Speaker 1: Thank you so much for making the time. I will first ask you a bit about your thoughts and experiences already with modular thinking and publishing. And we try to identify a few barriers for like why certain alternative ways of publishing are difficult for academics to use. And then at the end, I would like to think a bit about what kind of a tool we would need to make this a really helpful and effective tool for us academics to do so. But first, would you like to introduce yourself a little bit? What kind of research are you doing? And how did you come to be here today?

Speaker 0: Okay, so my name is XX. I am doing research in cryo-electron microscopy

Speaker 1: What was it that made you interested about today's topic?

Speaker 0: For me specifically, as I'm doing software development, I work a lot on GitHub and with one of my colleagues who is more of an IT technician in the group, but… All the software that I write is already open source publicly available on GitHub and also has the potential for other people to collaborate on. But what we anyways always do is when I make updates to the code, I make a so-called pull request in GitHub, which is essentially just a question like, I've made these changes to the code, which improved this functionality, and I would like to merge it with the main code repository. And then you can ask for a review of another programmer, in this case, the IT technician in our group. And he will go through my code and check it. And then we decide to merge it together. And vice versa as well. He has also made some updates to the code and I will check it. So I realized that at least for software development, this is already or it can already be a very open collaborative workflow. And I find it lacking in my field that there are still people that have fully closed source codes and I don't think it's a modern way of working on things and I think it can just help science as a whole a lot more if you have these sort of public spaces to do it, collaborate with people, because I also know people in the field who work on very similar stuff as I am doing, and I only recently just got to talk with them and work. If I would have started my research over, I would have, like, from the beginning, searched with people to collaborate with on these things. So, yeah, for me, during my PhD, I've, yeah, I've, I started thinking about it a whole lot differently because after I finished my masters I got educated in how academics work and in the start of my PhD as well. I think it can improve still a lot.

Speaker 1: Yeah, I feel that when I started it was also like, how do you do this differently?

Speaker 0: Yeah, exactly.

Speaker 1: So, thinking in a more modular way for you, it has, at least in your experience, some clear benefits from the possibility to collaborate much earlier. Are there any other benefits you see to such a more, well, timely approach of sharing the things you work on?

Speaker 0: Yes. Well, I think, for me, mainly collaboration would be, yeah, I see that as like the biggest benefit, I would say. But I also think that, yeah, we talked about it a bit during the workshop; having credibility for your work and having these DOIs and stuff, or DOIs, I don't know what is the normal term for it. But you can, yeah, I mean you can already at an early stage share IDs and work and have it indexable online and citable. So I think for that it's also a good idea too, I mean you don't need to keep all your research back, you can already put it in smaller steps online.

Speaker 1: Yeah, and I mean that wouldn't, right, not prevent a traditional paper at the end.

Speaker 0: Yeah, also.

Speaker 1: So there's a possibility of that.

Speaker 0: Yeah, I guess I also see the… Yeah, there's probably also some benefit at the end of some project, writing it up in one unit or trying to formulate it as the knowledge that you've sort of learned. Yeah, I guess that also still is, but it's good to also be able to publish earlier steps and be creditable for that. Or yeah, exactly, like teaching and stuff as well, let's say you have your ideas about the project and refine them a bit after reading literature, perhaps you make a presentation or something about it. If you can also put that online and link it together, that's also nice.

Speaker 1: Very much. And I mean, for me, teaching has been often more like presenting my research ideas in a presentation has been very helpful in like just, you know, understanding the things I want to talk about a bit better.

Speaker 0: Yeah.

Speaker 1: All of that is part of how a project becomes or like an idea at the end becomes a paper in a way. So when you want to publish, and this maybe comes also from me not studying or not doing research in software parts, software research, but if you publish things on GitHub, how do you present them then in a paper? Is there a traditional way of doing that?

