

Interview 21

Interviewee	23-User-D
Interviewer	Ashraf Shaharudin (TU Delft)
Date	7 June 2023

Interviewer

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

Interviewee

Yeah, I'm a <redacted> here at <redacted> in <redacted>. I've been here for, in my position as <redacted>, for about <redacted> years now. I've been here before as a student assistant and all kinds of other internship stuff, so been here for quite long some time and basically mostly doing GIS research projects, back in the day with ArcMap, now with ArcGIS Pro and also some deep learning, artificial intelligence type of stuff, mostly with external Python libraries. I'm working in a project where we try to develop an accessibility planning tool. I won't go into much detail.

But yeah, I've been working with ArcGIS products for, I guess, let's say eight years now. So I'm quite familiar with all kinds of platforms.

Interviewer

Yeah, very good.

Interviewer

And do you use open data in your work?

Interviewee

Obviously, yes. I mean, it's an important source, but is not the only source. Because the institute have also access to governmental data, but oftentimes you have licensing issues with these governmental data because it's licensed only for institute. We have to develop some kind of derived data set to give it out to the public. So using public data always has been an important part and it is getting more and more important by the day, basically forced too. So just as an additional data source, so to say, because the combination basically is what's important. So yeah, I've been using open data or researching open data, basically my entire career.

Interviewer

So it is indeed a necessary component of your work?

Interviewee

Yes.

Interviewer

And do you have any challenges or grievances or unmet needs with regard to using open data?

Interviewee

In generally speaking?

Interviewer

Yeah. Something that you find a bit tricky. For example, something that you don't like about using open data.

Interviewee

What I don't like about using open data. There can be cases where open data is offered in uncommon formats, which are not standard or you have to convert them by hand. Or something that's not easily machine readable. It's not often the case, but it can happen. So mainly the interoperability between different kind of systems is what mostly concern me. But yeah, I would say maybe in the less than 20% of cases that's the problem.

Interviewer

Do you also have issue with accuracy or timeliness of open data, meaning that open data is perhaps outdated or they're not accurate as you would like?

Interviewee

Not in comparison to what I would use otherwise. I mean there are always datasets that are outdated but that that happens also with governmental data, so that isn't necessarily a problem of open data by itself.

Concerning accuracy, that's quite interesting because colleague of mine is trying to assess question like how accurate is open data in comparison to other data sources and I'd say it's really hard to say because measuring the accuracy or something without having access to vast amounts of ground truth data is quite hard. So yeah. I mean they are different but I wouldn't say that the data is inaccurate, but maybe different between different data sources, but yeah, I wouldn't say that it's necessarily bad.

Interviewer

Just to clarify, when you say that you use a lot of open data, are this open government data, meaning data that's from government or like other open data sources like OpenStreetMap, or both?

Interviewee

Both.

Interviewer

OK. Now we're going to talk about Esri as an open data intermediary, how do you use Esri products in your work? For what purpose?

Interviewee

Yeah, I basically, I do, I'd say, 70% of my research is somehow spatially or related to spatial problems. And I do 99% of my work there with ArcGIS Pro because I'm just used to it, by using the platform, basically grew into it from university still.

And yeah, I can use QGIS but I only do this for specialized tasks where I know that ArcGIS doesn't perform very well or slow in comparison. Yeah, that's seldom the case. It's mostly ArcGIS because I'm used to the platform. I know how it works with all the pros and cons coming with that.

Interviewer

And what are those specialized tasks that you would sometimes use QGIS?

Interviewee

A recent case where I had to use QGIS is when I want to do some accessibility calculations because in QGIS you have this open route service plugin that is developed by <redacted> or <redacted> at it's called, [the plugin] is called ORS, OpenRouteService. So I use this for isochrone calculation. It's a well-known plugin. So I basically had to use that because otherwise I would have to set up my own network analyst and that's too much of a hassle because I can just load up QGIS, press calculate, and then it does the calculations for you.

