

1 **Interviewer:** Perfect. So, welcome. First of all, could you maybe just describe a bit your
2 research, and specifically about the data that you collect and/or generate?

3 **Participant:** I mean, it's a lot of different types of data that we collect and generate. For
4 example, we(--) when working with life cycle assessment we collect data from a lot of
5 different participants on the(--) maybe I'm going ahead, am I going away from the plan
6 phase?

7 **Interviewer:** No, just in general what type of research you do, we're not talking about
8 the *Vintergatan* yet, we're just--

9 **Participant:** Okay.

10 **Interviewer:** ...little bit, like give us a little bit of background about the research--

11 **Participant:** Okay, but so when doing life cycle assessment, we collect a lot of data
12 about emissions, energy efficiency, performance, materials used in different types of
13 processes, and that we try to collect from the different industry participants. Then we
14 sometimes do(--) Oh, I probably need to close the door here. Sorry.(--) and in in those
15 cases we get the output will then be the environmental performance data from different
16 from different processes and from a full life cycle and from different phases in the life
17 cycle. Then we have also, in another project we were talking analyzing electrification of
18 [vehicles] and then we want to collect data for [activity] [vehicle] activity types of
19 [vehicles] and when we also want to see in the [area] other electricity users and where
20 what we hope to generate from that type of study would be (..) Well, maybe there is not
21 so much output data, maybe that would more be a model.

22 **Interviewer:** Yeah.

23 **Participant:** Yeah.

24 **Interviewer:** Okay. This gives us a nice bit of background of, like, what type of research
25 is that you do. So, thank you for that. If we go look at the planning phase of, like(--) let's
26 just in your head take a typical research project maybe a research project that stands

27 out to you, when you in the planning phase, like how do you plan together your data,
28 and what works well in these phases, what kind of issues do you deal with?

29 **Participant:** I mean in the planning phase we are thinking of what type of data that we
30 need and who could have that data, I guess. And what we really(--) typically we struggle
31 with to get the data, the data that we want, the data in the right type(--) that is detailed
32 enough for what we want (..)

33 **Interviewer:** Yeah, so do you have any support in this, or do you have costs that are
34 associated with like, okay this is where we're going to get our data from?

35 **Participant:** We typically do I mean in most of my project we collect we try to collect the
36 data ourselves so we can we contact the stakeholders, we maybe have interviews or
37 questionnaires, or we could potentially even also purchase data from but we haven't
38 done that so far so much. So mainly we are doing the work and trying to collect the data
39 directly ourselves. Typically, you need to let me know if I'm not answering correctly.

40 **Interviewer:** Oh, well, I'll steer you in the direction, if necessary, but yeah, well, don't
41 worry.

42 [both laugh]

43 **Interviewer:** But you, so you have to make a list of, sort of like these participants, are
44 these individuals, are these companies, do you have a way to reach out to them, how
45 does this go in there, yeah (..)

46 **Participant:** I mean I would say that we typically probably start either from companies
47 or individuals if we already know individuals in some companies so it depends a little bit
48 on(--) I mean we have, I have two EU projects where we do similar type of data
49 collection and then we really start out with the consortium partners because they are the
50 ones that should have a lot of the primary data that we need, hopefully. But still, even if
51 that's the case and even if they are(--) we have a contact person, it's not typically
52 always easy to get the data from those companies, you know, anyway.

53 **Interviewer:** So knowing which people to contact is already(--) do you have a list of
54 these people or is it like there you have a list of people in your brain or is it, you know,
55 that you have a group or?

56 **Participant:** I mean, I think we build up a list, I would say, uh so it starts from probably
57 a list in my head, but when we will(--) we'll write it down who to contact, and if we
58 already know(--) have contact persons, I mean we already have established networks,
59 we write(--) that's like here we start from typically.

60 **Interviewer:** Yeah okay so that's all, to summarize it you(--) your data needs to be more
61 or less acquired over a network so you make a list of people within your network that
62 you think might have the data that you need, and then you start approaching them?

63 **Participant:** Yeah we also reach out outside our network if we feel that we there could
64 be(--) this company should have this data. I mean we typically expand our network as
65 well, when we need. Try to at least [laughs].

66 **Interviewer:** Yeah, how do you get outside of your own network?

67 **Participant:** I mean, I guess you can always email directly, even without any contacts.
68 It doesn't only work. You can see if you find people in your network that has contacts on
69 another company. So sometimes that's more effective. Yeah, I guess that's how we do
70 it.

