

## Interview 18

Interviewee	20-User-E
Interviewer	Ashraf Shaharudin (TU Delft)
Date	6 June 2023

### Interviewer

My first question to you, could you please describe your role in <redacted>?

### Interviewee

I'm <redacted> now and as part of the program in [country E], one has to <redacted>. So those are my roles.

### Interviewee

Do you want me to go deep into how I use open data?

### Interviewer

Yes, yes, that would be good, yes.

### Interviewee

So in teaching, I do actually have a lecture on visualization of data, and we often times use statistics on that and the data taken from <redacted>, which is the open data service. In research, we basically [use] open data, we rarely buy data from providers. That's what I can say.

### Interviewer

But these are mostly not geospatial data, right?

### Interviewee

Yes, I work with geospatial data all the time.

### Interviewer

OK, so the course that you're <redacted> also use geospatial data?

### Interviewee

Yeah, yeah, part of it. I do have to say at the beginning of this interview, maybe, that the idea of the research and the teaching is to go open data not only on the data that we use, but also open software. Why? Because, for example, in our projects, they are mostly international, we cannot rely on Esri, for example, because of the licensing. So we are moving towards either QGIS or <inaudible> that are doing most of the work at the moment with our models.

### Interviewer

OK. And how long have you been working in this or similar room?

### Interviewee

For about five years now.

### Interviewer

OK, now I'm gonna talk about the topic of open data. What do you think is the value of open data to society?

**Interviewee**

Well, as I said, as we move forward to open software, I think open data is open for transparency and most of quantification is enabled through open data only. And we have various experiences with closed data in [country E] and also international waters. So what we can see is that our models, what we develop, they're mostly enabled if we have the data, if there's no data, then there's no model, there's no prospect that we can build on and there's no research we can do. So, I will say it is transparency, it is quantification and development for sure.

**Interviewer**

And do you have any challenges or grievances or unmet needs in working with open data?

**Interviewee**

Yes. Most of our models, they compromise on detail depending on the granularity that we find in the data. So, some of our models will be very local like neighborhood based or city based, and some of them will have to be regional models because we cannot find local based data, if that makes sense.

**Interviewer**

Yes. So the granularity is an issue?

**Interviewee**

Yes. It is an issue for sure. Accuracy, it is as well and we, at least in the group, we tend to think that some data is better than no data, and some accuracy is better than no accuracy at all. We use a lot of OpenStreetMap because we need to identify building footprints, for example, and we know that that data has some issues but overall, if we take those building polygons, then we aggregate them in 1 kilometer, it's great, for example. And we match that with other type of data like for example population data and then we kind of cross validate and we make our model stronger.

**Interviewer**

OK, so apart from granularity, accuracy, are there any challenges or unmet needs?

**Interviewee**

This is also coming from the students, the services that we have now in [country E], for example, it's not easiest for the user, so it's not user friendly enough. It is not an interface where you go and click download and it's downloaded. So you need a user – there's a learning curve on how to use the service, for example, and that is oftentimes something that deflect users from using it.

**Interviewer**

OK, now we're gonna move to the topic of open data ecosystem, which is defined as a network of interdependent yet self-interested open data actors. What is your perception of the health or sustainability of the current open data ecosystem? In other words, how do you perceive the future of open data in terms of its potential access, use, financial sustainability, support from various actors, etcetera?

**Interviewee**

I believe that both the supply and demand of open data are becoming increasingly important. I do believe that we have even more data now that we had before and the providers are actually making an effort, at least in [country E]. But as soon as you go out of [country E], then you find out that [country E] is the Disney World for data. And that happens to us many times.

We had a project in Mongolia, for example, where we had absolutely no data, even though we had many partners in the country trying to access and fetch the data for us. So I do think that [country E] can be an example of how they share through a structure and legislation, for sure. I also have another example for the satellite data, it's becoming more accessible, more easy to digest and with machine learning and AI tools that can basically -- how can I say this -- communicate better that satellite data, that can provide us a potential that I cannot even describe.

At the business case, I think it will always be the case for a user friendly and make it easier for the user not to have like coding skills, and fetch data easier. So I think there will always be a business case for this. But as a researcher, I don't need that, so it's just a direct line to the accessibility of that data, for example.

**Interviewer**

How do you think open data ecosystem can be improved further?

**Interviewee**

Education first and foremost because most of the people, and I am talking about students as well, they do not really know how to and they don't know how they [data] are called, even if they exist, they don't know. So I would say education.

The interface is also have to be user friendly. There has to be some seamlessly profit to do so, so people use it in the most effective way. And also I would like to point out that some of the data do need more effort into metadata and documentation of that data, because otherwise it can be the most useful dataset but we cannot know how to use it. So yeah, I would say those.

