tedious technical questions that I would love to chat with you about over the coffee break, server, you name it.

>> David MacDonald: Oh God. Yeah.

>> But for this, I'm more interested in when you first introduced it, I thought you were going to go one-way which was I'm going to show you where on the page this person has edited. Then you went the other way instead, which is I want to show you where else they have edited. To me, part of this as social translucence is you're trying to lay bare the content, not lay bare the social structure of the content. So why the design choice to go to the social structure, rather than look at all the things that this person has done. This is actually produced by humans, not by you know [Inaudible] not by some sort of source out there that's completely inaccessible.

>> David MacDonald: So yes, we have both the relational aspects around this page and then we have the editing aspects of this person of which this page actually is possible that they're not. I mean in show up in the

top bar, obviously they wouldn't have actually have edited on this page, but if you chase things down its entirely possible they would have actually edited this page. I mean if you cascade following an individual. And so the issues I think we are showing many of the things that are happening on this page but around a particular person. And that actually was motivate by some of the early study we did around -- well, let me go back to this problem which is who do they hangout with? That was the main question posed to us by users when we did our RFA study and so we were actually trying to help the RFA Process. Because actually, The RFA Process is considered one of the most broken processes in Wikipedia. Despite the rapid scaling of participation in Wikipedia, the scale of Admins were just like -- it was just like flat. And there are people who have been D -- there are actually almost more D Admins. People who have been removed from Admin and have gained Admin in say like in the last five years. It's really close. So this is one of those processes in Wikipedia that Wikipedia is very interested in fixing and potentially having more people participate in at least doing the review. So potentially, more people saying you know this guy seems like he's done a reasonable job. I don't understand why we're not going to support him being an Admin. Maybe we can get that. So the main things about the things that we picked is the principle, the visualizations were to support this type of scenario. It turns out that it also happened to help support some of that styles type of scenario. And there are a few others. But in the redesign, we've actually been talking about well so what are some of the other viz that we might pick. And some of those viz are not anchored on a person's name but they would be anchored on a page. So you could imagine anchoring to the page who are the people who really edited this page in some way, right, or something about them. So that might be a larger network graph about the people for this page. But yeah, there might be also there's this idea of where has this person edited on this page. And we haven't done that one because there's a whole set of tools that had already been built in the area of validating content, trustable content in Wikipedia that actually does that. So that's actually already been done. So that's another reason why we didn't do that. But that doesn't mean it's not a good thing to show because those tools aren't actually available. I know where to get those tools but most Wikipedians can't get those tools. So there's this -- yeah. Really interesting piece there. Yeah.

>> Just a question I'm actually asking Nash and you at the same time because you both talked about visual, making things visible, and making things transparent, and making things available to people. And partially Wikipedia works because people can't see who is doing what's on a given page. And some people are willing to try this. That's a partial argument and that's how it works and partially scientific stuff works is because you can't find the most prominent person and have everybody in the whole world ask them questions because they would get overwhelmed. And so it's not inherently the case that visual, translucent, and visibility is always good. Do you have any feeling for where the traps are with visibility and translucence that we need to watch out for?

>> David MacDonald: Yeah. So I'll comment first and then I'll give Nash a chance. So first, I talk -- I know Tom Erickson really well and he and I have gone a few rounds on this. And I'll support him that you should not

confuse translucence with transparency. These two things really shouldn't be confused. Social translucence is really different. Transparency, the idea that you could really see everything that's going on completely, somehow that's kind of -- that's kind of creepy, right. Potentially. And so I think there's a -- if translucence is done well, it's not truly completely this idea of visible but interpretable, right. That a person can sit there and then when you think about that and move translucence kind of further down that logical conclusion the one thing that I would say is there's kind of like this weird problem of the way translucence sort of plays out. And I've taken heat on this from Tom. So he really thinks that everyone should see exactly the same thing. And that's his view of translucence and what I would say right now and we could test this right now. Are you seeing the same thing I am even though we are in the same social setting? No. So there's this weird tension around social translucence that suggests if you're doing translucence well, maybe we don't have to see exactly the same thing. We might come to some similar interpretations of what's going on here but we don't necessarily see exactly the same thing. And that carries through. I mean there's lots of systems that have regular user roles and Admin roles and what an Admin sees is clearly different than what users see. I don't think we would ever break that down so there's something going on in the space. And that's my long winded claim of same. There's something about translucence that current representation of the theory and the current way we think about it as a design idea sort of doesn't quite account for and so it needs some work to be done there.

