

Interview 3

Interviewee	03-Esri-C
Interviewer	Ashraf Shaharudin (TU Delft)
Date	11 April 2023

Interviewer

So let me start with the first question. Could you please describe your role in [Esri distributor in country C]?

Interviewee

I'm <redacted>. Our team now consists of <redacted: more than five people> and we have our manager. But I coordinate. So, he kind of know more about what will be, for example long term data sets that we will work on in the coming year or future, and I will more split that into tasks that can be divided amongst the team to work on and decide what will be more short term planning. Kind of together, we make a plan of how we will -- how to determine on which datasets to work on. But I make sure that <redacted>. And also I do <redacted> myself as well about half of the time or so. And half of the time it's more <redacted>.

Interviewer

Yeah. OK. How long have you been working in this or similar role?

Interviewee

In this role for I think <redacted>. And before that I also worked at <redacted> first and then I worked in the <redacted> now for few years, like <redacted> or so I think. And then the last <redacted> years, first it was only data processing, so data specialist, and now I'm the <redacted> for the last almost <redacted> years.

Interviewer

OK. All right. Just to make sure, yeah. How is your role related to open data?

Interviewee

So we have -- at Esri, we have the living Atlas and there we publish -- there we kind of share our data and it can be viewed by everyone, but we do have like a terms and conditions which say that the data can only be viewed from Esri technology. If you are asking my role to open data right?

Interviewer

Yeah. Yes.

Interviewee

Yeah, just like the rest of our team we are processing all the data to put into the Living Atlas and also share some best practices sometimes on how to work with the data or make sure that everyone knows that the data exists.

Interviewer

So if I may rephrase meaning that you get open data from data providers, yes, and then you process them and then you put them on Living Atlas?

Interviewee

Yeah, yeah.

Interviewer

Apart from the Living Atlas, are there any other open data services that are offered by [Esri distributor in country C]? For example, is there any services that you provide to data providers to facilitate them opening up their data?

Interviewee

Not that much. I think we do have -- we have a service with postal code areas that we created and for example <redacted> uses that as an input to create their data sets. But we don't. I don't think we have. Yeah. So the other users, they take it from the Living Atlas mostly.

Interviewer

OK.

Interviewee

Or we do also some data projects, but those are not open data.

Interviewer

So yeah. Yeah. So the main one is the Living Atlas. OK. And how do you think there is service that actually provides benefit users or attract users to using ArcGIS for example?

Interviewee

First, when you want to start working with GIS, if you don't have any data, then first have to collect a lot of data. So one of our biggest services that are used are base maps and of course they can help the customers a lot to even put the things on the right place. And I think another big part of what attracts the customers is that we have made the services very usable and we have a lot of data about a lot of different topics and you can also combine it easily.

Interviewer

In a single platform? So it's like 1 all in one shop?

Interviewee

Yeah. And now we have things like Arcade, which is kind of a programming language within the ArcGIS and you can really easily without adding a lot of data to a map, just reference to other services that are all available in platform and combine all those knowledge in another layer for example, so I think that's also a big advantage that the data adds to working with our software.

Interviewer

OK. You mentioned about one of your tasks, one of the main tasks is processing this data that you get from the data provider. What are the main technical or non-technical activities that you have to do as part of offering this open data service?

Interviewee

First, getting the data in, then making, for example, a model to process it, publish it as a service, and we try to do it as much as possible that the data is always the most recent. So also create models or scripts to make sure that the data gets updated or if it's not, for example, if the data provider doesn't have update frequency we have to make sure that we get it in our planning to check it once in a while to

update the data. And yeah, so processing is a large part and we have also part of the team who is mainly focused on making sure the infrastructure works some things we do in ArcGIS Online, but we also have ArcGIS servers and so that is not my task, but from other team members to make sure that it keeps working. It's also a big part of what the team does.

And then the other part is sharing the knowledge and data. So we have this hub page where people can view like when the datasets are updated or inspiration. We also have somehow tools with instructions for people how to do things themselves and or just writing a blog for marketing, for example. So that is also part of the job and also is collaborating with other colleagues, about what is needed as data, because of course there's a lot of data you have to choose which are useful to add, so it's also talking and making sure that you choose the right datasets to enhance our offerings or that we know that customers need the data. And sometimes a lot of -- you also have to of course adapt to changes in the source data because that happens a lot that things change, so you have to adapt to that or something is not working right so also some time is spent on issues and bug fixing, things like that.