**Interviewer**

OK. Now we're gonna talk about open data intermediaries. Do you think that open data intermediaries are playing an important and positive role in an open data ecosystem right now?

**Interviewee**

Yes, but the interview is for GIS and for Esri specifically?

**Interviewer**

Not necessarily. For this particular question, you can talk about open data intermediaries in general.

**Interviewee**

Then I do think that if I take for example <redacted>, which is like the one aggregating many data sources where you can go on and search data in [country E]. I do think that part of the whole system becomes easier for the user, so the user doesn't have to go all sorts of places to get specific data, but you just go there. It's like a centralization of a database of the country. Then you go there and you search the data that you want to access to. So I do believe intermediaries are important and positive because it's easier and better for the user. But of course it needs certain structure, right? And I am talking structure in terms of legislation as well because they are fetching from other sites, which also have the same structure. If they will be doing that in the individually without any structure, then maybe they cannot provide the service that they can, at the moment.

**Interviewer**

So when you say structure, do you mean like licensing and stuff?

**Interviewee**

No, the way the datasets are published, for example, the structure.

**Interviewer**

And how do you think open data intermediaries can play a better role in the open data ecosystem?

**Interviewee**

I will go with the things that I mentioned before with the user friendly and education and also the metadata. I think that [i.e., metadata] is a sort of like a main manual for the user to use the data. If we take geospatial, for example, I can also see a movement towards more importance to like maps and to use spatial data, and how can you perform geospatial analysis. So most of the data that you get in this service is that I mentioned to you, they're not geospatial. I mean they are separated, so tabular data and then they also give spatial data. People tend to move away from the geospatial because they feel afraid of coordinates, [they] are usually scary to people.

So I think a selling point is always a map, and a visualization of the data through map is always a gateway to inform what is the data set about. So if I take that as a recommendation, use visualization and use maps to explain what the data is about. That can that can be like -- when you show a map up to a person, then they'd feel more familiar with the data immediately because they recognize the territory. And if you add symbology to the dataset or whatever you're trying to represent, then the communication is easier. I will even argue that it's easier than documenting in word sometimes. Data visualization always helps in this intermediaries.

**Interviewer**

OK, so you think open data intermediaries can play a better role in that regard, like in visualizing and making data more accessible to users?

**Interviewee**

Yes. And communicate that data visually.

**Interviewer**

OK, now we're gonna talk about [Esri distributor in country E], as an open data intermediary. First of all, have you heard of [Esri distributor in country E] before?

**Interviewee**

Yeah, sure, sure.

**Interviewer**

And do you think [Esri distributor in country E] plays a role in enhancing access supply of flow of open data?

**Interviewee**

At the moment, no. At the moment I don't see any role in that unfortunately, but also because I struggled to understand that this is within their business case because [Esri distributor in country E] is licensing. At least they haven't made any approach towards services through data, for example.

They do provide some educational courses on how to use, for example, network analysis and for that network analysis you need road maps, for example. So if you use their dataset, then you're consuming some what they call credit in their system. So whenever you have a license for ArcGIS,

then they give you a credit which are like tokens. So if you use those geoprocessing tools then they will get some tokens from you. You will consume some token. I guess that is the getaway from them making money out of open data right, because that is open data.

I don't know to what extent they curate this open data, because sometimes business models are based on: OK our input is open data, but then we curate and we make models to improve it. That is the case with many businesses that sell weather data for energy projects, for example. So in that sense, I do I understand how they use open data. But to be honest, before this interview I never saw Esri as open data intermediary.

**Interviewer**

And what do you think about the ArcGIS software? Do you think that software enhance access, supply, or flow of open data?

**Interviewee**

Absolutely not. Firstly, to access the Living Atlas you have to have an account. Sometimes you do not have to have an organizational account which is licensed, but you do need an account for the data that you can access, for example, the national death per capita, national consumption of something per capita, for example, that kind of data. For research purposes, those data are not very interesting to look at, you know.

So if they are an intermediary, then the openness of it is very broad, like basically showing something in a map from a national level with statistics or something like that.

**Interviewer**

Do you think [Esri distributor in country E] plays a role in connecting other actors in the open data ecosystem?

**Interviewee**

I wouldn't think so, but I haven't also ask them about this. So no, I don't think so.

**Interviewer**

OK. How does [Esri distributor in country E] or ArcGIS help or support you in using open data?

**Interviewee**

That goes with the previous question. Absolutely not. No, no, no, no. They offer important support to us at the university as technical support, so if something is not working with the software, but it's merely software, not data.

**Interviewer**

OK. Do you have any grievances? I mean for sure you have a lot of grievances. So can you explain your grievances about using open data related services from [Esri distributor in country E] or ArcGIS?