71 **Interviewer:** Okay. Cool. Thank you, and then the data collection, and sort of(--) okay
72 so, but now we have like(--) we've gone through the planning phases, where we
73 mapped out how you decide where you're trying to go and get your data from, and then
74 you have your data collection process. How does this go? what do you do with the
75 data? Where do you get it from? Do you get any support or any costs? How does this
76 work for you?

77 **Participant:** We, I mean, we analyze the data, I would say. So we would(--) it will
78 depend a lot. So now I'm on the next phase(--)

79 **Interviewer:** Oh, we're still talking about data gathering. You have data. But before you
80 can analyze it. You reach out(--)

81 **Participant:** You mean the processing of the data? But I think we don't do a lot of
82 structured processing always of our data. Sometimes we do.

83 **Interviewer:** So what kind of data do you get? Do you get PDF files? Do you get Excel
84 sheets? What type of data?

85 **Participant:** Okay, okay. Sometimes of course we get(--) yeah sometimes maybe we
86 need to extract and structure our data. We have, in one project when we did interview
87 we mainly collected the data orally and wrote it down, so...

88 **Interviewer:** And where do you keep these notes before you could analyze it?

89 **Participant:** We keep it on our computers, in a folder [laughs] yeah, yeah, that's the
90 answer.

91 **Interviewer:** Yeah, okay but that is like, but you have a full(--) and is it this like, you
92 have backups or do you just only have the local folder?

93 **Participant:** I would say we typically always use one, right, so we have(--) we use a
94 backup system that is, well I trust that at least(--) I do that, maybe someone else in my
95 group have backups, but I don't make any backups. I hope it's safe [laughs], yeah.

96 **Interviewer:** You have one folder and whatever you get from the different companies
97 for example that you've reached out it just gets dumped into that folder?

98 **Participant:** Yes.

99 **Interviewer:** Okay this is good to know this is very good information for us to have,
100 because then you have your folder full of data and then, yeah, what do you do with this
101 data? Well, let me, sorry, let me backtrack what I wanted to ask. Like okay, you have
102 the collection, you've saved it in this folder. Does that work well for you, is this an
103 acceptable solution for you? Or are you not very happy with it?

104 **Participant:** I use the storing of data. I don't see a big problem with, as long as maybe
105 there can be that it's not safe but otherwise, from my purpose, I have access to it. I
106 know where I've stored it, I know where I have the data so it works.

107 **Interviewer:** Okay, okay, yeah, it's for some people it's like, this is the best we have but
108 I'm not very happy with it and so that is, but for other people it's like oh yeah this is the
109 best we have and I'm very happy with it, so...

110 **Participant:** Yeah, I'm okay with it. I think it depends a lot of how much data you collect
111 and maybe if you have a lot of data you collect, maybe you need more advanced
112 system of storage and handling, but yeah.

113 **Interviewer:** Have you ever encountered such a problem?

114 **Participant:** I'm trying to think of that (..) I mean, I have had problems when in(--) and I
115 should have learned better programs when I downloaded a lot of data from a database
116 which was(--) and went to an excel file which was too huge, large to be(--) work with in a
117 way which I had to give up, at that point, because I didn't have the time to(--) because it
118 was yeah fast so yeah. Then I had really big problems with storage capacity, it didn't
119 really work out well. It was so many data points.

120 **Interviewer:** Yeah.

121 **Participant:** It didn't work out well.

122 **Interviewer:** Can you give an aura of magnitude of how large this data set was?

123 **Participant:** It's some years ago, so I don't remember...

124 **Interviewer:** Are we talking like gigabytes, hundreds of gigabytes, terabytes?

125 **Participant:** I mean, it was still in an Excel file, so I still guess it(--) I'm not so sure, but it
126 was data from all countries. I think it was a lot of different types of statistics from all
127 countries and regions globally. So, it was a lot of data points, but I don't know. And it
128 was like human population, and it was also other things. (..) No, I don't remember.

129 **Interviewer:** Okay, perfect. Thank you. Then we go, I think, to the analysis, further.
130 How do you do your data analysis?

131 **Participant:** Yeah, of course, that also a lot depends on the different... (..)

132 **Interviewer:** Well, just think of the project and...

133 **Participant:** Yeah.

134 **Interviewer:** What is the research part where you really enjoyed doing the analysis?

135 **Participant:** (..) The analysis, okay, yeah, you're right, but it's also sometimes we also
136 generate our own data in a way, so we also analyze. And output data is sometimes
137 also, I guess, data from models, but maybe that's something else.

