Interview 10

Interviewee	10-StO-C
Interviewer	Ashraf Shaharudin
Date	18 April 2023

Interviewer

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

Interviewee

Yeah, currently I'm a member of the management team of <redacted>. We're also reorganizing a little bit, so in a couple of weeks I'll be <redacted>. And we have two members of the board and a director. So currently I'm more in management role within <redacted>.

I've been working there since 2013, I think, almost 10 years. I've always been involved in different topics, but one of them that is relevant for this research is that the advice report for <redacted>. Basically the open data for geospatial data for [country C] government.

I've always been a strategic and tactical adviser there, both to the <redacted>, who is the former owner of this platform, and the <redacted> who manages and runs the platform.

Are you familiar with the general role of <redacted> within [country C]?

Interviewer

Maybe you can explain it.

<Here, about 7 minutes 5 seconds of the interview is redacted from this transcript because it contains a lot of unique information that poses risks of reidentification of the individual>

Interviewer

What is the value of open data from the point of view of <redacted>?

Interviewee

I think we see two main values. On one end, it's the perspective of openness and transparency. Government should always be as open as possible in order to act and to function properly from a democratic point of view. With the economic drivers there as well. I think it depends who you ask within the government. We are both cooperating with for instance, <redacted> and they are really into the second pillar of that. And when we're working with the <redacted>, they're really into the first pillar. So I think both values are valid.

And we basically, we always try to do -- because we're a <redacted>, we always try to make standards actually work and function in real life and solve problems. So it's always really has a strong user focus. It's a user-driven development.

And it should [be] both cases are valid because basically, first, it doesn't matter whether you are a company that wants to create some added value on top of the data, or you're a journalist or researcher, or just carrying out your work process within government more efficiently because you're reusing data. It's cheaper. Basically trying to facilitate all kind of open data use.

Interviewer

OK. Thank you. I'm going to move to open data ecosystem, which is defined as a network of interdependent yet self interested open data actors. Before I move to talking about what you think about the current open data ecosystem, I want to refer to report <redacted>. Because the report mentioned that for SDIs to remain fit for purpose, they need to evolve towards flexible, open, agile and self-sustainable data ecosystems. I know that this is SDIs and not necessarily open data, but there's the word sustainable, which is quite interesting to us because we're also talking about sustainability, so could you please elaborate on this vision?

Interviewee

First small side comment, as you put SDIs are not necessarily open, I think within [country C], the general opinion is that SDI should be open. But it's true that there are variations between different countries within Europe. But I think also if you have a look at INSPIRE data, almost all INSPIRE data is available publicly, available from [country C] data sources, whereas in some other countries you have to request access and request the data. So I think the perspective that we're working with on a day-to-day basis is that the [country C] SDI, at least, should be as open as possible.

And to elaborate on the flexible, open, agile, and self sustainable data ecosystem, I think it's important to compare this to the current state of INSPIRE, because INSPIRE was always meant to be an open data ecosystem and that has been designed with specific purpose in mind, [e.g.] sharing environmental data in case of some kind of disaster, and the idea was disasters never stopped at administrative boundaries, so it should be -- you should be able to combine data across borders and it need to be available because as soon as the disaster hit you don't have time to start thinking about the creations of data or publication of data, it should be made available already. So in the case of emergency you can just access the data immediately.

But at the same time, although it was a specific use case, it was also designed to then how to drive the reuse of this open data, because they already have the idea: OK, there is a use cases more qualified, it's a high quality, it's something you should do because it's necessary. And not necessarily a quantifying use case where we say, OK, we are adding value and you can really measure it in euros. But it has always been idea that data should be made available for reuse as well. But there, some changes were made.

I think looking back, the INSPIRE, current state of INSPIRE is quite specific to geospatial specialists. You have to be specialist in order to understand the data, to understand the services involved in accessing the data. So basically it's quite a high threshold for reuse of open data. You need to be specialists, only focusing on a small niche of geospatial experts. And if we have a look at the infrastructure, that's basically -- the main cause for that is that the standards that they adopted already existing international standards, but they made more specific, they create almost an add-on, if you want to comply with OGC WFS service, you have to do this but for INSPIRE, you also have to do this, this, and this. So they come up with additional requirements and these additional requirements lead to the situation that you could not

use any tool that just supporting the international standards. You had to look for a tool that was adapted to this additional requirements that were INSPIRE specific.