Interviewer

So in most cases you would use ArcGIS Pro from starting with collecting data to analyzing and also to visualizing data you would use ArcGIS for all the processes?

Interviewee

For analyzing and visualizing, yes. Searching for data and getting data is mostly done by hand, I'd say, so by using Google services, web scrapping, whatever. That's how I'm used to it. Because yeah, a lot of a lot of stuff is actually -- this may be a [country D] specific problem -- because a lot of data providers, they set up their own small closed environment databases which are inaccessible from the outside.

So if you know what you're searching for and you know that's in this kind of database, you have to go there and have to download it there. Because oftentimes you also need an account, for privacy reasons or whatever. So yeah, I mean, I would have to do it anyway. So I'll just do the data collection and downloading of stuff, I do manually by hand, and then I import it into ArcGIS and for the preparation, analyzing, and the whole data pipeline there. But collection is out of the system.

Interviewer

So meaning that -- because you mentioned that in [country D], I guess every <redacted> agency perhaps would have their own open data portals, and so you would go to these different open data portals, but do [Esri distributor in country D] for example, tries to combine all and integrates all this data from different data sources and put it in their ArcGIS platform, say in Living Atlas for example?

Interviewee

What's the question?

Interviewer

So it is it something that -- can you access this data that has been integrated from ArcGIS platform?

Interviewee

I think I could, but I have not looked very much into it. I've been playing around with the Living Atlas in the past few weeks for a bit but yeah, I haven't used it seriously for data collection in the bigger scale.

Yeah, I know that it exists and I know that there is some kind of data and it's steadily growing. But I mainly use the Living Atlas actually for access to the deep learning models in ArcGIS, not so much for the data itself.

Interviewer

OK. Are there particularly reasons why you don't use open data related service on ArcGIS so much?

Interviewee

Habit. Just out of habit because I'm used to doing it differently. I mean, I could go into an example of what's the hindrance I'd say. I wanted to, for my calculations with this open route service, I wanted to do some preliminary analysis of the routing network and I want to download the entire street network out of OpenStreetMap for the <redacted> in [country D]. The thing is, in the living Atlas, you can access the street network somehow, but you can't necessarily download it. Well, I at least I didn't get right off the bat how to do it. It was some more complicated, so I just went to the usual data source, Geodata Fabrik. I don't know if you know this page, which is daily updated downloads of OpenStreetMap data. I just went there because I knew it and I just downloaded what I needed because it's prepacked per <redacted>.

Yeah, it's sort of a time related issue because I mean, I could obviously knew that I could somehow access this data in the Living Atlas, but yeah, it was too time consuming for me to get into it because it's not necessary like yeah, I search OpenStreetMap, street network, [country D] or whatever, and then it automatically has this prefetched packed downloaded pack like on the other website where I just can't download, unpack the zip file and have my data ready, which takes me basically as long as typing it into the search bar, downloading, and then I have my data, compared to then on the other hand, figuring out how the Living Atlas works and how do I get this data now. It just takes so much time that I don't want to invest so that's basically the main why.

Interviewer

Yeah, because you know exactly where you can get those data and out of habit, you can simply go to this original data sources instead of figuring out how to use Living Atlas.

Interviewer

Yeah, that's the case.

Interviewer

You mentioned just now that you only rarely use QGIS, only for specific task, but do you see any advantages of QGIS over ArcGIS?

Interviewee

I mean, obviously free. I could use it at home also if I wanted to do some stuff on my private computer, where I don't have license for ArcGIS. I mean it's the biggest pro that it has obviously. Some features seem to work better on QGIS, I don't know why. I mean it's deeply technical issue. I don't have a specific example right now from the top of my head because yeah, I rarely you use QGIS, but I noticed, for example, when I did, back some years ago, I somehow found out that the doing basically the same calculation in QGIS took 5 minutes whereas in ArcGIS, the entire thing took eight hours. So yeah, I don't know why that happens and it's not really reproducible for lot of operations. It's only in a certain, maybe for a certain kind of data set, I don't know.