Interviewer

So part of processing is it also standardizing the data, as you said because data can changes and at the same time, the certain standards are used in in certain industries, so do you do standardization as well?

Interviewee

Not really. We try to keep as much as the as close as the source as possible, and also for example keep the same fields or things like that. We don't want to adjust too much? But we do have some of our own standardizations in that we want our services to have the same way it works. For example, first we had a lot of boundary layers that were published every year, but now we change that and we make one service that always shows the most recent year and also we have one with all of the historic boundaries. We try to do that approach for all of our services, so it's easy for users to understand.

Interviewer

You also mentioned that that there is a hub where you update if there are new data and also like I guess the data stories or blogs. Is this hub accessible to everyone?

Interviewee

Yeah.

Interviewer

Including those who are not ArcGIS users?

Interviewee

Yes, it's <redacted>.

Interviewer

OK.

Interviewee

Yeah. So it has some instruction, documents, any change of the data sets when they are updated .

Interviewer

That's great. How long has [Esri distributor in country C] been offering open data services?

Interviewee

Yeah, I checked this right before the interview. I'm not completely sure, but I think our old services are from 2012 onwards, at least as web services. I think before that there were also some data but I'm not sure how it was back then.

Interviewer

And how many staff are involved in this service?

Interviewee

So we are with <redacted> people in the team. Yeah, that includes our manager who also has - also manages other teams, so he's not completely working on the data part, but he does a lot of the more product management tasks, for example, so he's also involved. So we're with <redacted> people we are now.

Interviewer

And the <redacted> people manages not only open data, but also other types of data, or just open data?

Interviewee

And also other types of data. I think the majority is open or at least like within our terms and conditions. And we have some products -- data products as well. Two of them are related to privacy related data so users also need to have an agreement with the [country C] <redacted> in this case or another party, and we process the data but that is like a paid service. And then we have some other data products which are, for example, one is a true ortho-surface of the [country C] and another one is a network based on OpenStreetMap. And we also have data that can be used to enrich your own data with demographic data, which is more based on credits within ArcGIS. So we do have a few products based on the data sources that we also process; so, often, it's kind of a combination for example high resolution imagery we offered in Living Atlas for everyone to use within ArcGIS and then we have this data product for people who want more, better quality data. For example, OpenStreetMap, we also have a base map and POI services and we have that in Living Atlas. But we also have this network which can be bought as an extra for people who want to use network. But the paid products are a smaller part of our team.

Interviewer

OK. Yeah. All right. And what are the skills of these <redacted> staff? Yeah. I mean, what are the expertise? Are they data scientists mostly?

Interviewee

No, not really data scientists, but more like data processing. We have <redacted> people who are more system engineers, we have one person who's more software developer. The jobs are quite technical. And <redacted> people who are mainly just have a lot of knowledge about how to process data, and then we also have one person who also has more creative backgrounds, and <redacted> also focuses on more UI kind of types. So we don't really program our own apps that much but we do have some -- we also try to make nice visualizations of the data so one person is more focused on that.

Interviewer

OK. Well, that's nice. What do you think is the value of open data to society from the point of view of [Esri distributor in country C]?

Interviewee

Yeah, to make sure that everyone can make better decision based on the best data that is available. And I think that's to everyone's benefits. Making sure the data is available and also usable to make sure that every question that needs an answer has the right sources available to work with in a good way and can be shared as well to make better decisions.

Interviewer

And what do you think are the benefits of offering open data services to [Esri distributor in country C] itself?

Interviewee

Why it has benefits to us to share data? Mainly we share data to make sure -- also some datasets are already available, for example the basic registration with the buildings, for example, it's already available, but we processes and make it also a bit more usable to our customers, but in the end it's also benefits us to sell more software. And you also kind of create -- when you have so much data available, then it's more usable and usable for people to also work with the software because it's already all available. And in some cases it might also -- you might have competitors who want to offer other datasets and you also want to prevent that from happening.

