# Appendix

**Tentative interview questions**

Background

1. Could you please describe the function of <organization>?
2. Could you please describe your role in <organization>?
3. How long have you been working in this or a similar role?
4. How is your role related to open data?

Open data

1. How long has <organization> been implementing open data?
2. What do you think is the value of open data to society from the point of view of <organization>?
3. How does open data implementation benefit <organization>?
4. How does open data implementation cost <organization> (financially or others)?
5. What are the key challenges faced by <organization> related to open data?

Open data ecosystem *(defined as a network of interdependent yet self-interested open data actors)*

1. What is your perception of the health or sustainability of the current open data ecosystem?
2. What do you think can be improved in the current open data ecosystem? You are more than welcome to describe more than one aspect.
3. With the development of spatial data infrastructures (SDIs) in Europe, especially since INSPIRE, what would you say are key lessons learned for other non-geo open data ecosystems?

Open data intermediaries *(defined as third party actors that enhance the supply, access, and/or flow of open data and/or relationships among open data stakeholders)*

1. Do you think that open data intermediaries are playing an important and positive role in an open data ecosystem right now? Kindly explain your answer.
2. How do you think open data intermediaries can play a better role in an open data ecosystem?

Table A: Challenges in the ODE and a sample quote from the interviews

|  |  |  |
| --- | --- | --- |
| ID | Challenges | Example quote from the interview |
|  | **User-driven** | |
| C01 | Different data/metadata standards | “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 DCAT and all kind of other formats” |
| C02 | Different open data licenses | “For instance, we are trying to combine governmental data with data from OpenStreetMap and [they are] different licenses. We have license on [our] governmental data which is very open, it's similar to Creative Commons. While for OpenStreetMap, you might know, it's an ODbL, the share alike, and you cannot combine open data with a very free license with share like.” |
| C03 | Siloed open data domains (e.g., across sectors) | “I think another aspect which I think should be very much in focus when we speak about, well, whatever we call them, open data ecosystem or whatever, is the aspect of the domain. We speak about a spatial data ecosystem, but what we see is that – and we have years and many years of building up an infrastructure, spatial data infrastructure – but what we can see is that much of the rapid development that is actually creating value right now, is happening outside the domain. So, I think the biggest risk is that we keep on being a bit siloed, and not well connected” |
| C04 | High technical threshold to use open data | “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” |
| C05 | Unfulfilled user needs | “[There is a] need of focusing attention on users and the use cases, the problems to be solved, rather than focusing on the data provider. Sometimes we forget the use cases and the user and their needs – I think that's an error” |
| C06 | Limited feedback from lay users (i.e., non-expert users) | “[Feedback] are mostly from professional users, so to say. So, rarely from citizens, that can be really small companies like independent, […] one man business” |
|  | **Circular** |  |
| C07 | Loss of open data providers’ revenue | “The income from the datasets that turned open data […] for the federal agencies in the states, it may be several millions. For them it's an issue” |
| C08 | Limited value return from data reuse | “Now, they [i.e. end-users] are just harvesting a lot of value free of charge. They're not really giving anything back. We have to develop the cooperation further. So it's more also in the mind of the private sector to give back to the ecosystem” |
| C09 | Limited use case visibility | “We want to work on more visibility of the reuse of data. We publish data as open data but we don't get in return what [others] did with this data? How they used it? […] This is good for [public agencies] to motivate them [to release open data]. But also to see what has [been] accomplished out of [open data]. Maybe there is something developed [by others] that [public agencies] can use for their work” |
|  | **Inclusive** |  |
| C10 | Limited open data from non-government sectors | “There are some information that we want to get from the economy [i.e. private sector] but we don't have this open data law [in] the ecosystems. The law is just for us, for the government to do” |
| C11 | Lack of incentives for voluntarily publishing open data | “If it's a voluntarily action, and there is no guy, no law forcing you, you need to have benefits and I'm not sure that this works in [redacted] at least” |
| C12 | Requiring viable business models | “You can't force any of those companies to do it [i.e. be an open data intermediary], because of course they have to find a way to earn money with that as well” |