Sometimes, it's a gut feeling that you get after having some years of experience in using both in parallel. So for example, a lot of raster calculations I feel are faster in ArcGIS whereas some vector operations are faster in QGIS. But that's only yeah, a very rough assessment. It really comes down to the specific problem you want to do because a lot of questions are so specialized that the you can't generalize very well on that.

Interviewer

And do you see any disadvantages of QGIS over ArcGIS?

Interviewee

I find it, especially since the introduction of ArcGIS Pro, the user experience and the UI of QGIS is just more cluttered and unintuitive, for me at least.

Interviewer

Compared to ArcGIS?

Interviewee

Compared to ArcGIS. I find the way of ArcGIS, the whole window setup is presented and the look and feel of the application is better for me at least, because I find it more streamlined. A lot of QGIS stuff is hidden behind drop down menus. But it could also be a like some kind of bias because I first got into the ArcGIS world and then into QGIS later, so I don't know. But just speaking from a pure visual point, also I find ArcGIS more pleasing because it looks more modern. I guess with the 3.0 update, they have introduced this kind of materialized floating window look and I like that a lot. QGIS, the UI looks basically like it's from 2010 still.

So that's it. I just like the look and feel. And a lot of -- actually that's quite an important point. I figure that a lot of stuff in ArcGIS caters from a usability point to Windows standards. So there a lot of things in ArcGIS work like they do in Windows, so you don't have to think where to find stuff. Simple things like the save button, it's on the top left, things like that. And so yeah, it's basic stuff, but it's important. QGIS also does this stuff, but not as good as ArcGIS.

Interviewer

You talked about the UI, but what about the end product of QGIS compared to ArcGIS, for example visualization, which one do you think are prettier, easier to generate?

Interviewee

I'm not an expert on visualization because I always pass this task out to our visualization experts here. I know the basic stuff, I do know. But yeah, I'd say ArcGIS, it edge a little bit more over QGIS, but I've done visualizations in QGIS so rarely that basically I can't say anything relevant about it. But of the few times I tried something, yeah, I just quickly handed it off because it was taking too much of my time because I focused more on the analyzing and data science and, not so much on visualizing because I'm not a cartographer, I'm the information technology guy.

But I'd say, even with no experience in ArcGIS, you can produce a map that's at least decent. It's not perfect, but it's decent and you can show it to people. I wouldn't print it somewhere, in a book, but you can use it in the presentation. So yeah, it's easy enough so that even people like me who don't have that much of experience in designing can get decent results.

Interviewer

In general, do you think Esri plays a role in enhancing access and supply of open data?

Interviewee

I think a year or two ago, I would have said basically no, but since I got into the whole deep learning system for ArcGIS, I'd say yes, with exceptions. I mean I think it's really great that they have introduced this deep learning library into ArcGIS because the offer of pre-trained models and

working models is steadily growing, and they are quite usable. For me, I would say this this does fit into the open data category. I mean the models are by themselves, they are almost always open.

But yeah, training them and giving access to pre-trained models which also has to be done on a lot of data, which is a computationally very intensive and basically you can only do if you have access to a lot of computing power and also having access to the data sets is obviously a problem for a the normal person.

So as a provider of these kind of models, I'd say, definitely Esri is an open data provider and it's getting more and more important. Because they can implement all kinds of models faster than I myself can do, because I have to get into the code and next, fix stuff myself, adapt to my questions. On the other hand, I could always just download the pre-model and adhere to some kind of adaptations also.

Concerning the actual raw data, for example, building footprints, I would say, not necessarily. Or I do not necessarily see Esri as a top tier provider of open data, but that's again because I for the most part, in the past, I wouldn't have even looked if they would offer any data because I just knew beforehand where, OK, I want to use building footprints and then I go to, for example, the BKG, because they are offering them and then I could, go to here or office and apply and then they would download it for me. But basically I wouldn't have to figure out where to get them because I already knew beforehand, so I just, yeah, took the beaten path, so to say, and went ahead with what I knew.

