

1 **Interviewer:** Yeah, so maybe we can go to figure number one. Maybe you can keep this
2 figure in your mind throughout this interview where you can see, like... different levels of
3 e-infrastructures of the research data. In level one you can see the IT-infrastructure
4 which [is] needed for data storage backup network. In level two, you can see the
5 research data creation process and some list of research infrastructure. Yeah, and you
6 can see in level three, more general, or general category about the research data
7 management. And the last level, it's... a kind of a guide and information for you about the
8 research data services. Maybe we can move to the next slide and... So here, actually,
9 we will... start by looking at this wheel. As you can see, these are the different six
10 phases of the research data management, and we will look into these capabilities in the
11 grey boxes, so you can tell us more about each phase separately. Maybe we can start
12 shortly without taking any notes with the... You can maybe tell us briefly about your
13 research with the focus on the data you collect or generate.

14 **Participant:** Yeah, so as you know I am [Name]. So, starting with that I think, I... could
15 start with what I do. What I did in, at least the last five years, is that I am kind of railway
16 dynamic engineer. We are modeling and also measure some train track interactions.
17 Basically, we are designing the railways, how they look like and how we could increase
18 the performance of railways. In specifically my case, we look at transitions: when you go
19 from a bridge to normal tracks. So, that transition, how it works, is a bit challenging. We
20 were looking at that, what are the challenges, what are the problems, how we could
21 increase and improve the performance of that specific section. So, it could be from (--)
22 You go to a tunnel, you go to a bridge, you change the type of the track, you go to
23 Germany and then suddenly you see you have different things. In general, we are
24 looking at that specific section, how we could improve it. Throughout my project we had
25 two-three phases: one was initial modeling and then afterwards we did the measurement,
26 which I think today we are going to talk about that section, and then we use those data
27 to calibrate our models and we use them for later designs.

28 **Interviewer:** Okay. Nice. Maybe it will be easier for both of us to talk about recent
29 project or section, as you mentioned. So, you can remember everything about your data
30 and the plan is to start with these different phases. So maybe you can start with the plan.

31 We have also different things we want to know about it. Like, if you get any support or if
32 there is any cost for this phase, it will be nice if you mention for us.

33 **Participant:** Okay, so I will start with the data. I will move to the phase of what data I
34 collected. Of course, we planned the(--) We had a bunch of sensors in the field. First, we
35 planned, "okay, we need to measure only one train passage", which I think was in order
36 of one minute of data with sampling frequency of, I think, it was 2000 hertz. So, you can
37 imagine a couple of megabytes for each train passage. But the problem was that in the
38 end, we couldn't only record one train, because we couldn't add a trigger level. When,
39 when the train was coming, we had to have some kind of trigger point and then
40 afterwards check, "okay, there's a train, start recording", but our system couldn't work so
41 we had to record continuously 24 hours.

42 **Interviewer:** Sorry, may I ask... So, you are now speaking about the Collection? Collect
43 the data?

44 **Participant:** Yeah, I'm just talking about the planning, how we planned it and make it
45 work, and then I will go back to the data. Basically, we couldn't make them(--) Like, the
46 system, I think, are the providers of the sensors, they couldn't prove, couldn't help us out
47 to(--) Help us to only record the train passage which (..) One or two minutes each train
48 passage. Then we had to record all the data for future measurements. Then (..) What we
49 did, first we planned, "okay, how many sensors do we need? We, I think, had in total 34
50 sensors in the north of Sweden, which was really cold and harsh environment. So, we
51 had to have a remote system, which I could do all these measurements through my
52 office here at [university]. What we did, we managed to have a system, a computer there,
53 a cabine [sic!] there, we put everything in that. And then also... From here I could
54 measure what I was doing. And also, another thing that we had as a headache was the...
55 Internet. We didn't have proper Internet there, we had I think 3G, not 4G, not 5G, we had
56 only 3G Internet. Then we managed somehow. I was able to log in and do some minor
57 works and do some uploading and downloading through this system even though it was
58 a bit difficult and time consuming. (..) Anyway, so yeah! That was the main thing that we
59 had so... I think I should... Is it correct now? [inaudible]

60 **Interviewer:** Actually, I just wanted to ask one more question regarding the planning. Do
61 you think (.) the planning phase, it's working well for you or... you face more issues than
62 what you mentioned...? [speaks slowly]