And as a result, the market was less interested because it was only a small usage that really specific INSPIRE, points are, I think almost valid from theoretical academic point of view, but for most users they don't care whether it's done like this or like that as long as they can access the data and use it. So basically in hindsight, the current INSPIRE infrastructure is too complex, it's almost too regulated.

And that hinders easy access, and it's really high threshold to enter the system and to be part of it. So what we saw in – oh, and another thing in the current standards of INSPIRE, all these different standards are part of the INSPIRE directive. So it's written in European law. You should use this standard and that standard. But changing the law takes years and years and years. Technology, it's going much faster and much higher pace. So, basically, the legal part of the framework cannot keep up with the developments on the technological point of view. So now we have a directive that basically says, you should use outdated standards, you should use complex standards. There are already lighter, easier, more generic standards available, for instance if you have look at the web feature service, the way how you can transport vector spatial data from one system to another.

Basically have to be specialist to work with it. It's a PDF document 500 plus pages and you almost have to read it as a newbie from the beginning to the end in order to be able to understand how works and how we should create a query to get some data out of the system.

Basically, the new generation of standards is completely based on the Rest API design principles, and basically everybody that's used to working with data on web, regardless whether they are geospatial experts or not, as long as you're working with data as a developer, basically within 5 or 10 minutes, you have this API up and running because it's a more predictable way of interacting with the data. So what we mean with being flexible and adaptive is that as soon as these new standards emerge, they are lowering the threshold to participate in this ecosystem. Then it's really important that you organize your infrastructure in such a way that you can adapt, that you can say: OK, we are no longer only using the old standards and it's also OK, in order to fulfill all the legal requirements, if you choose the new standards. So we have to be flexible.

And the same thing, I think, from really detail point of view, core additional requirements are valid. I mean there are not nonsense but basically, their mindset was: let's think of the most complex use case that we want to serve with this infrastructure. OK, what requirements does the infrastructure has to take in order to be able to serve this most complex use case. And as a result, there's a lot of complexity through the infrastructure that's only really needed in the most complex use cases. But in other 90% of the use case they are much less complex. They are quite straightforward and even if you want to do something that's quite straightforward, you're confronted with all this unnecessary complexity. So basically they added all the complexity for just the few rare occasions that also the use case is really complex. What we also mean with this adaptive and agile ecosystem is that it should match up, for instance, we see that within OGC standards nowadays being modularized. Otherwise, the standard was huge and basically it has so many requirements in order to show off all use cases. Now the new <inaudible> standards to have <inaudible> as the core, just a small set of requirements. And with a small set of regulations you can serve say 80% of the use cases. If you want to support more functions case, you maybe need to implement an extension or two extensions. So basically then, there's more of a balance. If your use case is more complex, your implementation will become more complex as well

because you have to implement more of these extensions or modules in order to also serve some more complex use cases. But the nice thing is that as long as you have simple use case, it's enough to work on with the core.

And that's also like being this adaptive thing, because it's basically the requirements to participate in ecosystem are now more in balance with the complexity of your use case. So if you want to do something simple, it has to be simple to participate in the ecosystem, and that's really I think the core message of this. Because it has repercussions for the way how it's organized from a legal perspective, because if you write a certain standard within a law, you know that it takes 5 to 10 years to change the law, you know that you will hit some problems within three years time.

So it's necessary from a legal point of view, say OK, there will be standards, and they are managed over there on that list, and that list can be updated twice a year, for instance. And make it less heavy legal process to change, for instance, the technical requirements. It also requires in the technical point of view that you have an open eye: Ok, what are the current developments in standards, in ways, in technology, and how people interact with data to be more in line with what data users expect, and really adapt more to the current state of user needs. I think one of the drivers also from <redacted>, I think also within Europe, we are one of the front runners in that we really try to transform from only our geospatial needs into making geospatial data available through the standards of the web. Because the web is already a federated system of data, it is an ecosystem, it functions with millions and millions and millions of users. Whereas the geospatial part is just a tiny, tiny fraction of that.

And the more that we can align with those more generic standards of the web and how we describe data and how we make data discoverable, how we describe it, it would make metadata easier to get. Because I think for most use cases, if you want to work with geospatial data, you need the developer that's used to working with data. But the geospatial aspect, which is really important in the actual collection of the data, for creating the use case for the developer, I would say 90%, it's not relevant to understand how spatial works but you can create of offer kind of query or whatever kind of service you want to build. Especially to make the data ecosystem much more open and inclusive, so that it should made easy for all data developer, basically everybody who has some kind of experience with data, whether you're a journalist or a developer or a company, you should be able to work with the data.