Speaker 0: Well, I mean, usually the code that you publish on GitHub has a certain functionality and or introduces some new type of functionality. It can either be a new algorithm, right, or you can apply an existing algorithm to a specific use case. So, at least in my field what people usually just show in a paper is this algorithm applied to a data set and then it shows what the results of the algorithm are. But people don't really describe the algorithm in huge detail usually, which I think in some computing more, like fully computing science fields that is also done more often, that you really share an algorithm. Yeah, although some people do, if it's really something new, they do sometimes add some pseudocode, for example, to show what the steps in an algorithm are.

Speaker 1: So if I understand correctly, then in the paper you engage... So you publish, if you write code, you publish it on GitHub, but in the paper at the end you don't really engage with that, except for how it can be applied to a project.

Speaker 0: Yeah, that is mostly, so it would more be… well it's changing a bit I have to say, but I think it was always that people would more describe the mathematics behind the algorithm, or the theory. So basically the methods of the paper would then describe this theory and then the results would show what the implementation of this mathematics as a software, what sort of results this produces from the data, but then it wouldn't usually go in too much detail about the actual software, which is a bit of a loss sometimes, but I do think people are; it is changing a bit in our field, so people are now more, and I'm also doing this in the thing I'm writing at the moment, like, yeah, describing a bit more about what the specific implementation is and how certain things are handled, yeah.

Speaker 1: So these changes you're talking about, do you have an idea where they come from? Is it something like a difference in what kind of information readers would like to have, or is it something that comes from publishing houses a bit? Do you have a sense where that comes from?

Speaker 0: I think more of what we as a research community like to have, because I think beforehand the field was, or my field was a lot more, at least the algorithms in the field were more coming from, how to say it, like it was more a theoretical thing, like okay we can develop this theory and apply it, and also the people in the field just, because that's more recent, weren't so knowledgeable about computing science and software development and what is proper software engineering. So I think we see now that more people are coming in with a software engineering background. People are getting more familiar with that and I think that is why it's getting more attention in the papers.

Speaker 1: Is the change fast enough, do you think?

Speaker 0: [thinking] Yeah, not really. I don't think, if I look at where this is published, so usually the like high-end publications that get in nature, nature method stuff. There it's more the traditional way, but it's more if I look at the more niche journals that people are having more of this focus. But I don't think it's really... I think it should be valued more by the community, yeah. I don't think it's there yet.

Speaker 1: So I hear it's like there's a traditional way of sharing knowledge and it's very much like still, well, the standard by the main publishing houses?

Speaker 0: Yes.

Speaker 1: And if you would try something with a bit of a different approach, do you think you would get a chance to publish it in one of the bigger journals? Or would you really need to make a case for it or ?

Speaker 0: I think you do, yes. I don't think it currently would be accepted quickly there.

Speaker 1: Do you feel that Utrecht University is supportive in any way?

Speaker 0: Uh, that’s a … [thinking] I don't really know to be honest. Well, my professor is at least pretty supportive about it. But the university, yeah, not really sure. I haven't, I have to say, haven't been, because I'm, yeah, I mean, I'm finishing my PhD at the moment and like a lot of, yeah, well, almost.

Speaker 1: I feel that.

Speaker 0: Yeah, yeah, yeah.

Speaker 1: Last paper, you know.

Speaker 0: Yeah, yeah, exactly. Yeah, I mean, a lot of the thoughts that I've had have more recently been developing for me. I've more recently developed, so I haven't really sought to engage too much in it. But, I mean, yeah, this workshop came at the time that I was also thinking about it, so that was nice. But what I do at least know is that for the last paper that I wrote, I added a lot of… I tried to put all my data available on Dataverse.nl, which is a Dutch website for datasets. And you need to annotate everything. They were quite... Like I got invited for something because I... No, yeah, more people, but because I adhere to the fair data principles, like fair publishing, which is like a way of annotating and making everything indexable and having metadata, rich annotations. So I know at least that's something they are focused on.

Speaker 1: Interesting. So now I would like to think about a bit how a platform, like thinking about database.nl for example, and if you would think that a modular approach in your own work or in your own publishing, right, if you would want to do that, if you now think future. What kind of a platform would you like to use, something that would make? So I have worked with OSF, for example, and I find OSF, well, it is quite useful in a way.

