Wikipedia in the Noosphere – Part Two

David Sloan Wilson: Well, let's get to governance here Anne. Two great challenges as I'm sure you know, one is a challenge of coordination, even if everyone is willing. And then there's the challenge of conflicts of interest where people actually are not on the same page.

And so how does Wikipedia deal with these two great challenges of coordination and conflicts of interest and where does your checklist come in? Because you are known for Risker's checklist. And I also wanted to say, is there significance to usernames the fact that people have a username, which is different than their real name? Is that an important feature?

Risker/Anne Clin: People choose their usernames for various reasons. At the time that I was joining Wikipedia, registering my account in 2005, there had been some very negative encounters. Some women editors had been quite seriously harassed by a few people and 2005 was a long time ago and I really had have always and very cautious about what I put out there on the Internet about myself. And I felt it was a little bit risky to join Wikipedia. So I called myself Risker. The interesting side effect is that, many people, even today, still feel that my username sounds very masculine and are often surprised to find out that, when they're on a video call, they're talking to somebody small. So that's very interesting to me, but on the whole, I would say that it's not really... People choose their usernames for reasons that makes sense to them. We do have a fair number of people who edit under their real names and many people who choose a username, but also linked to their real names. So it's very much a user choice thing.

2:20

DSW: So what you've revealed here, Anne is that Wikipedia is not immune from all of the pathologies that we associate with the Internet, such as trolling and harassment and predatory activities of various sorts. Not to speak of subverting information. I mean, everything we've talked about with fake news is something that's... I mean, some people must be dying to get their spin on things into Wikipedia for that to become the reality of seeing through the lens of Wikipedia. So we're back to coordination and conflict. And so back to you on how Wikipedia can accomplish these governance issues.

R/AC: We do a lot of things. And we have to keep in mind, this is a 20 year old website. So we've learned a lot of things over time. And the lessons that we've learned on English Wikipedia are shared with our colleagues on Hungarian Wikipedia and Igbo Wikipedia and Bengali Wikisource and all of our other projects. So they have access to that information as well. And the software tools that we've developed. We're working on making sure that they're able to be flexible enough to help people. So there's a sharing of knowledge and experience that is useful, not just within our own project, but to our many other projects as well. Some of it can be transported and some of it can't. Some of it becomes part of the hardcore software that runs the entire Wikimedia movement. And some of it is just very localized because it's a local problem that they're trying to address. So we have a lot of mitigation tools that we have developed over the years and this is where the coordination comes in. It is all self coordinated.

People will say, "Hey, I can do this," or look for other people who are willing to work with them on something. So that's how our featured articles started developing was that people who were interested in having the highest quality of work on our main page got together. And they just said, "Hey, let's establish some standards. Let's help each other build those articles." And that's how that particular little area developed. For a long time, and it's not as prevalent now because we have created, again, mainly by users, we have created software that we call bots, and I'm sure you've seen bots all over the Internet doing various things. Ours will quite often work on anti-vandalism. They'll automatically revert something that they've been programmed to take that edit out, to back it up. They don't formally delete anything. They just revert it to the previous version of the article.

5:34

DSW: Yeah, that's part of this is, is this fossil record is basically nothing gets deleted. You have an amazing, amazing fossil record of everything that ever gets done on Wikipedia, right?

R/AC: Yes. Wikipedia is quite a cache. And there are a few exceptions to those rules, but generally speaking, yes. You can go back to an article and read every single version of it all the way to the beginning, in most cases.

DSW: Just to clarify, you have bots that recognize vandalism, that there's some algorithm that recognizes, and vandalism is somebody that just comes in and messes up a piece just for the fun of it or what?

R/AC: Oh, it can be all kinds of things. It could be a test edit. You know, somebody who tries to edit something realizes they don't know what they did and saves instead of backing out.

DSW:Yeah. Just junk. Degradation. Yeah. I mean, it's not mal-intent. It's just, it's what happens. I mean it's entropy basically.

6:40

R/AC: Well, that's it exactly. And we have bots that will help with a lot of that. We also have something called a special screen that we call recent changes, which literally lists every single edit as it's happening. And people watch that recent changes list and will assess whether or not they need to look further at a specific edit to decide whether or not it's appropriate. And many of the editors who are working on Wikipedia today started off doing recent changes patrol. It's a very common entry point for our editors. People just start doing these things. It's amazing the things that people will just sort of fall into, or they'll notice that somebody else is doing it. And they'll say, "Hey, can I join you?" Or they'll just start doing it too. And it just sort of becomes.