63 **Participant:** We needed more help because at the end, you know, in this kind of project
64 you have a limited budget and limited time. So, I think a bit more support and help, how
65 we could manage this system with, through the [university] or somebody else, it would
66 have been so nice. Because, as I said, we were recording every single second and that
67 was, like, too much of data. I think I have four (.) hard disks full of data each. I don't
68 know, many terabytes here with me... That's why we (--) If we would have (--) If we could
69 have received some help that managed that only recorded train passage, we could have
70 had all the data in one or two hard disks. So, that was one of the things that in the
71 planning, we couldn't manage that, make that work. Of course, the planning was really
72 important. Also, another thing was that (.) we weren't sure that we can mana(--)) we can
73 do this task at the end. Because, you know, many delays because of traffic and so on,
74 so forth. We couldn't manage. But if we just look at the [university] point of view, I think a
75 bit of help (.) from the storage data, that how we could handle it, that was one thing that
76 we didn't know what to do with it. We didn't have any help, so that's why we went with
77 the hard disk. We bought some hard disk data and stored the data there.

78 **Interviewer:** Yeah, maybe this is about the Preserve. Later (.) we can talk about it.
79 Before we move to the Collect phase, maybe I have another question about the
80 Planning. Do you have any legal restrictions here with the planning phase or
81 considerations?

82 **Participant:** We have, as far as I understand. I think we have a bit because the project
83 is related to *Trafikverket*, and so we are not allowed to share some information in total.
84 Probably we could share some, but not all. And that needs to be asked if we want to
85 share anything. [speaks slowly] (.) It's a bit important, sensitive data... they call it.

86 **Interviewer:** Yeah, sure. Then if you think that you mentioned everything about this
87 phase, if you like, we can move to the next phase, the Collect.

88 **Participant:** So, yes, as you wrote there, yes, I was(--). So, we had 34 sensors, and we
89 were recording continuously for eight months, and sometimes I was just stopping it
90 because I didn't have enough storage (.) space, so I was just stopping it and then start
91 recording again after one or two days. The idea was to measure the long-term effect of
92 the settlement, so we needed all the data from day one to day end. Then afterwards,
93 that was the main thing that we wanted to look.

94 So, another issue was that since we had the sensors there and it was winter, sometimes
95 the sensors were broken, and we weren't able to know if the problem is in the sensors.
96 or in the recording or in the software. I had to travel there, of course [soft chuckle], to
97 check if there is a mechanical issue or something. So, that was a bit expensive. I think...
98 our system could have been smarter than that to find it out but, of course, we didn't have
99 that at that time. But that was one thing that I'm(--). I can say about the (.) collecting the
100 data. Then I was, as I said, we were recording the data in the field through a computer
101 there, then I was uploading the data to cloud and then I was downloading it here to my
102 hard disk. This process was always working. Record, upload, download in my hobbies
103 and I was doing analysis-

104 **Interviewer:** Actually, this... This will be in the next. If... It will be great if you tell us
105 about this collect phase, if you get any kind of support. Because you said you had to go
106 by yourself and check the sensors and if they are broken and you collect some
107 information and you check if the data that you are collecting is fine. You did not, like,
108 think about getting any support in this phase?

109 **Participant:** Support from [university] do you mean, or...?

110 **Interviewer:** Like, even if it is outside [university] or inside [university]. I don't know if
111 you try to get some support.

112 **Participant:** I think we had(--). So, basically, we had the sensor providers; they were
113 trying to help us. But of course, it was really hard to answer some questions, whether
114 the problem is because of the sensors or others. At the end, I had to make the final
115 decision by traveling there and just checking and observing what's the problem.
116 Sometimes, of course, there were mechanical issues because somebody (.) walked

117 there and put (.) their feet on the sensors and they broke the sensors. That was one or
118 twice that we realized that the issues, yeah, random things, happened. So, no, I would
119 say some support, but not all.

120 **Interviewer:** And is there, like, an extra cost for this phase or is it included in the project
121 in total? In general?

122 **Participant:** Actually, it was a bit extra. Because we had to control that the data is
123 correct. We had to do parallel measurements. So, we had to do some complementary
124 measurements, which then added some cost to us. And the number was high, actually.
125 It wasn't a small number.

126 **Interviewer:** Okay. So, if you feel satisfied about this phase, maybe we can go next to
127 the Analysis. Like, how you do analyze your data, what kind of tools are you using to
128 analyze them?