The whole idea is that by making the standards easier, by making infrastructure easier, it would lower the threshold of engaging in the system and taking advantage of the availability of open data. That's really a thing that we're standing for.

Interviewer

Yeah, that's fascinating. But I probably missed it. This new standards that you mentioned, is it already in the pipeline?

Interviewee

Yeah, basically OGC is working on it in really high pace. Some of the standards, they created specific website for it, ogcapi.org. What they're doing -- because the old family of standards, they were all W dot something service -- the web feature service, the webmap service, the webmap tile service, and basically they are all from, say, the early 2000, 2005, basically that time period. And they are now are all being transformed into API versions. I think the web feature service is also I think one of the two

important services within INSPIRE. There's already the OGC API features, but the core and two of four, I think they're <inaudible> for extensions. Two of them are already ready and released. Others, which is OGC API standards are almost complete or only the core is complete. So basically in this year, next year, I think the majority of the standards in this more modern API version will become available.

INSPIRE is already doing some small experiments. And how can we or how can you benefit from applying these new standards and basically everybody is enthusiastic from a functional point of view. But basically the huge task for the European Union is to come up with enough security for data providers. In the law, in the directive, we still mention web map service, web feature service, but it's also acceptable <inaudible> trying to create legal loopholes so that you can also use this more modern standards.

Interviewer

Alright, with the development of SDIs in [country C] and Europe in general since INSPIRE, what would you see are key lessons learned for other non-geo open data ecosystem? What should and shouldn't be done?

Interviewee

I think one of the key lessons is keep it as simple as possible because geospatial data, it goes back to land surveyors, and they're really focusing accuracy -- centimetre, millimetre -- so it's always -- there's this drive, they're driving for even more accurate, more reliable data, almost striving for perfection. And if you want to serve as many users as possible, and also if you want to serve use cases that are relevant from a societal point of view -- for instance in [country C], we have this housing problem, we have nitrogen problem, climate problem, energy problem -- if you want to be able to use open data in that field it will -- because basically everyone says that they're all spatial problems, where you want to build, you don't have room for nature or to create solar farms, so basically we don't that much space in [country C], so it should be data-driven approaches. But I think still not enough people realize that it will only happen if the data is so simple to use that climate specialists or agriculture specialists can use it.

As long as we say, oh, we have great data, but you should ask us to help you, then it's never going to run. So the whole idea of having successful infrastructure from a user perspective is to lowering the threshold to use it as much as possible. And I think going to the way of web standards, more generic standards. Spatial is not special anymore. That's also really valid for technology and for standards and how you interact with the data.

Reference discoverability is also a nice thing. I think for a long time, and also the INSPIRE directive states you should describe your metadata according to ISO standards. ISO metadata is brilliant in describing spatial metadata but basically it's non-existent for generic administrative open data, they're working with DKAN and all kind of other formats.

So I think the lesson is always to have a wider focus, understand that you're part of a broader ecosystem. I think there was also a period where we thought they would have to basically let go of the geospatial standards altogether and focus on these more generic standards solely because it's the only way to gain access to this wider audience of data user.

I think we now have also concepts like federated architectures that we see that it's OK to have multiple platforms, to have multiple access points for data, and maybe some access points is more from

geospatial perspective with those kind of standards, and others more from administrative point of view. And it's OK that they are both there as long as they interchange as much as possible with their data. They link to each other, for instance. So there's the principles calls in some [country C] government documents: no wrong door principle. It doesn't matter where you enter as a user, the important role of the infrastructure, it should help you regardless which door you enter. So for instance if you enter the administrative door because you don't know that there's also a geospatial door when looking for data, you should be able to discover also spatial data behind that door.

And maybe, for instance data in metadata DKAN format, where you will need really specific method that's only in ISO, it should link you to the register where the method is recorded or published in ISO standard so that you can have a look at it small additional fields that are maybe really crucial for you as a user. It's also the other way around, if somebody enters the geospatial door and said, well, but I'm looking for administrative data, you should not say, well, so we don't have data. You should still link them and help them, and basically it's not the fault of the user to take the wrong door.