7:40

DSW: How much apprenticeship is there? You say, "Hey, can I join you?" So if you wanted to get into this, you could do it solo, but you just said, you could actually attach yourself to somebody and be an apprentice. "Hey, can I join you?" Could you elaborate a little bit on that?

R/AC: We do have some users who will act as mentors to new editors in various ways, whether actually doing some editing training with people. We have courses and what we call edit-a-thons or project similar to that, where we help people to learn some of the background, some of the little tweaks and things that they can do to be able to integrate themselves into the project. Something as simple as how do you format a heading, or how do you save your changes? What should you say in an edit summary, those sorts of things. So it can be very basic or it can be very complex.

For example, when we are dealing with some of our more high level user rights, we'll want to actually provide a structured training course and a person practically to hold the hand of a new person in that area to walk them through the entire process, because it may be a complex process or a lot of steps that happen. And a lot of it you just sort of observe and you watch things happening as well. This is a culture where trying something is accepted. If you try something, it can always be reverted back.

DSW: Yeah. I mean, that's one of the great things, it's safe to try. There's a phrase, safe to try, and then you could always go back and you could do something else. So it's inherently evolutionary, in my way of thinking. Variation and selection, variation and selection, in a way that's safe, because you could always go back, which is not always true in the real world, but in this case it is. How soon when a person needs help, do they connect to get connected to a real person? As opposed to something like a frequently asked question, disembodied frequently asked questions list?

R/AC: They can actually put a help notice on their user talk page and someone will come and help them. It may take a little while. I mean there aren't always people who specialize in helping people who are online or are available at that precise moment, but they're going to get an answer sooner or later, usually within 24 hours. The other thing is that Wikipedia is based on eventualism. Not everything has to be done right this minute. There are exceptions to that quite often, when we are dealing with a breaking news story that is getting a lot of observation and review, and everybody is coming to Wikipedia to see what's going on.

Those articles are sort of, or areas are dealt with a little differently, and they will be very, very highly watched and very carefully monitored and edited. But for, the article on, I don't know, Madonna, there are always going to be some people who are watching it as somebody asks a question on the talk page, there they'll get an answer. It may take a couple of days, but there are probably, I don't know, 1,000 editors who have that on their watch list and will notice that a question was asked. So it takes a little while.

11:41

DSW: So what happens if somebody's identified as a kind of a bad actor, that's not really operating in the interest of the community. What happens to them? What's the response?

R/AC: Over time, we have developed a lot of mitigation practices and they sort of fall into two categories. One is to manage a problem with an article. And the other one is to manage a problem with an editor. And sometimes they interchange, they cross over, but we have dispute resolution systems that are involved. There are some what I would call bright lines that if people cross, they're not going to be allowed to continue. We will shut down their account. We will block their accounts. Really harassing content, adding a lot of harmful information, undisclosed paid editing will usually result in having an account blocked, repeatedly damaging articles will get an account blocked. We do those sorts of things, very routinely.

DSW: How common is that? Does that happen all the time, or just only infrequently?

R/AC: All the time, all the time. During the course of our interview today, which would be what an hour and a half, we will probably block 50 editors or accounts. I won't be doing it personally, but it will happen. We also block the IP addresses of VPNs and similar processes, similar Internet access processes as open proxies. And that is because, so that we have a better control on being able to link users to each other. And to be able to block those users, we have found over the course of many, many years, that most of our problem editors come from those sorts of editing or online processes. And we really have found that it works to our advantage to just say, no, sorry, no VPNs. And then we wind up having to exempt people from those because we have certain countries where the way to get to Wikipedia is through a VPN. It is the pretty much the only way. So we have a balancing act to do there too.

Bad players exist. We know that they exist. On English Wikipedia, we're not very tolerant of bad players. Other projects have taken different approaches or have just done other things that prevent bad players from getting to them. For example, some projects only allow people with registered accounts to edit. It's something of an experiment, but it's very important to them. Or they have something called flag revisions process where anybody can edit, but the publicly viewing version of the article will remain with the last edit of a registered user until somebody reviews what the unregistered user has done first.