129 **Participant:** So, starting with the software or language that I used, I used MATLAB and
130 Python. Of course, the data that I recorded, they were binary data. I was converting
131 them to physical data, like acceleration, displacement, so on and so forth. So, I was
132 converting that first. And then I was doing some filtering, signal processing, and so on,
133 so forth to extract what I needed and also remove the (.) data that they weren't useful,
134 like there were no train passes. I had to find the trains and then just extract the train that
135 I want, do the signal processing, look into the (.) interesting data that I wanted to see,
136 and then record those that I wanted at the end as clean data for next analysis. So, that
137 was the thing that I was mostly doing as analysis.

138 **Interviewer:** And what kind of issues you faced during... analyzing the data?

139 **Participant:** So, one thing that it was(--) I think a big one, was that I was expecting to be
140 able to convert acceleration to displacement, which is just a double integration. But at
141 some point, we realized that the ratio between the noise and the signal was high. I
142 couldn't get rid of the noise in the accelerometers. Basically, those data that I recorded
143 in the accelerometer, they weren't useful, but the rest were. The main thing was that I
144 wasn't able to, or nobody was able to extract the data from the accelerometer. Because
145 the accelerometers were, I think, not those that we wanted. And normally, it's the case

146 that accelerometers are tricky. So, that was the main thing, I think, in the analysis I
147 realized.

148 **Interviewer:** Okay...

149 **Participant:** I couldn't do that in a pre-phase. I couldn't check it before. Because before
150 I installed them, they were working as I wanted. But... You know, sometimes you have
151 noise because of this electromagnet from the rail and stuff like that. So, there is some
152 white noise, which comes from the electricity and so on, so forth. That was the thing that
153 we couldn't get rid of.

154 **Interviewer:** Okay. And during this phase, did you get any support, or you did
155 collaboration with others? Let's say with [university research data support] [titters]. You
156 know, they have this kind of support.

157 **Participant:** To be honest, I didn't get any help. No, I think I was single person on that.

158 **Interviewer:** Since you said that you faced some issues while analyzing your data, so
159 what did you do after you faced these issues? You tried another way, or you never
160 thought about getting support?

161 **Participant:** Mostly I was (.) looking at books, papers and literature, how they do, how
162 they deal with it. And what are the tools available in the, like, yeah, the Internet, of
163 course, and the people, my supervisor, they knew. So, in that level. We didn't ask any
164 externals for help.

165 **Interviewer:** Okay, I understand. Then! Make it available for the others... What did you
166 do in this case? And what kind of data are you making available to the others? Is it raw
167 data or after processing it, or only results, or... What kind of...? [speaks slowly]

168 **Participant:** So, basically there are two kinds of data that are available. Some are
169 published like final, final analysis. It's just a curve that people, they can use it. It's
170 published and it's out, it's in a paper. Just one needs to (.) ask I will share that, it's like a
171 couple of values. And I think that's open, that I could share it but still, I need permission
172 which I think I asked for permission once, I shared it with one person and that person
173 used it for a paper. Also, there are some train passages that I extracted myself, which I

174 think in [university] at least some master students and one or two PhD and postdocs
175 have used the data. But of course, again I asked for permission and I shared it
176 individually, so it wasn't like online somewhere, I haven't uploaded them anywhere.

177 **Interviewer:** So, it's up to requests for the students?

178 **Participant:** Yes, exactly.

179 **Interviewer:** Yeah, so these kinds of data, like, in the level of results after processing
180 the data you are sharing it, not the raw data.

181 **Participant:** It's just clean, clean result at the end. Just need final responses. [inaudible]

182 **Interviewer:** Okay. Did you face any issues with this phase, like legal issues or...?

183 **Participant:** No, as far as I think the source was *Trafikverket* so I could ask "okay, you
184 paid for this data" and the data is this which I have already published. So, it's out. "I'm
185 going to share this figure with someone, is this okay" and they said "yes, it's okay".

186 **Interviewer:** Okay, are you sharing that only with students from [university] or are you
187 also able to share it with people outside (.) the university?

188 **Participant:** There is one from [European university] but that was a shared collaboration
189 between myself and that person. I think it's only, mostly [university] or(--) Because this
190 project was part of a European project, so if anybody from that European project asks
191 for the data, probably yes, we could share it with those people. But outside of that
192 project, I'm not sure if we can.

193 **Interviewer:** Okay, thank you for the information. Then if you don't have any other thing
194 to add here, maybe we can move to the next phase. It's Preserve. How you preserve
195 your data. And again, it would be nice if you tell us, like, what kind of level of the data
196 management you are reserving? Is it, like, the raw data or through the processing or
197 only the results again?