138 **Interviewer:** What type of models are you talking about here?

139 **Participant:** Yeah, we have global energy system models where we do(--) not
140 supermanning, but we do Monte Carlo runs with it. We get a lot of data from running it.
141 We don't have resources to run that, and it takes a long time, so we don't do, but say
142 maybe 500 runs. So, we have for varying different parameters, how the results changes
143 in(--) (..)

144 **Interviewer:** Would you like to do more of those Monte Carlo simulations or because
145 you say you are basically constrained by the resources that you have to do those?

146 **Participant:** I think I lost(--) I think you(--) some sound was lost.

147 **Interviewer:** I was asking if you're(--) would you like to do more of those Monte Carlo
148 simulations, because you eventually, you were a bit constrained by the resources you
149 had?

150 **Participant:** In that case we would really like to have the capability of running much
151 more runs. Then we are(--) don't have the right computer power for that in the group. So
152 there would be fantastic. But it takes(--) running this model takes about 10 minutes, and
153 if we want to run it, we would like to see if we can run it in parallel, so it goes faster, and
154 then run it maybe a thousand times.

155 **Interviewer:** Yeah, I see.

156 **Participant:** But that we haven't started work on, but we had a plan to try to do that.

157 **Interviewer:** Yeah. Okay.

158 **Participant:** Yeah, I'm seeing that I'm jumping a lot, because I probably work in a lot of
159 different projects, and some projects we don't have so much data. And so, okay, we
160 always have a lot of data, I guess, but it's more or less. But I was on analyzing of data
161 that we have in a way collected. If I take the example of when we are collecting data for
162 lifecycle assessments, when we typically collect data for a process, what we do in the
163 analysis phase would be that we compare that data to literature sources, see if it's so,
164 and maybe to different types (..) maybe to stoichiometric models or to checking so that
165 there is(--) input-output is a mass balance. So maybe we analyze it in that sense. And if
166 we, if we collect data from several different sources for the same process, which we
167 have not done so much, then we would, like, decide on which data which we think is
168 most suitable to use in our models and then put it in the life cycle assessment model.

169 **Interviewer:** Yeah.

170 **Participant:** That's one example.

171 **Interviewer:** Okay yeah, and for example if you have these Monte Carlo simulations
172 what type of analysis do you do then? Because I guess that also generates a lot of data.

173 **Participant:** Yeah since, when we have(--) I have not done it myself, but some of my
174 PhD students have gone on to call simulations also on lifecycle assessment and then
175 done it(--) looked at the range of the data so the median value and the 90th percentile of
176 these data set and see, I mean, how it could vary for when we have done this with our
177 energy system models, we have not run it so much that we have like statistic(--) we
178 cannot really analyze the data from a statistical(--) so we have mainly looked at how it
179 varies connected to different parameters. We have looked at the output values, how it
180 varies with the different types of input parameters. So more analyzing the trend than
181 actually done real statistical assessment of it.

182 **Interviewer:** Yeah. What kind of tools do you use for these analyses?

183 **Participant:** Excel. [laughs]

184 **Interviewer:** No? Alright.

185 **Participant:** Mainly, I think. And except from the, I mean, the PhD students use Python,
186 but I mainly use Excel.

187 **Interviewer:** Yeah. And are you happy with how this process is going, or?

188 **Participant:** No not super. I'm happy with some things I'm happy with this too, but I
189 know excel well myself, so I know how it works I'm happy with that. And if you don't
190 have too much data, it's fine. Otherwise(--) so yeah. In in my future life I'm saying that I
191 should learn some better tool. [laughs]

192 **Interviewer:** And so, the PhD students use Python, you said. And so how does this
193 work, like you know, one person working in excel, one person working in Python. Do
194 you encounter any issues, or not?

195 **Participant:** Of course, if they want help from me, it will much more difficult for me to
196 help them. That's a problem I guess and it's also a problem for me to analyze their
197 code, or if there's something wrong in, doing that. But otherwise, if we do it separately,
198 in different projects, it's not a problem in that sense.

199 **Interviewer:** Yeah, and can you get support for this?

200 **Participant:** I'm not aware of how to get support for that. I don't' think there is a way.

201 **Interviewer:** Okay, I think we then move towards the making available phase of the
202 data lifecycle. You have done your data analysis, and then you want to maybe make
203 this data available. So, the first question is: do you make your data available?