Basically everybody that's working on this ecosystem, they have the responsibility to make it work. So it's a collaboration and I think especially the nice thing about this whole idea about federated architecture is that there is not one single central platform. And because if you think my ideal solution is 1 central platform, yeah, as long as there are two well, which one to pick? Is this one more important for the administration, is this one more important for geospatial, which one? With federated architecture, we understand, they both have valid reason to exist. They both have additional value to the users. The only thing that we should take care of is that the user is not hindered by the fact that there are two separate solutions. So they should link, they should cooperate and they should serve the user regardless their background. So I think that's an important lesson.

When talking about federated architectures, the Internet is the federated architecture. It's proven it works. It works for documents and it's working for data more and more. It's one of the big steps for the Internet from the web document where Web page change to web with data where an object links to another object which links to another object. This building may link to university, but it also may link to this address, it may link to the municipality. So basically how you can navigate through pages, you can also navigate through data.

And think that we should understand for instance, the INSPIRE, the old INSPIRE structure, they say, OK, we'll start to create something almost like a green field, we start from scratch, let's build something. And now we understand, we should be really cautious of what's already there. The Internet is functioning, so you have to adapt and make sure that it functions as much through those kind of standards. There are already a lot of data usually user, so instead of saying we should have more data users that understand this geospatial niches, no, we should make it simpler. Accept the fact that the majority of users is not geospatial expert. They're just interested in data that might or might have no spatial component. And it's just an attribute. So you should deal with that.

And I think the last importantly from infrastructure point of view is that we could basically say governments that are running these kind of infrastructures, they take the initiative and of course they design it and companies participate in the ecosystem. But basically they're quite strict dividing between public parties, they come up with the design, they make the rules; and private parties as long as they play according to the rules, they can be part of the ecosystem, but they do not have a real influence in the way in how the infrastructure is going to develop.

And the problem with that the governments are, in general, not the most adaptive organizations. I think private parties are much more capable in adapting to changing circumstances, to changes into technology. So there should be -- I think you cannot create completely equal playing field, but it should be made more effort than now. We should acknowledge the strong points of these private parties.

Because I think the old way of thinking within government, which, for instance in [country C] we have the base registry large scale topography and I think of all the trees that are in [country C], only about 4% of the trees is registered in this open data set. Now if you want to do an analysis on heat, heat islands within urban environments, basically, if you have only 4% of the trees, you can't say anything about the impact of having more or less trees on these heat islands within the city. There are at least two private initiatives as run by small combination of Lidar and all kinds of data, they created data sets of basically 99.5 or 99.9% of the trees. So the data is already there. I think the whole reaction of our government, oh, we lack data, we should collect it. Whereas, especially in this federated system, OK, other parties have already have it. They are willing to make it available. Just make clear how they can be part of this federated system and then some of the data sources are public, some of the data sources are private, maybe some data sources are released in some kind of public private partnership.

I think the realization is really down that that we need all these data sources in order to solve these actual factual problems. Now, for example, if you want to have a look at mobility, the charging stations for electric cars, they are currently not in the base register because at that moment when it was created 10 years ago, you had phone booths on the streets, but no electrical cars. There are already quite a number of parties have the state-level available, but if you want to make policies on municipal level one, do we have enough of these charging stations to make sure that we create more sustainable mobility? But when you're looking on your own data, we do not know. But the data is there. So I think it's really important that the next step of this ecosystem, private parties are not only allowed as a user, but also as a data source. So that's important lesson.

And I think this whole concept of federated architecture, federated structure, it's really beneficial to organize that. And I think it's -- because I see this idea emerging all over national governments, but basically the concept of the data space, from Europe data space, also the federated system. If both private and public data sources, they should be able to connect all these sources and I think the laws should make it more user friendly.

I think in the beginning, the [country C] open data policy was really strict. You can only publish open data as is and as soon as you make it a little bit more adapted, a little bit more to what user needs, there's this discussion of only private parties are allowed to do that, you should not compete as a government with private parties. I think the way and how people think about what's the role of public, what's the role of private, it's also shifting a bit.