15:35

DSW: Yeah. Yeah. So this is so very interesting to me. Do you think of it like an immune system? I mean, you and I have immune systems do to help keep diseases away. And does anyone actually use that metaphor that Wikipedia needs an immune system for bad actors at all?

R/AC: Oh yes definitely. Definitely. We've used that metaphor. That very specific metaphor. We will never be completely immune to it.

DSW: No immune system is. No immune system is. Always vigilant, often challenged, sometimes overcome. That's the way our immune systems are.

R/AC: One of the major reasons for our success is keeping ourselves open, making ourselves available, allowing ourselves to be edited, our projects to be edited by people whose motivations are completely unknown to us. And that's how we exist. We need to have that input coming in all the time or else we just become, some other site on the Internet that falls out of its currency and its usefulness.

DSW: Absolutely. I mean, for an organism to exist at all with all of its amazing physiological processes is just amazing and it's always being subject to disruptive forces. Forces of entropy forces of disruption, conflicts of interest and so on. And there's a selection process that's required in order to maintain order. And in the absence of that process, then disorder results. And so, it's so interesting. I could talk with you all day, Anne, but I think that I want to make sure we cover two things.

17:44

One is your checklist because I think that's part of the immune system is your checklist. So I'd like to know, first of all, the story of how you proposed it and how it became widely used and what it is. And then I want to run through Elinor Ostrom's core design principles and see how well they kind of fit what's evolved over there at Wikipedia. So, first your checklist, what is the story of your checklist?

R/AC: Risker's checklist is a checklist for software developers. It's not for editors. It was developed because I had good connections with the software developers from the Wikimedia Foundation. I count many of them amongst my friends, and they kept creating all kinds of neat bits of software, what we call extensions and then they would ask me to test it because I was somebody who was willing to test new software. And I had all of these advanced permissions, so I knew what I was looking for because one of my positions was that I have to be able to use my advanced permissions for everything that's publicly viewable. And I kept running into the same problems over and over and over again. And I just got frustrated one day and just wrote the checklist and then linked over to a couple of developers and said, is this of any use to you guys?

And they said, "Oh my God, somebody finally wrote this." They were very thrilled to have it, and it was written from the point of view of somebody who was experienced and running into new software that was applied to Wikipedia, or one of the other projects that we couldn't take care of from the perspective of the community. We were not able to moderate it properly and that's sort of a core principle, is that the community has to be able to moderate it because the foundation has 500 staff, and they shouldn't be moderating content. And starting to tell people, these are the kinds of things that you could wind up with if you don't give us the moderation tools built right into the software. And so I wrote that checklist and then, passed it on to a few people, asked their opinion and the head of engineering at the time was somebody I was working with on something else. And I said, "Oh, take a look at this. Is this is this useful." And he said, "I'm going to print this out and post it on everybody's desk, so that they know when they're developing something."

DSW: What are some of the items on the checklist?

R/AC: The checklist says that everything that creates a publicly visible version of Wikipedia, if it can be seen by the public, then the editing community has to be able to moderate it. We have to be able to delete it. We have to be able to edit it. We have to be able to find other things to do with it. We should be able to revert it. We have to make sure that we can do something called suppression, which is removing even from the view of administrators, certain content. And we do that with content such as,

very personal medical information or telephone numbers, things like that, very personal information normally.

DSW: So privacy, I mean, there's so much, we haven't even gone there, Anne, the use of personal information and the protection of it. So actually maybe we should, but please continue with the checklist.

R/AC: So we talked about that and then we say, "We have to be able to track those edits as well." So we need edits to show up in recent changes and we need them to show up in various tables that we would look at and things like that. And that has actually become a major design principle. And even when they're writing in software, that isn't going to directly go onto Wikipedia or isn't the base software for Wikipedia, they do their very best to match up as closely as possible. I've been working with the group who is developing a project where we can put all our bots, sort of host all of our bots, from all of the different projects and people can see what else each other has. And it's written on slightly different software, but they've used exactly the same principles as mentioned in the checklist. And we often find that if people are not following the checklist or are not using it for interpretation, that problems ensue.

DSW: You have a do no harm clause, talk about that.