198 **Participant:** So, I think(--) One, we were thinking about it two ways. One was that to
199 save them in the cloud or somewhere – I think you were thinking about that – but at the

200 end we realized it's a bit expensive for the project. So... At the end, we came up with the
201 (.) hard disks which... I just have them here; there is one here.

202 **Interviewer:** Okay.

203 **Participant:** Always next to me, there are four more here, so I have a bunch of them. (.)
204 We recorded in them all the raw signals and also those that I extracted myself. So, if
205 someone is, let's say after me, is interested in the data and they want to go back to the
206 raw signal, still the data is there and they can use it, of course. But also, the extracted
207 data is also there, so both are available and saved here, but they are in the hard disk.
208 And [inaudible]

209 **Interviewer:** But are you putting any documentations about (.) this data somewhere?

210 **Participant:** So far, I hadn't any time to do it. Mostly if someone needs, they need to ask
211 me so that the documentation is actually with me.

212 **Interviewer:** Okay, so in some years maybe you will forget about that. [laughs]

213 **Participant:** Probably [laughs]. I guess I will forget about it, but I hope that I have some
214 time after my PhD to spend a few days and give some guidelines for that.

215 **Interviewer:** Okay. But you did not get any support from your department. You know
216 there is, like, a different (.) storage and locations from, for example [university research
217 data support unit], there is small storage for free. You did not think about that?

218 **Participant:** I think that was too small, the storage space that we had. I think it's in order
219 of gigabytes, but we were thinking about terabytes here. So... That's why it was... I think
220 it wasn't enough for me. I think the only thing that I use, I use the OneDrive, my own. I
221 disconnected my laptops and everything from the OneDrive. I think I just connected the
222 OneDrive there from the [inaudible] and the OneDrive from my personal laptop to just
223 make it free for the data. The data was coming to my OneDrive and I was downloading it
224 to the disks. That was all I did and I didn't get any help for that, because as I said we
225 discussed that and then they came up with that "it's a bit expensive, maybe you should
226 find another way". Then we came up with the hard disks, which was much cheaper.

227 **Interviewer:** Okay, so (.) as I understood you did not get any support in this phase as
228 well.

229 **Participant:** No.

230 **Interviewer:** Yeah. Any other information or issues you are facing in this phase?

231 **Participant:** No, only as I said I wish(--). Because I see that there are smarter ways to
232 save the data and record, and I think preserve them, but unfortunately, we didn't do that.
233 Or maybe we could have done it with some help and support.

234 **Interviewer:** Okay, thank you. Maybe now we can speak about Reporting and... it's
235 different than make it available. Like, how you publish the information about your data.

236 **Participant:** So, as I said, the papers are open access based on the European project.
237 So, everything should be public, open access. So, we... published the final results in
238 papers. And those are journal papers and my thesis, of course.

239 **Interviewer:** Okay. And it's also possible, up to the request, you can share the data with
240 the external people?

241 **Participant:** As I said, yes, upon request. And that I need to ask for permission from the,
242 again, *Trafikverket* for each single person. So, I think that I need to always get
243 permission for it.

244 **Interviewer:** Okay. Thank you very much. Maybe you can just... take a short look at the
245 notes that my colleague wrote and please try to tell us if there is something you need to
246 correct or change in these boxes.

247 (...)

248 **Participant:** Yeah, I think, in general, that's good.

249 **Interviewer:** Okay, we have another request now for you. If you look in the bottom you
250 can see three smileys there. Can you please use those smileys to describe how you feel
251 about these different phases? Like, are you satisfied, for example, with the planning or
252 you don't have any opinion, or you are dissatisfied. [titters]

253 **Participant:** So, for all six phases? I should–

254 **Interviewer:** Yeah, one by one. Like, for the Planning, are you happy about planning or
255 do you think you face (.) issues and you are not happy with it, something like that.

256 **Participant:** I think I go with “no comments” for the Planning.

257 **Interviewer:** Okay.

258 **Participant:** Because we could do what we could do [laughs]. We did what we could do
259 based on that, that’s why I think (.) no way to change it. The collected data, I think...
260 There, I’m a bit unhappy because... Because I think we could do it smarter, so that’s
261 one thing that I can say. Analyze the data, I think it was fun, so I’m happy.

262 **Interviewer:** Okay, good.

263 **Participant:** Yeah, and then Make it available... Mostly it’s extra work, so for me, which
264 I’m lazy, I go with I’m unhappy with that because, yeah... (.) And also Preserve, yeah,
265 probably no comment–

266 **Interviewer:** This is obvious, it will be nice if you be more honest about that. Because
267 you said you have some issues with the Preserving. [smiles]

268 **Participant:** Yeah, that’s why I have no comments about that because I think–

269 **Interviewer:** But... You don’t have comment or you are happy or unhappy?