I think the last thing is making data available in a way that it's more user friendly. Because still the biggest user of governmental open data are other governments, other governmental bodies. So in a way it's quite weird in the way of thinking, the government was not allowed to publish data in such a way that another government was not able to use it -- to say no, if another government wants to use my data, you should hire a private company to adjust it, make more flexible or have a little twist. It's just weird. So I'm really hoping that in this current movement into data space, or federated architecture, whatever you want to call them, there's much more collaboration. And it's an understanding that private parties need the government to sometimes say, OK, this data, this is most important data. For

instance, if you are applying for building permits or do your analysis based on this data source, this data source is the data source. We will guarantee you that -- we will not say you're using the wrong data and as a result, you're not getting your permit. I think we're really creating a more level playing field for all actors and they're, I think public and private parties, really need each other.

Interviewer

I have two follow up questions to what you said. So the EU data space that you mentioned, is it the same vision of federated architecture that you envision?

Interviewee

Yeah, basically the concept of data space is a bit more from thematic approach, the European Green Deal data space, the idea is that INSPIRE infrastructure already holds a lot of data free and make it more rich. They basically have the data-driven approach for the Green Deal, there's a Green Deal data space on mobility, on healthcare and I think 2025 or something like that. So although it's described more from a thematic usage, I think if you're going to compare it, they both say we need to have access and to organize and to make common operational rules and regulations to make it possible that we have access both to public and to private data sources. So if you have a look at their architecture document, then it's basically the same as a federated architecture. In [country C], it's a bit more architecture driven because then it's seen as a solution because I shouldn't compete your platform or my platform, oh wait a second, both platforms are valid, which you should just improve the way and how they interact and how they make each other stronger and easier to use. In Europe, it's a bit more user focused or use case oriented. But in the end I think if you compared to the architectural ideas they're pretty much the same.

Interviewer

But do you see the risk of arranging it based on themes would create ecosystem that do not interact with each other for example?

Interviewee

There is, I think an actual risk. At the same time, I think you can – one of the best ways is to adopt standards of the web. If they all use comparable standards in how to access data and how to discover data, use the same API patterns this time, maybe in five years time, it's something else, maybe it's linked data or it's both, I don't know. I think both approaches are valid and they will find each other because basically they are trying to do the same thing.

And I think it's almost use case dependent. In the current [country C] case it's quite helpful, but it's coming a bit more from an architectural point of view because it solves some of the current problems that we see in [country C] or the obstacles that we encounter in [country C]. But I think it's European goals that they are more thematic this is also valid point of view. Basically, they're both valid. They're more user interface let's say, more of the higher <inaudible> of why are you working on this data, on this architecture or anything -- it's a synergy between the two.

Interviewer

OK. You also mentioned about allowing private parties not only to be data users, but also data providers. Allowing is one part, but do you also think like the incentive for private parties to share data is also not there.

Interviewee

I think the way you know how private parties are looking today is changing and that's the result for open data policy. I think 15 or 20 years ago, data was an asset and people were selling data. And even quite generic data, base maps, also <redacted> was selling their products. Now all these are in base register and are usually free. So if you see business models with private parties to add value doing smart things with data. And I think that's really good. Because in the end, the real value of data is in the use of data. So when they come up with smart solutions for digital twins or whatever, when they shift to more advanced operational data, where is modeling or analysis or visualization, that's where these companies excel. And that's where we need most of the developments. And I think private parties are better at these developments and improving technology. That's really what they have to offer to us.

So I think for a company, it's attractive to be part of this ecosystem. If lots of people use their data because basically if they want to do use your data, you can sell them some of your functionality. I think that's also the case with Esri in [country C]. I think in the beginning, for instance people in the <redacted> was thinking, Esri is a competition for us because we try to distribute as much open data as possible. But every Monday they come to <redacted>, they get the base registers, they shoved in their own serving their own clients. Ah no, all clients should use our system. Why? Basically the results of open data is the sum of the use. We have billions and billions of hits in <redacted> and about the same number again in the Esri environment and the sum of it is the success of open data. As long as the data sharing is a part, they are paying for server in the cloud storage, everything. Let's help each other and from that point there was really a change of mindset and basically <redacted> started talking to Esri: we see that you update now on a weekly basis, but we're getting more and more data updates, it's getting bigger. Is it okay the way we serve you? Is it still OK for you or you need some tweaks so it's easier for you to update your data and transform it into your own format. It shift really way of changing and how you think about data?