I would be more open into getting into this all kind of Esri data providing stuff, but yeah, at the beginning right now, I've planned on checking up when I would start a new project and need to look for data. Currently more on the closing phase of some other stuff, so I don't need to download new data right now. But next time, when I will be setting up something [new], I will check out what they offer. Maybe also I would be open to using the service because I'm in the platform anyway. So if it would be easy and I could just search in the Living Atlas, and click on download and then it would download the data already prepared into my database or whatever, I'd definitely be using it's saving time when it's easier for me as a user. If it's saving me time or making my work easier than I would definitely be switching system.

So until then, it's right now it's still easier to do stuff by hand because it's not that simple as just clicking download and the data is ready.

Interviewer

So also in your organization or in your community in general, people, especially in [country D], people typically go to the sources instead of utilizing open data in Living Atlas. Do you get the feeling that is the case?

Interviewee

Yeah, also because a lot of people don't use the Esri system. I feel in [country D], it's more like, I don't know exactly but probably with people I'm working -- let's say the institute here because my colleagues are obviously all working more or less with ArcGIS because it's licensed for the entire institute -- but when comparing within the institute, I'd say 30% roughly are using the ArcGIS ecosystem, Esri ecosystem and a lot of them are using QGIS or whatever alternatives they are just because of financing reasons especially -- I've noticed this in when I was dealing with administration staff back in the start of the year, I also did some interviews with people from administration, from

varying levels, municipalities, <redacted>, whatever and a lot of them actually migrated to the QGIS world because of financing, limited budget, budget limitations, because they needed to save money and so they moved away from the paid stuff. But I guess that's a whole another -- A lot of people, I would say, didn't do it by choice necessarily out of their personal choice, but it was forced onto them. In university, it's common to work with ArcGIS I guess, or at least in those universities where I've been in or I know people from.

Interviewer

And you also mentioned about deep learning models. Do you get access to the training data?

Interviewee

Not directly, no.

Interviewer

Yeah, but do you know what training data do they use? Is it transparent?

Interviewee

That's transparent. Yeah, it's always -- before you download the model, you have to go to Esri's website and then there's a documentation on what they've used for training, how many images, for example, or which datasets they came from. Also, if you know what you're doing, you can open up the packaged model and look into the code. But yeah, that's advanced stuff.

Interviewer

OK. And do you think that Esri also plays a role in connecting other actors in the open data ecosystem? Not necessarily through the ArcGIS platform, but also perhaps through events or anything that they do that connect other actors that use open data.

Interviewee

Probably yes. I mean, I haven't been to much of a lot of Esri hosted events and but also I've been to this one summit, there was quite a broad interest in them so I'd say yes. Because just having a platform and opening up to your user base, which is coming from all kinds of different institutes and the professions and whatever. Yes, basically it's given. It's not like that Esri is a small company, so their user base is quite large internationally, so there will necessarily be some kind of networking between actors at least.

Interviewer

Do you see any negative or less than ideal impacts of Esri on other open data actors, such as data providers or data users?

Interviewee

At the moment, I'd say no, I don't see any negatives. I mean, from a user experience, it's always nice to having different access points to the same kind of data. For example, if

it's basically my choice where I can get access to the exact same dataset, I can just integrate into my workflow, it is the most efficient way. However, it only would become a problem, but it's an if, if the data source would be limited to, for example, only being accessible via one way.

For example, let's say if Esri would start providing any data set basically for free, that they they've collected or have access to or compute themselves or whatever, and they would offer the open data

for free but only inside of the platform Living Atlas, that would be a problem because then I'm limited to the platform. I have to use ArcGIS which is paid, to access the data, so it's not necessarily free either. They at least should offer the option to access anything via the Internet on their website just as a simple download zip and go. So the data access shouldn't be limited to being inside the program. But it's not the reality.

Interviewer

So does it also mean that it would be good to have various data formats, so not limited to formats like shapefile, that is proprietary format, but instead also open standards?