| C13 | Overlooked non-government open data | “If you want to do an analysis on heat islands within urban environments, 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, by some combinations of Lidar and all kinds of data, they created data sets of basically 99.5 or 99.9% [rough estimates] of the trees. So the data is already there. [But] I think the [typical] reaction of our government: we lack data, we should collect it” |
| C14 | Practical constraints in multistakeholder engagement | “We can always be better in touch with each other. And we are open to it, but also to a certain limit. We can't talk every week to every software provider or intermediary or whatever” |
| C15 | Lack of data awareness\* | Input obtained from the validation exercise |
|  | **Skills-based** |  |
| C16 | Limited knowledge of open data providers | “Some organisations don’t know the proper regulation to apply in every case. And that happened, for example, in the case of metadata, which there is a European regulation directive. Sometimes they are difficult to implement for some local level organisations” |
| C17 | Limited knowledge of (potential) open data users | “I see the ICT domain is really developing a lot of interesting, also ecosystems you could say – [from] cloud, to edge [computing] and whatever technology, and [they’re] super fine systems. But when you then need to add the content, the data, it doesn't fit because they [i.e. end users] didn't have any data knowledge” |
|  | **Foundational** |  |
| C18 | Poor open data quality\* | Input obtained from the validation exercise |
| C19 | Incurring maintenance costs for open data provider | “It does cost something. We have to get the information, we have to make products and we have to service it and platform like [redacted] costs quite a lot of money too. So as long as there's enough budget to keep that running, then it's very sustainable. But it depends on the budget” |
| C20 | Incurring development costs for open data provider | “Users get used to have more and more information, they demand the information to have great quality, great updating. The technology goes so fast and they asked to have it more quickly and it is a cost for us to be up to date in the newest technology and progressing day by day” |
| C21 | Technical difficulties in establishing open data management systems | “The reality is that a lot of the municipalities and governmental bodies that have to work with data, they are not capable of creating everything from scratch. They don't have the fund, they don't have the staff that's capable of creating a local infrastructure on their own” |
| C22 | Complex and/or rigid open data standards to comply with | “We see that OGC [Open Geospatial Consortium] standards nowadays [are] being modularised. Otherwise, the standard was huge and basically it has so many requirements” |
| C23 | Heterogeneous data administration | “At the moment we have situation that approximately half of [administrative units] has open data, others have not. The perspective is that, with HVD [high value datasets under the EU Open Data Directive] from June next year, in theory, any of the states would have to. However, we notice that some of the federal states try to escape and they find gaps” |
| C24 | Privacy concerns | “Now there's quite a lot of information available and people start combining this information and through combining you can draw more specific conclusions, you get more specific results. [But] which also enter into the privacy of people. We've got maps and there's open data of buildings and addresses. Combining buildings, addresses, maps, aerial images and whatever, you can easily get to where people live and how the environment is – [anything] is possible” |
|  | **Broad** |  |
| C25 | Inflexible/unclear government-market boundary | “For instance, our aerial image is open data, and there are companies who are providing services with that open data. We now encounter the situation that we want to do similar services for the whole country and we are not allowed to because we've got a law in [redacted] which says there has to be fair play between governments and companies. They make the products and services, they earn money with it, [and] the government is not allowed to give it away for free” |
| C26 | Reliance on a political agenda | “If you really want to have a successful digital government, you need serious funding. It needs to be on the political agenda as well. For instance, if you have a look at [redacted] and [redacted], in [redacted], everything about data and digital transformation is a political issue. Their Prime Minister has an IT background. And he's not only the Prime Minister of [redacted], but he's also the Minister of digitalisation. So as a result, there is a completely different mindset within government” |
| C27 | Inflexible governance/law (esp. with evolving technology) | “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” |

Table B: Potential contributions of open data intermediaries and a sample quote from the interviews