204 **Participant:** Yes, some data we have put up on Zendo [sic!], it's called, right?

205 **Interviewer:** Zenodo, you mean?

206 **Participant:** Zenodo, yes, thank you. I forgot it. The EU(--). And then other data I think
207 [inaudible] put on GitHub, or plan to put on GitHub. So yes, we try to make our data
208 available. Sometimes we just make our data available as supporting materials to our
209 articles.

210 **Interviewer:** Then it is hosted by the article website?

211 **Participant:** Yes.

212 **Interviewer:** Okay. How does this availability(--) Are you happy with this workflow? How
213 does it work for you?

214 **Participant:** I mean I must say that I'm not(--) I think it(--) if you want to share data or
215 models you need to work a lot of extra with it, to make it easy for other ones, other
216 people, to understand and work with. And we are not so used to and train with that, and
217 it takes extra resources. So maybe that's challenging. So what we did, like, in an EU
218 project we were just, we needed to make the data available so we just put it up on
219 Zenodo, but I'm not sure it would help anyone really, in a way. So, if it was useful for
220 anyone else, I'm not sure.

221 **Interviewer:** Yeah. Are you aware of any support that you might be able to get for this
222 or? Do we(--) have you used support?

223 **Participant:** We have not used support, no. I mean maybe there is(--) I mean I know
224 where there is some collaboration or initiative around open data and [university] but I'm
225 not sure if there is any support available i guess that we also have [infrastructure] but I
226 haven't really figured out if they have support for that.

227 **Interviewer:** Okay very good any other(--) are there also any costs associated with
228 your(--) with making the data available?

229 **Participant:** Yes, our resources I think, so our time.

230 **Interviewer:** Yeah okay.

231 **Participant:** ...as far as I know.

232 **Interviewer:** Yeah okay, and other have you encountered any other problems while
233 making the data available, or?

234 **Participant:** (...) I don't think so. Maybe I've forgotten something

235 **Interviewer:** Okay well if anything comes to mind then you can always, you know, make
236 it known or let others know. Of course, availability and preservation are different things.

237 **Participant:** Yes, I would say we have not worked with preservation really [laughs].

238 **Interviewer:** There is no data archiving?

239 **Participant:** [thinking] Not much at least. I mean probably we archive our own data on
240 our own computers, but not structured.

241 **Interviewer:** Okay and how does that work for you?

242 **Participant:** I mean, maybe it's inefficient probably but (..) but it's not, I mean, it's
243 always, I think, also sometimes maybe you need to reuse old data or understand all
244 data but sometimes you just move into new topics or new areas so you don't really need
245 preserving it in a way I don't know. Yeah. I haven't really, I haven't been thinking so
246 much about this. It's also like, in, I would say, like the data we collect and the data we
247 maybe generate, it get also old after a while so the question is how useful is this? I
248 mean, since we typically try to figure out either how is the process looking today, and
249 emissions today, or how could it look in the future. Our data is constantly depending on
250 that we in a way collect new data and we have connections with industry so... but of
251 course, we probably should be better on preserving data.

252 **Interviewer:** You say your data gets older, or gets out of date basically, but would it be
253 helpful for you if you can establish a timeline of things? If you have historical data to
254 compare the current data to or is that...

255 **Participant:** Perhaps, I mean, sometimes I think it could be helpful. I think we don't(--) I
256 mean we did, we collected when it was collecting data actually, it's from other literature,
257 from other articles, so we collected data from scientific literature and when it comes to
258 cost and efficiency of the [facility] and all that, I did the collection and everything in
259 Excel, and then we(--) I still have that stored on my computer, I have it stored for my
260 research group to see that, in a way, database that we built up. And then we did a new
261 research paper where we updated data, but then we didn't do as detailed data collection
262 with new literature, but we have that also in Excel, stored with new data, but we don't
263 synchronize them fully, I guess, yeah.

264 **Interviewer:** Yeah. Okay. And the reporting of data, unless you want to add anything
265 else about the preservation?

266 **Participant:** No [laughs] I'm feeling bad at that.

267 **Interviewer:** No no no, don't feel bad about it. It's like I(--) we really don't care, like,
268 what your process is, we just want to know what your process is looking(--)

269 **Participant:** No no, It's just my feeling(--) I don't say what you are saying, yeah.

270 **Interviewer:** Like the reporting, so you report on your data, how does it happen?

271 **Participant:** I mean, so I would report(--) I would say that this is typically happening in
272 scientific articles, or maybe you are aiming for something else?