23:06

R/AC: This is really important, and that in particular came from a very well intentioned, but extremely controversial software decision that was made, I don't know, around 2012 or so. A new editing software was being developed, which we call visual editor. And it was very early in its development. And a decision was made for whatever reasons to make it the primary editing focus for English Wikipedia and many other projects. And the problem was that it didn't work very well, because it was very new software. And more importantly, about 30% of the edits that were made with this software were actually inserting information that was harmful, damaging somehow to the visual appeal of the article. It was taking things out that was supposed to be left in. If you tried to enter, you would get all kinds of strange characters going on.

And it took a very long time and an awful lot of arguments with the Wikimedia Foundation who had put this software in to withdraw it. And we were actually at the point where we were about to do something that could potentially have been very harmful to the project just to protect the project from this software. And they finally, agreed that, "Geez, maybe this is not working as well as we thought." We actually call this the snowman example, because one of the things that this software did was it would put little tiny snowman characters all over the place, in the middle of a sentence in the middle of a word. And this wasn't even a character that we had on any of our character sets. So we had no idea where this was coming from.

DSW: But it wasn't and malicious, right?

R/AC: No.

DSW: Not malicious in this case.

R/AC: That's it exactly. It was not malicious. They were trying to improve things very quickly but the fact of the matter was it was just too much for this project. For a project that gets, thousands of edits a minute, or tens of thousands in an hour. We were busy fixing the mistakes constantly just of that. We couldn't build the encyclopedia because we were too busy cleaning up after these projects. And it was really, that was the doing no harm. That was definitely harming the project.

And we learned, both from the foundation end and the software developer end and from the user end that we had to be a lot more accepting of, some changes, but at the same time saying, "You know what, it's okay if it doesn't work." And developers quite often have this idea of move fast and break things, or

we'll just upload it and then, we can fix it from there. And the fact of the matter is this is a live site. It's being used by millions of people a day. It's an international resource. We have to try our best, not to mess it up too badly. And we have other ways of doing a lot of those things. So backing it out and then working again. It's resulted in things that what we call our development train, where they try a new version on a couple of projects that are willing to go through and do a lot of testing.

DSW: Yeah, there's got to be something like A/B testing or whatever. I mean, you must have refined that to a fine art. I would hope.

R/AC: Yes. And certainly this particular problem with visual editor, which is now, and I will tell you quite honestly is an excellent editing tool, several years later and improvements later. And I use it all the time, was something that was unusable when it was first there, but it's a whole different ball game now. It's a very different project and a very different editing tool.

DSW: So what does the do no harm clause item of your checklist cause someone to do. Basically to try to anticipate unforeseen consequences or something, and just basically think about the systemic impact of what they're doing, and then to have that view in mind, is that what happens with the do no harm clause?

R/AC: That's part of it, so taking care of things in advance, the slow progression of application of revised software or updated software. So they start on a couple of wikis where people... those particular projects have said, "Yeah, we want to be testers." And they have people who actually spends a lot of time testing out all of the new changes to make sure that it'll work properly. And they also do something called revert... they will revert their changes as well. And it depends on the nature of the harm that's being done. So if it's something that makes a site unreadable, obviously they to revert a lot faster because they need to have the site readable. That's a primary organizational goal. If it does something that really harms something or makes something very difficult to use, un-editable for example, they will revert their software change just as we would revert a bad edit or a problem edit.

And they've gotten a lot more willing to do that. And they've gotten a lot better at doing things step by step, so that before it gets to English Wikipedia, they've fixed a lot of the problems before it gets to us. We're usually the last site that they upload software to, out of the hundreds that we have, 800 or so that we have. And that's simply because it's the most heavily used, and if something goes wrong, you want to take care of everything that you possibly can before you hit your number one site, the one that's getting all the hits and is bringing in the readers from around the world. But as I say, we're one of 800 sites, so they've got a lot of sites where people are testing things and finding things out and letting software developers know if there's a problem.

30:21

DSW: I asked if you use the immune system metaphor, and you said that you do very much, are there any other organismic metaphors that you use to think about Wikipedia as some kind of superorganism with an anatomy and a physiology and a nervous system or as there more use of the organism metaphor in addition to the immune system?