Speaker 0: What is OSF?

Speaker 1: At least in Groningen, that's the pre-registration platform we have to use. And a lot of the social sciences, a lot of the publishing houses require to have like a registration of the project at the beginning on OSF, and then supplemental materials on OSF, and then... So it's a bit, the idea is that it is a bit modular, but in practice, for the majority of people that I talk to, and for myself as well, is like, you pre-register, then you ignore it, and then before you submit your paper, you fill everything out. So it's an extra step. The idea is great, but the step is extra. Or if you think of my reflexivity practices, in my field, that's something that you need to do, but it's something that it's enough to put these two sentences at the end, really. Rather than continuously reflect, we are supposed to do that, of course, and we all try to do that as much as possible, but there's no space for that in the paper. It's like something you write at the end and you tick the box. Right? So thinking about a way, for me this is very important, or like I'm passionate about in this project, is to find a way to create a tool for us academics that would make a modular approach to research and publishing, Well, accessible and effective, and not an add-on.

Speaker 0: Yes.

Speaker 1: That gives us more work, right? Yeah. We're all busy. So thinking about, for example, for you incorporating GitHub more into your publication practices, what would you need from a tool or a platform to make that possible for you? Also maybe in comparison to Dataverse, if that's helpful to think about?

Speaker 0: Yeah. I mean [thinking] Yeah, like I can just describe how I would see, like, an ideal publishing platform or something.

Speaker 1: Yes please.

Speaker 0: Because, I mean, it was also, well, of course, we also talked a bit about it today, but... Yeah, I mean, well, I like to be able to share steps in my research and also be available within this thing to discuss it so that others could potentially react, engage in a discussion publicly, to give feedback, for example, or reiterate on ideas. I think that would be really nice. Yeah, and it would be cool if you could do that in steps. So first of all, describe or make a thing where you introduce the aim of your project, or your initial aim, and then in steps try to update it, for example. I think during the research, for me, it's always pretty clear what are sort of building blocks within the research that I take. And I mean, I guess in software development, it can be quite easy because, for example, you write a specific function or module or something, and you test that module. And that is already something that you, for example, could put on such a publishing thing. Yeah, so I guess I could first of all see it that you sort of build it in steps. I guess at the end you would probably want to make something formal of it and also make a final version or something. I think that is something that is nice. Probably should have some sort of version control as well that after you have updates on it, yeah, that you can go back to previous versions and stuff. Yeah, I mean, yeah, I don't, I wouldn't like fully envision, I cannot fully envision what all the, how you would incorporate like all the steps of research. I think for a final product, it is easier than to go through versioning because once you have a sort of, or I mean, what now is like a draft of your paper, once you put that somewhere, it's quite easy to invite people to engage with it and then answer to these questions and maybe update it to a newer version. And with GitHub, for example, that could quite, I think, easily be linked because you might have a version of your manuscript and then have a new version of your code and you can update the manuscript to a newer version, which I think is something that could even continue, like now it's just these review processes are maybe half a year or something, which is already quite long actually, but I mean for code, it can easily be an ongoing effort that you release a new version and update your manuscript accordingly. I think that can all be linked. This is then a bit separate, I would say, from having, like, sharing the individual steps of building something. I don't know how you could link that really together. If we're talking about versioning of code and your manuscript, I think that is quite easily tied. Yeah.

Speaker 1: So I also understand you already have an idea of, like, how you build a project, of the different steps, would it be helpful for you to have a platform that ,where you are able, for example, to create a template and then when you have a new project, you can just open that template and go through the steps which are flexible and adaptable to your project and your needs. It would mean, well, you would have a starting point that guides you. Would that be something that'd be interesting to you?

Speaker 0: I mean, Yeah, I'm not really sure, maybe. I mean, I guess you would normally [thinking] adapt it a bit on what you are doing. I think it's most easy if you can just attach parts or modules together and sort of... I don't know. Yeah, I don't know if templates would... I think it's really... because it changes so much generally per thing you are doing. So, I don't know. Yeah.