Interviewer

So yeah, yeah, if I may ask a question regarding that. So for example, QGIS, which is the open source for GIS, they don't have that kind of service because they don't have a team like Esri. So do you think that what Esri is doing is basically providing not only the software, but also like this additional supplementary data to make it easy for customers to use instead of like with QGIS, they may have to look for that data themselves?

Interviewee

Yeah, exactly. And it takes a lot of time also and people might not always realize that. So also we try to make it as easy as possible for people to use and then indeed they might not look at other options.

Interviewer

But what are the costs of offering these services to Esri?

Interviewee

Yeah, I read this question. I actually I don't have any number or insight in that. But of course there are a lot of costs involved because we are with <redacted> people processing the data and we have a very large infrastructure to store it. And it's getting bigger and bigger. So there are definitely a lot of costs involved and -- but I think we also see it as an investment because in the end it will lead to more software sales. It will in the end make sure that all the all the costs aren't lost. And we also we do sometimes have some side projects to do some data processing for customers to kind of also earn something within our team. And also we have now these few data products that I mentioned before that now also yielding some revenue to make sure we can keep doing all of the data processing for also the open data part.

Interviewer

If I may ask regarding storing the data, do you also get support from like Esri headquarters for example with regard to this?

Interviewee

No, not really. We do a lot of things ourselves. We do a lot of -- try to also store lots in ArcGIS Online and I do think we get a little bit discount for storage compared to customers. And other than that, no, we set up a lot of things by ourselves, actually.

Interviewer

Yeah. Meaning that Esri in different countries would see open data services as different investments. So may they may perceive for example, the cost is too high and they probably wouldn't do what [Esri distributor in country C] does, for example?

Interviewee

Yeah. Yeah. So of course, it also helps in each country the open data policy is different, like in the [country C], we have a lot of data, so make sense that we also provide that data. But I know it differs a lot between Esri distributors. I don't know who else you will talk to but in other countries, they almost only do, for example, paid data sort of services. For example, in <redacted> now the government is, I think, providing more open data, so that kind of gets them a bit more in trouble now because if they first deliver the data for a fee, then it's now more competitive. And we already had it. So for us it's kind of makes more sense to first deliver as much open data that is already available. We will also process but on top of that we are starting to do more and more also paid services because we also have a lot of knowledge on the data and we have the infrastructure to also offer something more for customers who want more than just the basic data.

Interviewer

OK. Do you face any challenges in using open data or in offering these open data services?

Interviewee

I think the main challenge is changes that happen in the sources. For example, now you have this basic registration in the [country C], the <redacted>, and we also use that data in our base maps but now they will change the data to <redacted> and it will have a lot of changes in the data structure and also some elements are being removed, so we would have to readjust all our processes to those changes. So that is, I think often the difficult part in data processing. Of course it will have less impact the more time you get to adjust to something new, but sometimes something just changes and you didn't have time to adjust.

And also some data sources aren't really -- don't really have good system to update the data. That can also be a challenge to offer data. Or you have some topics that don't really have a standardization yet, for example the gas pipes, you have different providers and they do sometimes share their data as open data, but there's no really a structured way to do it for the whole of the [country C]. Because we also kind of have as a standard to only do whole of countrywide data, so not -- we don't really process data from one municipality for example. So, also with those kind of data sets we want there to be a uniform data sets of the [country C] and then we might process it.

Interviewer

So you mentioned, for example, these gas data, are they provide provided by the utilities or provided by the municipals that are not standardized?

Interviewee

The utilities.

Interviewer

So that's the industry wide has no standardisation yet.

Interviewee

Yeah.

Interviewer

OK, makes sense. And you also mentioned about the irregularity of update. Is it a common issue from data providers here in [country C]?

Interviewee

Not that common, I think like the bigger important data sets, I think they are processed well like they have a lot of standards. But sometimes you have smaller datasets, for example we have this service with wind turbines, which is important topic now with energy transition but they don't seem to have a regular update frequency, so we often have to check to see what's happening. So it's always better if a lot of thought has been put in before the data set has been made available already about when it will be updated.

Interviewer

Who are the key open data actors that [Esri distributor in country C] engaged with? Including, apart from you mentioned data providers and also like the users of the software. Are there any other partners that you regularly engage with to provide these open data service?