So I think really from the open data point of view, the mind shift is already there. And where there's still a bit more debate or the debates undecided is, OK, when governments publish open data, they're always in open formats. For instance, Esri is using open data that should have proprietary formats, and they are serving their own users. That's their own right but from a reuse point of view, if their client use it in their own process, well why would you care about. For instance, if it's municipality that's using Esri software, they use this open data that's maybe originally published original sources <redacted>, it's harvested by Esri, it's served in their proprietary format to municipality and municipality creates some kinds of open data service for their inhabitants based on a proprietary Esri format.

Then the open data or and the derived open data set is a little bit less open. I mean, the data is still open, but it's not an open data format. So then suddenly people, for instance citizens if they want to interact with this data, and there's a lot of web technology now, so it's not that complex, but the problem is there are municipalities that use Esri software, so if you want to apply for a permit for an event, we have a tool that we provide base map and then you have to draw, OK, here is the podium, here is the first aid thing, if something go wrong the ambulance can still enter from here and from here

but there will gate over here, and then suddenly people have to use a proprietary format and proprietary system to do it.

Is that a problem or not, I think that's still undecided. I think the people that are most strict say no, it should always be open format because they're still in a way, if they're lucky, if you knew if it's getting too expensive, I want to go to their competitor. It's almost impossible to change vendors because it's everywhere in the processes and in a way that's may hinder -- not saying that it's actually happening -- but it may hinder the way our government works may because it's almost impossible to go to another vendor, it's becoming more expensive. Because there's no real competition on price anymore because the cost of changing is so high that you accept that basically, your monthly fees are 10% higher than with your competitor because it will cost you an enormous amount to have the whole transformation process done. Then it's something that is frowned upon within the government.

So if that's still in a way difficult. On the plus side, and that's really true for Esri, I think 10 or 15 years ago they were not interested in open standards at all. They said we do have the standards, it's called shapefile and it's the standard almost everybody uses it, everybody can read it, it's the standard. It's not something OGC standard or ISO standard <inaudible> Everybody uses it. Same thing with them. I think now they are more interested in metadata, they created some of them, they have developed on their own. They made it OGC community standards that is recognized as an open standard and also it's judged as an open standard in the way how it's organized.

And their uptake of modern OGC standards is much better than old generation standards. And I think in the old generation, they also had point, sometimes their own format, which are own proprietary formats had better performance. Yeah. For the sake of doing derive thing, offer something with less performance. Yeah. I mean, I completely understand why they did it. And that's also a thing that is a challenge to standardization organization. It should not only be a derived thing from a theoretical or academic point of view, but it should also work in real life. And not only on one machine, but the standards should work in the cloud, it should be scalable and it should also work when there's like an immense peak usage. For instance, <redacted> in the in the beginning, there's some quite nice examples, there was a map with the quality of surface water if you want to go swimming. In the summer it was first day of 25 plus degrees, and one of the news I think at <inaudible> starts with an item, there's now these websites, it's <inaudible>, so you can check your local swimming spots, you know the water is actually healthy or not. Within 5 minutes, they had two million users. So you need to be scalable. And then it pops out, OK, some of these standards, they work nice with one user, with five users, with 1000 users, maybe 100,000 users, but 2 million, it's not going to scale anymore. Basically if you want to have a real sound and grown up ecosystem, you have to be able to do that. You have to be able to scale, to deal with these kind of big requests. And I think new generation of standards does it, and so being model. Esri now is OK, we have our own standards because we know them in and out and we can optimize it for our own range of products. But if you want to publish something in open format, just click the box, it's possible as well, so there's also a real uptake of open standards within Esri as well. And that's important to mention. And if you again -- good news for certain municipalities that run most of their spatial processes on Esri platform, it's becoming more safer to do it because they're still in line with all kinds of legislation about open standards and open formats. So I think they're shifting in the right direction.

Interviewer

Alright, yeah, I'm going to ask you more about Esri, but before that I just want to ask you about open data intermediaries in general. Do you think that open data intermediaries are playing an important and positive role in the open data ecosystem right now?

Interviewee

Yeah, I think the most important intermediary I can think of in the [country C] case is <redacted>? And I think <redacted> are service providers, not data providers. They're service providers so that the individual data providers don't have to take care of the scalability issue, the performance issue and time issue. The nice thing is the data provider is really into their data, they have some more knowledge, but what you see happen in reality is that <redacted> is much more aware of the user needs.