Speaker 1: So, network view is something that you would appreciate to see how steps are linked.

Speaker 0: Yeah, for sure. Yeah.

Speaker 1: Do you have any kind of, let's say, ethical concerns or safety requirements for such a platform? Right? Picking up to what, whose name I unfortunately have forgotten, said about their worries on, well, someone stealing the research.

Speaker 0: Right. [thinking pause] Yeah, I mean, potentially I mean it's always it is for sure annoying if you if you put something out online and and someone runs off with it, however, I mean I think it's also a change in mentality ,because it's a scary thought, let's say like that, but I don't think it, it does not need to be that scary because it depends a lot, I think, on what sort of licenses you put on your thing. Because on coding then, of course, I use that as an example. But I mean, you always put a license on your code. And if someone uses it, they need to legally cite your code. Well, it depends on the license, of course. You can give a super free license and people can just do with it whatever they want. The company could even commercialize it, but if you put a more restrictive license on your code, people can use it, but they always need to quote where they get it from, and it cannot be commercialized. So I think for academic practices that is great. I mean, there's of course the thing that someone does use those ideas or develops them further, but I think that's also as a community. then a bit of, yeah, credibility thing. If everyone does it in this way, it's also a way to tell others, like, hey, you used this code, you should reference where it comes from. You can always, of course, go to people and tell them, like, I think you used my code, could you please reference it? I mean, that's more of an annoying thing to do. And people should just do that. Yeah, I don't know fully, but... Yeah, I think it doesn't need to be... And it can also have the benefit that you actually... That's the other side of it. It does have the benefit that if you then put something online, it is already your thing that you have on your name instead of... working on it in secret and having someone else work on it in secret and then right before you're about to put it online the others put it online first and then your whole thing is useless.

Speaker 1: So a more timely collaboration or a more timely information about what is going on would be very beneficial for you.

Speaker 0: I think so, I think for the whole research community, yeah.

Speaker 1: Let me see if I have some last questions. Ah, yeah. So you write mainly then for other academics in your field, right? So you have like less of a, you need to share it with other people who are not academics.

Speaker 0: Yeah.

Speaker 1: But if you share it with other academics, are there other ways you would like to present your output rather than a written presentation, for example? So if they're like, I don't know, would you like to present it as, during COVID we all had to do a video presentations of our stuff or there's the infamous and scary poster presentations or stuff like that, right? Like this, a normal presentation, right?

Speaker 0: Yeah.

Speaker 1: Are there like, is that something that you would, do you feel you would like to do more of or less of, which is perfectly fine as well?

Speaker 0: Yeah. Yeah, I mean, I'm not a huge fan of writing, so I mean, yeah, if I can present it in another way, that is definitely nice. Yeah.

Speaker 1: Do you feel that there's a shift as well? Do you think this is something that is recognized more, like the ways of presenting it? Or do you still feel like the written word still reigns?

Speaker 0: It definitely still reigns. But I do see people now online. This is not fully from research results, but I do see people now that online put a presentation for a course they might teach on Zenodo, which is, I guess you know this, right? It's also a place to archive data, code, and also put a citable link on it, identifier. And yeah, I know this one person, at least, who consistently always puts his presentations on Zenodo and tweets about it, like, hey, I gave this course, and here's my slides on Zenodo. So yeah, I do think there's people that try to do this more. But yeah, I don't know. It's definitely not the way to do it at the moment.

Speaker 1: Do you think it would be beneficial to do that more? Or like to enable such a... well, building a repertoire in a different way?

Speaker 0: Yeah, I think so, yeah. Like, especially for these kind of things, like I don't think researchers get that much credit for giving a poster presentation, a presentation at a conference, or teaching a course, right, or that sort of stuff. Yeah, if you can put that also there and have it citable, like you can, yeah. I think it's a good idea and it shows value in those things that you did.

Speaker 1: Alright, thank you so very much for answering all my questions and for sharing your insights.

Speaker 0: You're welcome.