Interviewee

You mean besides the data providers?

Interviewer

Yeah. For example, is there a standard body or it's just providers?

Interviewee

No, not really. Mainly we have -- we look up data sets and we just process the data and for some datasets we have more contact with the providers also. For example, with the imagery data and the height data, we also have contact with providers of that, the [country C] <redacted>, we have a lot of contact with them. But sometimes we just process data sets without having contact with the providers. With the end users, sometimes we just get emails with questions or things like that. But other than that, we don't really have any other kind of contact that I can now think of as an example.

Interviewer

I also wanna get back to the Living Atlas. I just want to clarify do other people can upload their data on Living Atlas themselves?

Interviewee

Yes. So customers can also upload their own data. There are like some requirements for example that when you look at the service, you have this item description and it should pass certain requirements on that there needs to be a terms and conditions, there needs to be a credit for the owner, things like that. And we do have now one other organization that's putting data in Living Atlas and it's <redacted> from the railway. So they do the train tracks. So they put data in Living Atlas. So then of course, they can determine their own use, accessibility and user restrictions for example. And we do have contacts also

with other parties that might want to put their data in the Living Atlas. And I think -- and hope that will grow it much more that many others will also contribute because then the data is closer to the source and that I think also has a big advantage too.

Interviewer

If they put their data on the Living Atlas, can they also put their data on their own platform or on somewhere else?

Interviewee

Yes.

Interviewer

And do you have to check this data that they put on living Atlas?

Interviewee

Yes, someone can request the data to be uploaded in living Atlas and then we as a team curate the data. And so it has to meet this certain requirements, and when it does we also often -- we also check the data when they put it in, so then we might say also for example, maybe if you cannot click on the objects we can say: oh, you can also put in a pop up with more information. So sometimes you give some suggestions as well and just check if there is really data and if it's really working. And if it does, and we think it's a good enhancement, then we will put it in Living Atlas. And then after that and I think we do an annual -- every year we have to check the items in Living Atlas, so then we will check if they are still working correctly but in between, we are not really sure what the data provider provided us with the data because it's their responsibility.

Interviewer

OK, do these data providers have to be institutions or organizations, or can they be individuals like researchers, for example if they have data and they want to put it.

Interviewee

I think we don't have requirements on that but we would prefer if it is an institution and preferably like a country wide institution but also university can also of course participate. But for example, not a single municipality because it's too small and so it should also offer something to more [audience].

Interviewer

OK. Next question, do you have any example of projects or cases where Esri's open data services demonstrated impacts?

Interviewee

Yeah, so you mean what people did with our data?

Interviewer

Yeah, that you can think of that actually provide real life impact.

Interviewee

We have a lot of examples. I think one big project that we also worked on was the <redacted>, which <redacted> provided <redacted> maps. And I know that is also used, for example in archaeological research. So that I think has a lot of impact. Also, our height surface is used a lot in research I know. We now also have the data set with energy labels and we know there's also a lot of people use that, for

example, municipalities to check where -- which houses or neighborhoods require more energy labels. Those are a few that I can think of.

Interviewer

The one that you do with <redacted>, that's not part of living Atlas?

Interviewee

Yeah, it's also in the Living Atlas. Yeah. It's like completely free data. It's distributed by the <redacted> but we also put it in the living Atlas. So a lot of services, we have our restriction that it should be used in ArcGIS, but those <redacted> maps are kind of an exception that can be really freely used.

Interviewer

Now we want to move on to the topic of Esri in the whole open data ecosystem. Do you think Esri plays a role in enhancing access, supply, or flow of open data?

Interviewee

Yeah, I do think that. And I think mostly it is by enhancing the data and providing examples on how to work with it as well. Because a lot of data is already available so you could say in that sense we don't play a role because people already get it from <redacted>, for example. But I think we have advantage on also providing what you can do with the data and make it more useful.

Interviewer

Do you think Esri plays a role in connecting other actors in the ecosystem?

Interviewee

Yeah, I think so too. Yeah, of course we as Esri wide, we have like our conferences to connect a lot of users and they also exchange information on how they work with data or share experiences. I do think that we can play a bigger role in the sense of, for example, I mentioned with the gas pipes, for example, we could also play a bigger role by, for example, setting -- trying to make a standard or take a leading part in that sense. I think we do play a lot of role to connect users but within the data we could play a bigger role by bringing together parties and organizing maybe more events and talk about the data and how we can combine it better for example. We don't do it that much yet.