And there you say, OK, of course I want to access, for instance, the building registry, but I want to combine the data with another source because I have this question a million times a day. Those kinds of user request always pop up on a service provider and not on the individual data provider. I'm only concerned with addresses, so combining it with other data, yeah, don't ask me. So basically this intermediaries playing a really important role in the transformation being the supply driven infrastructure into a much more demand driven infrastructure. Because now users have an entry point where they can post on request where they can say OK this is not nice format but it would really be helpful if the data is available in that format as well. And when you're serving it for hundreds of data sets it is much easier to see those kinds of trends and to understand what kind of formats you should invest. So I think these intermediaries like <redacted> are really important.

And at the same time, an intermediary like Esri is also really important because although they're a huge driver towards open standards and open formats and maybe even more open source tooling in the future, the reality is that a lot of the municipalities and governmental bodies that have to work with the data, they are not capable of creating everything from scratch. They don't have the funds, they don't have the staff that's capable of creating a local infrastructure on their own. So they rely on basically, these full service providers like Esri and other competitors. They can say, OK, we know that your task is this, this and this, we can help you with it. We take care of data storage, we take care of visualization, we take care of how you combine all those, all these different data sources. So I think Esri also played an important role in lowering the thresholds of using data.

Interviewer

OK. What do you think <redacted> and Esri for example can do better?

Interviewee

I think the shift that's currently topic of debate is offering services that integrate data. We have a lot of different data sources, we have different base registers or even more data sources of course that do not have this formal status but basically they are still organized almost like independent silos. Addresses optimize for addresses alone. Large scale topography with large scale topography. But the user is not interested -- there is not single user: oh I want to access the address registry or I want topography. I want to know if they have information about buildings and maybe if you want to know everything about this building, some of it is starting the address registries, large scale topography, the small scale typography, the real estate, tax data set.

So basically to serve this user you have to almost play the role as the intermediary to know where to collect all the tiny bits of information and combine it to really come up with an answer because the current open data infrastructure is basically do it your own infrastructure. You have a problem? OK, you can search for data. We can point you to different data set. You can make a connection or download the data which you have to do the analysis yourself. We give you the building blocks for solving your problem, which you still have to solve your problem with yourself. So basically you should be capable of using QGIS or Esri software to combine this data to do your analysis and to get your answer. And the data specialist is perfectly fine with it, but a climate specialist, he could not be bothered to use that kind of system. They want to know how many charging station for electrical vehicles are in city center. And he doesn't care if the data is coming from different data sources. He just should be able to ask the question and get an answer.

Because then you're really helping the customer. I think that's something that <redacted> can do much better. And the <redacted> is still a bit reluctant, is it the task of a public platform or is that something that if you want to do that, you should go to a private company and pay for it? And that's still a debate. Bit it hinders the actual uptake of data-driven approaches within all this fields that are typically not that tech oriented, or that data oriented. I mean climate specialists, some of them know data, but most of them don't know data and they don't care about data, so it should be as easy as possible. Yeah. Why should why using Google? Because you can ask something and they're really helping you. You get an answer instead of building blocks instead of hey, good luck, mate.

Interviewer

Yeah, but do you think this debate that <redacted> has whether they should do this integration or not? Is it because of resource or because of the legal around it?

Interviewee

Both. Legal concerns is really a blocking issue. At this moment in time and the Ministry of <redacted> also thinking about changing law on this point and making it a specific task because <redacted> has its own law where it says what <redacted> should do. The role of having a data platform and access is already in the law, but integration of data is not yet in the law. They could change the law and they're thinking about it. So legal is really a valid point, but also they're struggling with funding because open data is a success and if you make calculations about added value for open data, it's always a positive business case. But the place where the benefits go are different places than places where the costs are being made.

So <redacted> needs more money because I have more data, more users, they want to provide better services. But basically who's going to pay because their current, all their open data services are free to use. And as soon as you say, OK, you have to pay for it, even if I charge you one cent for open question, it's a psychological issue because I'm now confused, I'll try to do my own, even if it's 1000 times more expensive than to do it.

As soon as you start charging money, people tend to back off and say, I'm not going to use it anymore. So then the problem is the government should come and say, in the bigger picture, the value being created, so there is some value added tax, so we have more tax revenue, so we are going to pay

<redacted>. But basically, the closing of the loop, it's not working in government. And I think most parties have problems with that. It's not only <redacted>, it's true for the entire <redacted>.