>> Okay. So I think that was a fair enough question. I'm still [Inaudible] translucence not the level that you are. But I would argue that starting from the premise that we're trying something new. If you rewind a little bit, you know, things were more translucent or transparency, when we didn't have a [Inaudible] and everyone was face to face. People knew who was friends with whom. David Krackhardt, for example, has this very interesting concept in networks called Cognitive Social Structures. It's who knows who knows who. Right. So it's our mental map of who's connected to whom. So there's an actual map of who is connected and there is a section of who knows who. And when you're in face to face settings, you have a pretty good idea of who's connected to who for the most part especially in the work settings because you see who hangs out with whom in lunch, who's at the water cooler with whom, when you walk by people's offices or cubicles, we see who is hanging out with whom, et cetera. When you went virtual, we lost that. So part of what this is doing is using the same technology that shrouded those kinds of [Inaudible] I'm trying to now unshroud it if that's a word and we're trying again to bring that back. So it's not like we are experimenting for the first time. Yes, in that stage before we had virtual technologies, transparency was good and bad. And you know, that's why we copied social structures is one of the ways in which I talk about it is that one of the ways in which we decipher if a link exists is by asking the two people. But then there is the Brad -- what is it? The Brad Pitt, Angelina story where if you asked the two people if they were having an affair they would say no at one time. Right. And everyone else would say yes. So what makes it yes is not that anyone else -- if the two of them said no, [Inaudible] social structure. So of course, it could be bad for them at the time. Right. They didn't like all this transparency, it was

driving them crazy. So that can be true in science as well. It is true in science. But I think the problem is not a new problem. That problem was there face to face and now we may have cursed the problem as we increase translucence slash transparency or transparency in the world.

>> Obviously one of the main contributions here is to raise visibility and to show things to human who can make decisions. Have you done any work to kind of take it to where the computer, you know, might do some sort of model that could help make decisions, maybe even to for instance, in this case nominate someone who might be a good Admin based on their similarity or any other patterns?

>> David MacDonald: So that's actually almost explicitly what this work is modeling Wikipedia with decision making. So they actually have the -- what I would say is -- and this. So first off, I like Mora. She does great work. Bob Kraut, complete admiration. They are quintessential statisticians. They know how to do that well. And so this paper does a really good job outlining all of these really interesting qualitative traits that are really important to do Admin decision making well. Oh, and then they go well we throw all of these things out because these things are accountable. So there's a great statistical model. It is really good. I mean and it's accurate. I mean I tell students, I had another student that was working on something similar. I said you should just adopt that model. It's a reasonable model to start with. But recognize that there's this whole space here of things that they acknowledge but they can't count so that's one thing. There's a weird thing about thinking about the statistics. And so in terms of helping people, there's some other work that I've done on Wikipedia vandalism. And one of the things that's interesting about vandalism is it's actually -- it takes a little work to detect vandalism well when it's subtle. And people are actually bad at this. And something that people are actually bad at when its subtle machines are going to be pretty horrible at. And so the student that I was working with, the argument I was making with her and that I was trying to get her to kind of push more forcefully was that automatic vandal detection, that modeling, that's going to cover a really interesting say like 90 percent of stuff that you just you get it right. And then there's this really interesting 10 percent of stuff that if you know that the model has a horrible time with it, then that tells people where a person's attention may be the most valuable. And so I think there's really some interesting spaces here where maybe we can do some really cool things but then if we could recognize where trouble, we could actually leverage the power of the crowd even more. Right. Because attention is a limited resource. And so we should value it. It's the systems that aren't valuing the persons' attention that may be ones we should be trying to squelch. These ones that are valuing the attention, that are saying that this is a difficult problem and I really need help with it. Right. Because the machine can't do it. That's a huge value potential.

>> Jenny: I wanted to open up the translucence and transparency discussion a little bit more. Because my understanding, or certainly my interpretation of the Erickson and Kellogg work was that they showed us a certain participation in a chat system and then some of that -- I think the translucence actually was the same as transparency. And one of the

things I wanted to mention particularly was that people who were lurking on that system or people who were I prefer the term not participating were shown up and at that time it was thought to be sort of a pejorative activity. So I wanted to sort of point that out and open that out. And then following on from that I wondered if how much controversy there is about the translucence that your systems are operating. I'd have to say that unlike Ben, if I was a new person coming in, I don't want to be shown up. I want to get to use the system and build up my confidence first.

>> David MacDonald: Yeah.

>> Jenny: So you know in a way, it's revealing some things that people might prefer not to be revealed for good reasons.