273 **Interviewer:** No, so this is one way, like, okay, you have your data. How do you report
274 on the data? It can be articles, but it can also be something else. Maybe you write a
275 report for a European Union working group or other ways in which you communicate
276 your data to the outside world.

277 **Participant:** Yeah, yeah. Well sometimes we also write reports maybe, but mainly
278 articles.

279 **Interviewer:** In the articles, yeah. And are there any associated costs with that? Or do
280 you have support for that?

281 **Participant:** I mean, I would say it's the same. It's our time, but it's also really our, I
282 think it's part of our work to be a researcher to do this. And no, we don't have any other
283 support. I think it's part of our work.

284 **Interviewer:** Yeah. Okay. I think that is it there. Then we go to the smileys, like just the
285 planning phase. How happy are you with the planning phase? Is it happy, neutral, or a
286 little bit sad?

287 **Participant:** It depends, I mean just planning how to get the data, it's probably neutral.
288 Otherwise, it's typically like a planning phase but if it's just(--) that part of figuring out
289 what data you need, that's rather creative. So maybe it's in between neutral and happy?

290 **Interviewer:** You're gonna have to pick.

291 **Participant:** I have to pick.

292 **Interviewer:** Yes.

293 **Participant:** But then it's probably happy.

294 **Interviewer:** Happy, okay.

295 **Participant:** Even if just maybe(--) Yeah, we're very(--) practical things so it is maybe
296 neutral but oh, happy.

297 **Interviewer:** The data(--) sorry

298 **Participant:** The data collection, creation, it's maybe almost(--) yeah, I never feel sad
299 but it's not super fun to collect the data. Maybe it's neutral.

300 **Interviewer:** Okay, and the analysis phase, how do you feel about that?

301 **Participant:** Well, I like it even if I from(--) I could be better on it but I like it. I would say.

302 **Interviewer:** Okay.

303 **Participant:** Yes, so, I'm happy.

304 **Interviewer:** You're happy.

305 **Participant:** Make it available - you can put a sad smiley there. [nervous laugh] Maybe.
306 It seems like just [inaudible]

307 **Interviewer:** And the data preservation?

308 **Participant:** Yeah, you can put the sad smile as well [nervous laugh].

309 **Interviewer:** And the reporting?

310 **Participant:** I mean if it's in a scientific article, I like it. So if it(--) but if it's like, maybe
311 just from doing some extra additional work, maybe, it's neutral or sad. But if it(--) since
312 the main thing I do in reporting is an article, maybe it can be an happy face, I guess.

313 **Interviewer:** Okay perfect. And one more question. We've now gone through the
314 entire—

315 **Participant:** The dream

316 **Interviewer:** What is your dream scenario, if like, you know resources whatever were
317 no object, what would you(--) what would your dream scenario look like?

318 **Participant:** Yeah, my dream scenario is that I could plan what data I need and then I
319 just get it.

320 [both laugh]

321 **Participant:** And then(--) I mean, I am happy to analyze, but it would be nice to have
322 some help as well along the way. To collaborate with someone who is a super expert on
323 data analyzing, then data analysis would be nice.

324 **Interviewer:** Would you(--) Sorry, keep on, I'll ask my question at the end.

325 **Participant:** Okay and I hope I(--) someone else can went and then I can just 'okay,
326 please put it up', where I could tell someone else. I don't need to make it available,
327 someone else can do that. And preserve it, 'please someone else do that', and then(--)
328 And I still like to be involved in the reporting, but of course you maybe could have some
329 support there as well. In how to report(--) maybe there is(--) well because reporting is a
330 bit abstract for me here, since I see it mainly as an integrated part of something I do...

331 **Interviewer:** Yeah okay. This help with analysis would have to be someone who
332 understands your research area, or do they just have to know how to do statistical
333 analysis, et cetera?

334 **Participant:** That's a good question. It's probably good with some understanding of the
335 research area, but I think (..) yeah, it's probably helpful, but I think it could be good also
336 with just someone who(--) but maybe my limited understanding, just someone who's
337 good at doing the statistical analysis, if we can understand each other, maybe, maybe
338 not. Sure. But to some degree, I could do some analysis myself and then I could just

339 ask for help and maybe it could work with any expert, not just someone that know my
340 field. But I guess for some problems, maybe it will be easier.

341 **Interviewer:** Yeah okay, that is the dream scenario then. I think that's where we end the
342 recording. So, I will stop the recording.