Everybody's positive about what they're doing but they still have problem of getting funding and especially additional funding for these developments and bringing it to the next level. I think that's something where -- because it started like really technical thing and only tech minded people were concerned with it but basically the big decisions about money are made in politics and those kind of people, they are not involved in data, they have no clue about technical issues whatsoever. They think oh, it's boring. We have much to do with policy, we create policies and this just the actual operation, putting the policy to work out. We don't want to be bothered with that.

And that's a big issue. If you really want to have a successful digital government, you need really serious funding. it needs to be on the political agenda as well. For instance, if you have a look at <redacted>, everything about data and digital transformation is a political issue. Their Prime Minister has an IT background. <redacted >. So as a result, there is a completely different mindset within government.

And as a result, <redacted> government is developing in a ridiculous pace, if you compared to the [country C] base because they have the political support, the political backup, and they have the funds. So they are working with solids and with data poles and in all kinds of technical solutions to make the old data network more ethical again and make it work with regard to privacy as basically try dealing with these issues. In [country C], we see that we have a problem, but nobody's building a solution yet because of that issues.

Interviewer

OK, I want to ask the last question on Esri [country C] as an open data intermediary. But do you mind if I take another 10 minutes of yours?

You mentioned a lot about Esri, but do you have any suggestions or like ideas on how Esri business model should be changed or could be changed for the benefit of the whole ecosystem, so not only to Esri.

Interviewee

I think it's basically, along the part that they've already chosen. Indeed, Esri was really a software company. They sell software and installed it on machines and it's shifting more and more to a service company. And I'm a strong believer that -- for instance, we also are working on the idea of how an SDI can evolve into basically not only spatial data infrastructure, but into a digital twin infrastructure, because then you don't do not only have to deal with data, but also with all kinds of functionality, models, analysis, visualization and again in the federated architecture. So in ideal world you would be able to pick data from private services from public services and maybe I want to run model at TNO and I want to put the output in a service, in a web viewer that's offered by Esri and then the type of platform is used to interact with it in game environment as that people can use it for participation meeting.

Basically, that's the ideal world that's Esri continue in and compartmentalize their functionality even more. So that you can have visualization as a service, or having spatial analysis at a service. Even if I'm not running ArcGIS whatsoever from my computer, but I want to run basically on the web and check for my use case, I need that data, that's the best model for me, I want to engage that kind of functionality.

And if they really play part or become a part of that ecosystem so that they can decide, maybe they have a great model on shadow analysis, then I want to access that as a service. And then basically, in order to make it work, we need the standards on how to interact because we have lots of tension: How we can make data interoperable? and how do we make functionality interoperable? How to speak to models and it doesn't matter whether I ask Esri to do a shadow analysis based on a couple of days or competitor model? how they run their own model, I don't care. Optimize it because I won't see in the end how it works together.

Interviewer

Yeah. OK. Just one last question. But do you think that the problem right now is because Esri is having this monopolistic power, especially in geospatial ecosystem? And so do you think that in this instance, for example, tighter regulations regarding this, you know, should come from the government?

Interviewee

Now regulations is one side of it. It's the question of the carrot and the stick. There should be something in it for you, but also, if you're really ignoring regulations, there should be some consequence.

So the regulation part is the stick part, there needs to be some consequences, but it should also be made interesting. I think it should be made interesting for Esri to compartmentalize their functionality even more, because basically for them it means that the 30 or 40% of those who do not use Esri, they might use Esri as well on specific parts.

And I think that's an important thing because then they can do business, they can generate money from it. Because people are willing to pay for it. That's the issue. There's no need for free functionality. At the same time, it should make certain criteria, so for instance, if I apply for building permit and then there's this data analysis, they use the official data sources, I use the Esri model for shadow analysis and if they can show me OK, I'm certified to do this according to the official regulations. OK, it's valid, I meet all criteria regarding shadow, then I should get my permit almost automatically. And so this guarantees that the way how they run it that it meets certain criteria, I can trust the outcome of the model, For instance, if now I use an Esri digital twin, the outcome maybe "ah yes, you should get permit" but if I use digital twin of their competitor it should not be "it's not allowed". And that's a problem because then we are using technology for diverging cases – for one report it says it's an excellent move, another report says it's a bad idea. Which reports to choose? If you make it more complex and more tech savvy than we're not solving the problem. Basically, we want to find common ground and understanding that we agree on the outcome because then they are really helping.

Interviewer

Yeah. OK. Thank you so much.