>> David MacDonald: Yeah. Yeah. So just to kind of sort of reiterate the issue that Jenny was mentioning is that the original visualization is basically like this one in the middle here I believe that they called the cookie. And it represented a chat room. And when people were contributing to the chat they were colored in and they would come to the middle. They would like -- that they were kind of like joining up to have this kind of close intimate conversation so they would come to the middle. And when they were not typing into the chat, when they were not sort of participating in the chat they'd actually hangout on the edges. And so lurking was someone who was signed in and just sitting on the edge. Right. So that's how you would interpret lurking. So they would claim that -- so again, I'll channel Tom. They would claim, this isn't transparency. It's not transparency because I'm not telling you actually in the context of this what they're actually saying and who they're actually speaking to. Right. So there's a certain thing here that is not quite transparent, that's not completely exposed.

>> Jenny: So here's the thing, they are not speaking. That's it.

>> David MacDonald: Well, right. So I'm just kind of -- I can't completely support everything they say. C'mon. They are unique. I am not them. Right. And so I'm just kind of saying that's sort of the line of argumentation they make about why it's actually translucence and not transparency, but great. So the translucence is going to reveal some things that in the past you would have needed to kind of scope out a little bit more. So in IRC, if you watch enough of IRC, you can figure out who has not said anything in the last few hours. Right. It's very hard to be completely hard on IRC because you can see the names. And so it's not that it wasn't available. Right. You could see who is lurking but it takes more work. Right. Because the interpretation and the work to do the interpretation, the system's not helping you do any of that interpretation. Right. And so their system is making a little more of that interpretation easier to make. Okay. So that's the first thing about the nature of translucence versus transparency. And you can tell, there's certainly kind of a line here, potentially thin line, right. Okay. So then about the kinds of systems that we have. So, yeah. You know, it's kind of interesting. When we did the talk at Wikipedia Foundation, there was a few people out there that were like you should never release this. Wikipedians are going to hate this. Because there are some that don't

want their behaviors inspected so easily. And then, of course, the other half of the room pops up and goes, yeah. But they're just trouble makers. So even inside. Right. You could tell that inside there's these kinds of debates. You know, is lurking good or bad? Are these people who would be so -- feel somewhat violated, because remember the edit history is all there to inspect. It's just that the tool makes it easier to inspect, right. So it's not that you can't see these people's behavior and the really good Wikipedians know how to find that persons behavior. It's that you're making it available to sort of the broad unwashed masses that now can see that maybe I'm not completely the honorable Admin. Maybe I'm an Admin that takes a few too liberties or whatever right. Yeah. Those things become more visible by lots of people. And that's not, that translucence or heading toward transparency, not every system wants all of it. Right. So I think there is something there. Yeah. I mean -- I don't have a way of gauging that's sort of a traditional resistance to new features or new systems or not. I don't have a way of completely thinking about that. But when you think about systems that compile massive amounts of behavioral data, it's going to be become easier and easier to do this. So I think in the space of the complex systems that we're going to be building, this -- whether it's just resistance to a new feature or whether it's the new domain are the kinds of things we're going to allow people to see an online. I think there's something going to be going on there. And it's good for us to stay on top of it thinking about it rather than let Google do it or let Facebook do it or right. Yeah. Where we don't have an idea of what's really going on. There's actually -- okay. Yeah.

>> There seems to be kind of a conflict between what you just said and the point you made earlier where the people who are kind of generalists Wikipedia aren't the ones who are necessarily interested in the information you provided. So how do you necessarily gage between the work that you are doing as valuable per se, and

>> David MacDonald: Right.

>> Kind of this attempt to inform the community itself and gage what they're into?

>> David MacDonald: Right. So it's not completely inconsistent because what was designed for doesn't necessarily align with like who might actually take it and why it might be taken up. Okay. So those aren't inconsistent. It was designed for the RFA process. And the way in which people engage the RFA process. And by and large, the people who participate in the RFA chat discussions, they're generally very experienced Wikipedians. I don't have to tell them about name spaces. All right. They know it. Matter of fact, they would complain that I would call it context. Right. They would say no, those are name spaces. You should use name spaces. Right. That kind of experienced user. But these tools because they make potentially and that's the issue. Potentially, they make something that's very difficult to do. As an experienced Wikipedian, you kind of need to know how to churn through all of these edit histories of all of these people. They make something that's difficult to do that maybe now only experts are doing. Because that's what we're seeing experts are doing it. They make it potentially a person

that isn't quite an expert that makes it possible for them to do it. Now do they or will they I don't have the data for that yet. I just don't. All right. And so, I don't know if that will happen. But that is what I'm -- when I'm talking about that I'm talking about the potential. Because it does open up something that's difficult and make that difficult thing a little easier. So potentially, a less experienced person could pick it up. But it wasn't designed for them. You're absolutely right. Yeah.

>> Jenny: A really wonderful discussion. Thank you very much. [Applause]