270 **Participant:** No comments... Because they were buying hard disks, so...

271 **Interviewer:** Okay, thank you.

272 **Participant:** About my “make availability” I’ve been unhappiest, because I needed to do
273 some stuff there... (.) Work... Yeah, I’m happy about that because my thesis is published
274 now.

275 **Interviewer:** Great! Then... I think we are now approaching the end of the interview, but
276 we have one last request from you. Maybe by looking at the notes that you helped us to
277 write and these smileys that you put for every phase. Maybe you can describe for us
278 your... Like, dream scenario of how you deal with your data (.) in your (.) work.

279 **Participant:** So, I think this was an experience for me. If I go... If(--) With this knowledge
280 that I have now or the experience that I have now... If I go back (.) I would do it in
281 another way, of course. So, I could do it really much, much, much better because I know
282 that there are good tools available outside. And also, I could select (.) different sensors,
283 different companies, different providers, different acquisition system. Also, I would push
284 it to do the measurements not in the north, maybe somewhere close by, that I could go
285 and have better access. Or I could initiate a collaboration with the [university located up
286 north] which is closer to the location. I could have somebody there, which they do some
287 hard work which I didn't need to pay extra. I'm sure that they would have been happy
288 about that. Because, of course, you give them something free and that you want
289 something back from them, that would be good. I know that... As I said, also for storing
290 data, of course, I have seen these (.) cheaper things which are coming out nowadays.
291 Different class system and other things. I think there could be also useful because by
292 recording them, you can always have access to them. But hard disks that I need to
293 connect to my computer, look into and find. It's really tricky to find data in your hard disk
294 because if you are working from your office, then you need to take your hard disk, which
295 is heavy always. Or if you work from home, then it's a bit, you know, headache. All the
296 things I would say, yeah, maybe there are better ways to do it. And I would say, yeah,
297 this is my comment. I would have done it in a better way.

298 **Interviewer:** Yeah. Thank you again. And do you have anything else you want to add
299 before we end this interview?

300 **Participant:** Yeah. One thing that I realized during the interview, I think you mentioned
301 that [university research data support] do support. What I understood is that is just too
302 great. I think it would have been nice to somehow let the PhD's, researchers know about
303 this because I feel like we don't know about it. At least push (.) head of divisions to
304 mention that in division meetings because I think there, people they listen. If you put it
305 on a website probably less people read it. If you put it on somewhere that it's in the
306 website of... your website, I'm not sure if people read it because they don't know that it
307 exists.

308 **Interviewer:** You are right.

309 **Participant:** So, one way I would say because in a division meeting they normally
310 announced what is what, what is available, what is not. People and each division has it
311 every one month, so. It's good to know and "okay, you are a head of division, we are
312 going to give this kind of support" by listing some "okay, we have a space, we have this,
313 we have this, we do this kind of support with short description like in one or two
314 paragraphs, a few bullet points. Then they just copy it in their presentation they said
315 "okay, we have received this information from there" and then they do these kind of
316 things for us. Yeah, it's good if you do measurements if you play with the data, it's good
317 to call them and contact them, and get support. I would... I didn't know about many
318 things which I realized today, of course. I will let others know about this, but I think it
319 would have been nice if(--) I know that other people also have these issues in, at least,
320 in our division, I know that people are struggling with the data, storing it, recording it,
321 analyzing it. So, if such a good thing exists, I think it would have been nice that it would
322 have been shared in a better way.

323 **Interviewer:** Okay, thank you very much for the feedback. We are putting this
324 information in the website, as you said, maybe it's not readable by everyone. Also, we
325 have also newsletter, I think, every two weeks or one month. But this is a good note,
326 maybe in the meetings or the divisions. Yeah.

327 **Participant:** I think the best way is the head of divisions, because when we want to
328 even buy a simple coffee machine, they mention it in the division meeting. So, we know
329 that, "okay, there will be a coffee machine soon", or it's broken or whatever. So simple
330 things or really advanced, and we hear it there. So, I would suggest there is a good
331 place to share something really useful. Thank you so much.

332 **Interviewer:** We will transfer your notes [smiles]. Thank you very much for your time
333 and talking to us about your data.

334 **Participant:** Thank you.

335 **Interviewer:** And I will stop recording.