R/AC: I think probably one of the metaphors that we use, has to do with the levels of development. So as I say, English Wikipedia and several of our fellow projects, Spanish, Italian, German, French that have been around for a long time and have large editor bases, and a lot of articles and experience are the mature projects. We continue to grow and develop, but they're fairly mature projects. And then there are the middle sized ones that are still developing and still learning how to operate, and quite often we try where we can, and through different channels to help those projects to grow and to support themselves to become the next level of mature project.

DSW: Okay. So it's a parental care sort of a thing?

R/AC: Yeah. Sort of. The very young and small organizations and projects get a lot of support in helping them to develop. We actually have an incubator project, which is used for very, very young projects where we only have three or four users in a particular language, particular type of project so that they can get to a certain point before they get on the list of the official Wikipedias or the official Wiki sources or whatever. So they have enough content to actually be usable. So we do have that hierarchy of projects.

And the key is trying to learn from each other and to learn from where others have been before. And it applies also to some of our social groups, our chapters, and our other affiliates, our youth user groups, where we have levels that they are expected to meet, to come into certain opportunities, shall we say. An annual grant only goes to a chapter that has formal recognition and has certain financial stability and so on, and special grants will go to user groups who are going to hold special projects.

33:23

DSW: So financial resources. So much is based on volunteer. But now I've just heard that if you're doing a project, there might actually be some funding available, even if it's modest funding. And so tell us a little bit more about the financial resourcing of some of these projects.

R/AC: We have to keep in mind that... As I said, it is volunteer based entirely. There are about, around the world, about a thousand people who work for Wikimedia Foundation or one of our affiliated chapters or similar organizations who are being paid a salary. I mean, if this is their job, they should be paid a proper salary. And that's fair. I think that's something that everybody would agree to. If you were being hired to do this for eight hours a day or whatever, you should get a proper salary.

They probably don't get the biggest salaries in the world, but they do get salaries. And whatever is required in their country for additional support, depending on where they are. So for the groups that actually hire people, they have to have some kind of funding. Some of it is raised locally, and some of it comes from the general pool of funding that is raised on the Wikipedia and Wikimedia sites. We are hitting December 1st. We're about to see on English Wikipedia, our biggest fundraising kick of the year. So you will see those, if you don't have a registered account, you're going to see that the next time or in the next few days, when you log in. And that's one of our biggest fundraising processes.

DSW: And that gets distributed then in this fashion. So then throughout the system, basically, so you could apply, there's some application process or grant application process and things like that?

R/AC: Yes. And I mean, the applications need to be related to Wikimedia obviously. Some aspect of our motivation, our mission. So for example, there'll be money for prizes. For our Wiki Loves Monuments competition. There will be prizes for photos of the year competitions, Wiki Loves Africa.

36:08

DSW: Now, a little more. Do keep going. I can't help myself. So the competitions you have, competitions with prizes. Talk about that a little bit.

R/AC: A lot of these are based on improving either the content of a project. Quite often, these competitions will be to create or improve an article on various projects or in certain languages. It may be illustrating articles about a certain topic. It could be taking photos and who has the nicest photos. It could be focusing on a particular type of content that we want to improve. So for example, we may very well have a competition on writing articles about women scientists. We could have a competition on improving articles about Ghana.

DSW: So let's take the first of those. Writing articles on women scientists. So there's a competition, and first of all, what are the prizes and what's the response? I mean, you get dozens of articles? Hundreds of

articles? Is there a single winner? Is the prize distributed? How many articles does that result in? These are all questions that rush to my mind.

R/AC: Well, a lot of it depends on how the individual competition would be built. So right now we have a competition going on English Wikipedia for improving the opening paragraphs of articles. Because when you look on Google quite often, you'll see the opening paragraph of a Wikipedia article. So we want to improve those, make sure that they're good and they're consistent. And what will happen is that the competition is how many did you do and you're awarded certain number of points. How many did you do? How many bytes of information did you add? Did you add a photo to the information box? All of these. So there'll be various points assigned to various tasks. And then the competition runs for a certain period. And then at the end of it somebody will win. And the prizes are usually relatively small. They could be money towards buying a resource, a book that you want or something like that. Or they could be a t-shirts, or stickers or...

DSW: Small change, small change.

R/AC: They're usually small time. I mean, it's unlikely that... I don't think we've ever seen a price that was worth more than a $100 US.

DSW: You're getting a lot out of that.