Interviewer

Well, with only <redacted> people in the team do you think it's possible?

Interviewee

No. And also of course we have a lot of technical people in our team and playing this kind of role, it's different kind of skills and capabilities as well.

Interviewer

Do open data providers ever need to pay to obtain data or services from [Esri distributor in country C] that are developed based on the providers' open data?

Interviewee

It could be if they would want to use it but I don't have an example for that. So the data providers can also view the data that we processed within ArcGIS and we do have some products based on the data.

But in this case I don't think the data providers that release the data also use it. Yeah, they could. I'm not really sure if I understood the question.

Interviewer

Yeah, meaning that they can freely use the data if they use the ArcGIS software, right? But what about, for example, if you find any mistakes in the data that the open data providers provide, do you provide feedback to the provider?

Interviewee

Yes, we sometimes do that. Sometimes you get of course questions from customers, like there's something wrong in the base map, for example, and then it turns out it's a mistake in the <redacted>, for example. And I know that [country C] <redacted> has a site where people can give input on where there are mistakes so often try to say to the end user that they can provide that feedback there, but we also have direct contact with data providers if there is -- if we think there's an issue or something should be different or, for example, something has changed and we don't agree with it or there's an error in the data or something like that, so we provide the feedback.

Interviewer

OK. Yeah. OK. So when you provide feedback, then you will also correct the data?

Interviewee

No, we don't correct the data ourselves.

Interviewer

OK, so unless it is corrected by data provider then?

Interviewee

Yeah, because we don't want to -- we only process the source data and we don't take any adjustments to the data itself. Sometimes we add something, for example, if we have a dataset and we want to add the municipality name then we can add that as a separate column and add that to the datasets, but that is already something we don't do that often. But we don't, for example, the house number of the house wrong, we won't correct it in the data because it should be improved at the source we think.

Interviewer

But there are big mistakes, then you would alert the <redacted>, for example? The data provider.

Interviewee

Yes.

Interviewer

Do you have any suggestions or wishes of how Esri's business models or way of doing things can be improved? Are there any things that can be changed?

Interviewee

Yeah, I think we have quite a good way that we do things and try with offering a lot of open data try to make products that use the data as an extra service. But yeah, I think we could have more -- it's really difficult sometimes to really get the input from the users what they need and not only what they think they might need, but like really having a better overview of what is needed and how the customers will use it, to have the best understanding on what to provide and how.

Interviewer

Do you think events like a <redacted> is an avenue for you to gather this input from users?

Interviewee

Yes. Although we do have -- we have a lot of presentations then as well and we could do some more interactive sessions maybe, with user groups. I think that can be an advantage, but now we already -- there are so many days and plans and we don't have that much time to really set it up at the moment, but I think we can benefit from having more user input or bringing users together.

Interviewer

Do you have any wishes from other actors in the open data ecosystem on what they can do better? For example, do you have any things that you wish open data providers can do better to facilitate you in providing the service, open data service?

Interviewee

I think mainly just making sure there is a data structure that doesn't change. I think just one source as much as possible -- don't provide data that already another data provider provides for example. And making sure that if the data scheme changes that they will inform us and that there's a lot of time to go to the next data set. And I think collaborating with a lot of other data -- people work in data and make sure that people kind of look at it the same way.

Interviewer

What about from data users? Are there anything that they can do better to improve the whole ecosystem, you think?

Interviewee

Yeah, I think of course, yeah, like you mentioned also when there are errors in the data to alert the data providers. Also, some end users might not realize they are also data providers, for example, municipality has to deliver data to the <redacted> so they also have their own inputs in enhancing data. And I think users should just think of how they can already use the data that is already available as much as possible, and if they have new data also think about how they can share it and if it's possible to also to make that available to other users.

Interviewer

Yeah. OK. So that's the end of my questions. Yeah. Do you have any other things that? You want to share with me?

Interviewee

No, not. Really, I think, yeah.

Interviewer

OK. All right. Thank you. Then I would end the recording.