**Interviewee**

That it is not open. The granularity, it's very broad, so we couldn't use any of that. I mean for example, I checked when we had this challenge about Mongolia, given us zero data. I went in to check what Esri or ArcGIS Online had on Mongolia. The only thing I could get is the Mongolian boundary of the country. So what do you do with that?

Instead, I could for example use the satellite imagery to identify the building polygons or use OpenStreetMap. You can find other places where you can find this thing instead of Esri.

**Interviewer**

It then relates to the next question, how do you see a the negative or less than ideal impacts of your info an artist on other actors. When I say on other actors, not only to users, but do you see any negative impacts to the data providers as well?

**Interviewee**

Well, I will always be critical about and making profit from open data because it's like you take something that is free and then you shape it into a bit better thing and then you sell it. Some people can argue that it's legal and it's OK. We have had many conversations with them about researchers accessing data, which is not the same as industries accessing the data because it is research purpose. Data should be open for at least research purposes. But I do think that if they start with the idea of a software being closed, if you compare it with QGIS, which is something that you can compare really closely, there is negative impact because you are segregating people. You are not making it democratic enough so people can access it.

**Interviewer**

And with QGIS, do you find it easier to use, to integrate data into the software?

**Interviewee**

Yeah, I think so. For example, they have a plug-in for OpenStreetMap data. They connect directly to that server service and it shows in your map and then you can download right away.

Esri even has a problem with file format and some of the formats do not work with Esri but work perfectly fine with Python script or QGIS files that are much more light handle , you can handle lightly with those formats, they take less memory and space.

**Interviewer**

What are the example of file formats that uh cannot uh be used on that as free software?

**Interviewee**

Geopackages, for example. Geopackages are very difficult to use in Esri and very easy to handle in QGIS and also in Python.

**Interviewer**

OK, so personally you always prefer to use QGIS and look for the software and look for the open data yourself instead of using ArcGIS software, because you think even like the data that they have on their software, it's not better than what you can get yourself?

**Interviewee**

Yes, absolutely. But I have to say, though, I see that it's not in the questions, but I will tell you later how I use ArcGIS Online.

**Interviewer**

Or you can also tell me now if you want.

**Interviewee**

OK. So in my work I divide sections. First, I have to research about what we can do, what data we

have, then I shape a methodology and then I make a model. So after I make a model, I decide whether this should be done in Esri or in Python script. So once I do that then I for sure use ArcGIS Online, which is something that QGIS doesn't have and something that Python script can provide but I am not a proficient user on how to build apps, for example.

So I think Esri is going very much forward in the industry with providing an easy way to create a map application, [i.e.,] the templates or how to build your own app to be published online. So I for sure use ArcGIS Online for publishing, but not for handling or processing or making my model. And in that way, I am trying to make it open, because I open access to everyone to see the app. I open also the layers so people can download the results that we have in different projects, for example. So I make my way into making it open for other people. So in a way I am an intermediary.

**Interviewer**

OK, interesting. And so, for example, if you develop app with ArcGIS, this app is only mostly accessible to people who also have a ArcGIS account or accessible to everyone?

**Interviewee**

No. Because then you can set -- when you upload a layer or you create a web app or a web map, you can always make it as anyone can access. I'm not sure how it is called, but it's open for everyone. So what we do is we have a web page and then we link to all the apps that we have on ArcGIS Online, so people can just see it.

**Interviewer**

So my last question, do you have any suggestions or wishes of [Esri distributor in country E] business model or way of doing things could or should be?

**Interviewee**

I think they can look into the potentials of discovering data and so forth, but as I said, I don't think the data is their core business. I think the software is, but if they can provide data that is better quality and better granularity and localized data -- I do know that they have included some geoprocessing tools for imagery, for example, so if they can make that seamless for the user to download the image satellite data and be able to use it, I think it's a great potential. People will pay for that for sure, but it will not be open. It will not be free. They will use it, but it will not be free for the user.

**Interviewer**

Yeah, because even now, the data that they provide in the software is not better than what you can get from outside?

**Interviewee**

Yeah, absolutely.

**Interviewer**

That's my last question, but do you have anything else that you would like to share that will be that you think would be useful for my research?

**Interviewee**

Maybe you can also - I know that QGIS users are much more critical about Esri for example because

QGIS was basically built, so people do not have to pay for Esri. So if you can have it, I don't know any -- oh, I do know one person, but he works in a company, he doesn't work in a research environment.

**Interviewer**

And he uses a lot of QGIS?

**Interviewee**

He's actually a QGIS developer. He has written some books about it as well. He for sure has the insight on the openness discussed.

**Interviewer**

We can talk about this after we finish the recording.

**Interviewee**

Oh yeah.

**Interviewer**

Anything else that you would like to share?

**Interviewee**

No, no, I think I have that many things.

**Interviewer**

Thank you. I'm gonna stop the recording now.