Interviewee

Definitely. The broader the range of offered formats, the more applicable your data is. I mean, if I don't have to transform data before using it, it's always, from a user perspective, easier because I can use the same data set in different kinds of applications and streamline my work and spend less time in data preparation, which is always a dead time because it doesn't produce any meaningful results, it's just preparing analysis, obviously. So yeah, obviously I mean the more the better.

Interviewer

Oh, I didn't ask this question before but whenever you get a data from sources outside Esri environment do you sometimes face issues with the format? Are there some times that the formats are incompatible that you have to transform the format first in order for you to use it in ArcGIS environment?

Interviewee

Mostly not. Mostly you can use directly or you can -- there are functions for transferring. Well, yeah, there are some cases which don't work near, namely being 3D City GML models, which I have recently tried importing into ArcGIS. There is a there is a function or toolbox provided by Esri or some kind of users. It's downloadable on the Esri side but it works semi good. It works but doesn't work perfectly. It has some kind of errors. Some buildings don't get imported correctly. So definitely this is the most prominent example from my perspective, the integration of 3D City GML models is not as good as I'd like to be. It's not just simply download this file, city GML, which is basically the open standard for 3D city models. There is a huge catalogue of cities offering, there are 3D model with textures in the city GML format, but it's not like drag and drop into Arc and then you have a 3D model. [Instead] you have to work around and download this toolbox first and then as I said it works semi correct only, so, that would be definitely an improvement that I would wish for direct integration of City GML into ArcGIS, [i.e.,] just drag and drop, having an import function that works built into it would also be OK. But yeah, working with 3D data. [I don't mean just] for visualization but also having a more advanced toolbox or set of tools for analyzing this data -- yeah that would be some kind of wish I'd had for the future.

Interviewer

And what are the typical formats you use when you sourcing data?

Interviewee

Can you repeat the question?

Interviewer

What are the typical formats? You take data from these sources, what are their typical formats?

Interviewee

Concerning these 3D buildings or everything in general?

Interviewer

In general.

Interviewee

The most common format I use is some kind of... yeah, it's hard to say the most common one. I mean we have stored everything here on our internal servers ready made for ArcGIS. That's why I draw a lot of from our internal, but this data was prepared at some point from a different format.

But let's say, if I focus on open data, the most common format is still the shapefile, but it's on decline, I'd say. Some years ago it was really dominant, but more and more it's because also probably because my work is changing, but I'm using CSV so just basic tables XML. And yeah, sometimes even an Excel file, but it's that's rare. GeoJSON, JSON in general is also getting more common. But, yeah, I would say those are the top ones. So I'd say shapefile, CSV, JSON, any type of JOSN format and then XML, that's the main format that I use.

Interviewer

So I guess, like, even if you go to [country D] agencies' open data portals, you can get data formats in shapefile? Even though it's a proprietary format they do provide in shapefile?

Interviewee

Yeah, actually, oftentimes, yes. It really depends because [country D]'s really splintered between <redacted>, and then every sub or municipality if it's a big one, has their own way of doing things and then cities have their own way again. So yeah, it's a quite a jungle of data providers here in [country D].

But I'd say if you're looking for -- yeah, shapefile is still the most common, even if it's a proprietary format. Yeah, because it's just that people are used to this kind of format and a lot of time in the administration offices, people aren't really crazy experts on GIS. So dealing with -- for them -- dealing with shapefile is the best they can do because they have sometimes heard about it. For example, if I would have to send them something, I wouldn't send them a GeoJSON, I'd say 99% of cases, they would just ask me if I could send in a different format, probably shapefile.

Interviewer

Because they also use ArcGIS?

Interviewee

Not even not necessarily. Just because this format is so common and it also works with QGIS. So, it's some of the most interoperable and easy -- or it's the most well-known format. Basically because back in the day, when there was no QGIS, people had to use shapefile necessarily. And yeah, bureaucracy in [country D] is very slow. So people adapt really -- adaption to change is, yeah, even slower.