|  |  |  |
| --- | --- | --- |
| ID | Potential contribution | Example quote from the interview |
|  | **Technical** | |
| P01 | Implement federated architecture | “I think we now have concepts like federated architecture 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, 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 principle called, in some [redacted] 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” |
| P02 | Integrate data (e.g., across sectors/ administration) | “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 organised almost like independent silos. Addresses optimise for addresses alone. Large scale topography with large scale topography. But the user is not interested in [a single dataset], [for instance,] I want to access the address registry or I want topography. [They’d say,] 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 [from] the address registries, large scale topography, the small scale typography, the real estate, tax data set” |
| P03 | Transform data into open standards (esp. web standards) | “They [i.e. intermediaries] are playing an important role since they have been adapting their tools to open data format although it is true that sometimes it has been difficult for them to get out of proprietary formats” |
| P04 | Customize data (based on use cases) | “One of the things users of our open data always ask is to give it in different formats. We provide information in internationally recognised standard formats so that it's open to everybody. We don't do any industry specific formats or company specific formats. We just don't do it. It's not that we are not allowed to, but if we do one we have to do them all and it gets messy and troublesome. So what we want is we provide it in an internationally recognised standards and we want the companies, the markets, to provide it in all the industry standards because we also recognise that our format is not always the most useful for all kinds of users – [for example] architects and building companies and whatever they want it in DXF and anything. Now we've got 3D and they want it in BIM IFC and whatever” |
| P05 | Offer process automation | “Once we provide the information, we have to give a step more, allowing to taking advantage of the big data based technologies and to create automatic processes through artificial intelligence and using clouds as technological support” |
| P06 | Develop open-source tooling | “An intermediary like [redacted] is also really important because they're a huge driver towards open standards and open formats. Maybe even more open source tooling in the future” |
| P07 | Provide direct technical services\* | Input obtained from the validation exercise |
| P08 | Offer freemium data platform\* | Input obtained from the validation exercise |
|  | **Non-technical** |  |
| P09 | Foster public-private collaboration | “Public-private partnerships are, in my view, a key mechanism for enhancing the role that data intermediaries can play in an open data ecosystem. They could be defined as long-term contracts between a government agency and a private entity with the objective of providing a public asset or service, and in which the private party assumes a significant portion of the responsibility, risks and, generally, the potential benefits” |
| P10 | Foster public-civic collaboration | “We want to know what the civil people need so that the government can implement it. We want to know where a lift doesn't work so we can repair these lifts. So we want to collect this information, this data, from the civil people, so we can do better services for the people” |
| P11 | Implement multistakeholder collaboration | “So our role is to create the standards and also do that always in an open process so that all stakeholders are involved and are participating from the start, because it's really important to invest in participation from the beginning of the process because it influences the uptake of these standards afterwards. Because otherwise it's always – ahh it's not invented here, my standard is better, I'm already using this” |
| P12 | Perform open data advocacy\* | Input obtained from the validation exercise |
| P13 | Invest in open data-based civic tech | “My goal as an open data officer for [redacted], is to improve the work of the civic society, the civic tech, to have certain kind of an exchange where they can tell me on which civic tech projects they're working and which civic tech projects need to be implemented [by] the government themselves. […]. Therefore, I have these heck days where I have a good exchange [on] which ideas they're generating to profit from it and go to the government and tell that these are projects where [government] can put money into and what we [i.e. civic organisations] can implement” |
| P14 | Showcase open data value | “Show benefits. Show countries where it works. I think <redacted> would be a good example because they have open data, they have it nationwide, they have good quality” |
| P15 | Promote open non-governmental data | “The intermediaries can work also from the economy [i.e. private] side or the civil side on working with open data, on returning open data to the government” |
|  | **Combination** |  |
| P16 | Provide consultancy | “We also have a consultancy, [redacted] for example, consulting the government [redacted] [following] the law. But a lot of companies from the economy [i.e. private sector] asked whether there can be an open data consultancy for small and medium-sized companies and also for startups, showing them how they can work with data and why open data is very important” |
| P17 | Streamline cross-administrative processes | “Disasters never stopped at administrative boundaries, so 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” |
| P18 | Facilitate internal reuse of open data | “We have been running this geoportal only four years. And in four years we became the main information system within the City Council […]. Our colleagues have been giving feedback to us. They say that running their own businesses have changed since they started to use geographical information systems because -- they had numbers, […], for example, it's very clear when you see an image of what you are spending money on and how it's related to [impacts in] some boroughs” |
| P19 | Facilitate feedback on open data | “Those kinds of user request always pop up on a service provider and not on the individual data provider. [From data provider’s point of view:] I'm only concerned with addresses, so combining it with other data, don't ask me. So intermediaries play a really important role in the transformation [from] being a supply driven infrastructure into a much more demand driven infrastructure. Because now users have an entry point where they can post a request where they can say, this is not a nice format [and] it would really be helpful if the data is available in that [other] format as well” |
| P20 | Provide education on data literacy and skills\* | Input obtained from the validation exercise |