So yeah, people just are used to using shapefile formats, and even if people know how to use other formats at some point in the chain, someone would say yeah, can you send me in shapefile because they don't know better. So everybody just [has a] consensus to use shapefile because at some point, you will have somebody [who] will have to convert it into a shapefile anyway. So you could just

basically, you will write from the start -- if you know you are inside your spoiled little project team, and then it's obviously can be some kind of different format, -- but if you just send stuff out onto a broader range and you don't know the people, then you commonly use shapefile and that's why also a lot of data providers offering shapefile because they know that is the format that easiest to use for people. Oftentimes they offer shapefile and some other format as well, so that's quite good. In most cases, not exclusively only one.

Interviewer

My last question to you. What is your wish list from Esri, especially in terms of providing open data related services, what are the things that you think they can do perhaps better or they can start doing it?

Interviewee

I mean, obviously, as I've told the integration of 3D City GML is a big point. And then I'd say from me, a very very nice addition would be a in direct integration of OpenStreetMap into ArcGIS, not only as a visualization as a base map, like it's right now, you can visualize it, and as I've said, it's somehow it's also in the Living Atlas, but it would be nice for having just a simple click and add type of integration of OpenStreetMap but on the database level. So that, I can select features, for example, click on the building and select the building and view its attribute table, export it, use it as an input for my standard ArcGIS function, so having a fully integrated database access level into OpenStreetMap. That would be amazing I'd say, so that I don't have to pre download everything packed but just I can define an area of interest and then ArcGIS automatically pulls the most up to data -- yeah, doesn't have to be live, but maybe live from last night, it gets harvested every night -- it pulls the up to date version of OpenStreetMap with full access to the underlying data and not only the visualization. That would be really helpful.

Interviewer

OK, so I'm not a GIS specialist, but I didn't know that now in ArcGIS, you can only use OSM data mainly for visualization? So you can't do analysis with it? Or you can use it, but it's not direct integration, meaning that you have to do it manually yourself?

Interviewee

Yes, you can visualize it as a map, that's already in the program. If you want to do calculations, for example on the street network from OpenStreetMap, you have to download the network.

Interviewer

From OSM?

Interviewee

From some kind of OSM provider.

Interviewer

Uh-huh.

Interviewee

Then you have a zip file. Unpack it. And then you can load this file into ArcGIS. And then you can do your calculations, that works. But then, you have to do the step of looking and downloading the data.

What I meant [i.e., wished] is, if I would go into the Living Atlas, for example, and then I just search for OSM [country D]. Then I click add to my project and then I have the OS ready-made in my project, not as a visualized map, but the data that I downloaded before. Or, I can type in street network. So that it can, directly inside the program, it loads the street network with attributes as an editable feature that I can use with basically any ArcGIS function like I would do. So, I don't need to go to the Internet and download. I would just save myself one extra step of downloading and unzipping and converting the data. So it would also be nice because then I could directly export the street network from ArcGIS into a different application. So yeah, for example, if I wanted to do some calculations, I'll import, calculate, exports -- now I can just use the provider to do my calculations and export and directly with the street network in the correct format. It's just a time saver.

Interviewer

That's a very good point. I didn't notice it obviously. Anything else that is in your wish list?

Interviewee

Generally speaking, not concerning any open data provision, reduce the number of times error 9999 shows up. It's still a problem and it has been for years. Sometimes the program just doesn't do what it's supposed to do, and it just throws the error. Yeah, something went wrong and you don't know. It's getting better from update to update.

I see it less frequently now than some years ago, but yeah, the entire-- it's not even the single error, even if it shows up a correct error, the description is basically useless. So yeah, the error messaging in general of the program is still not informative.

Interviewee

No, it's not very helpful.

Yeah, that would be the main things from the top of my head.

Interviewer

Anything else?

Interviewee

Yeah, they're that's the main ones. City GML, OpenStreetMap integration, and the error messaging.

Interviewer

Well, thank you so much. That's the end of my questions. I would stop the recording now.