Episode 5: The Future is Transparent: Moving towards Open Science

KEYWORDS

Open science, code, early career researchers, reproducibility, openness, people, data, diabetes, publish, tools

SPEAKERS

Gretchen Repasky

Luke Johnston

Ivo Grigorov

TRANSCRIPT

Gretchen Repasky 00:06

Many researchers are concerned about the possibility of public trust in science declining. Modern science is complex, and there are many examples of experts changing their assessments or reporting seemingly opposing results. But most scientists favor transparency, openness and reproducibility in their work. But the current academic research environment doesn't always reward such efforts. So what can be done? Is open science a path forward on improving the public's trust in science? Today, I get to sit down in the studio with two experts to learn more about open science and why researchers should embrace it. First, Luke Johnston is here. Luke recently transitioned from a clinical researcher postdoc position to a team leader role at the Steno Diabetes Center in Aarhus and Aarhus University in Denmark. Luke is an epidemiologist looking at the role of early life on diabetes risk. In recent years, his work has taken him to looking at the adoption of open science practices. And, Ivo Grigorov, a research coordinator at the National Institute of Aquatic Resources at the Technical University of Denmark in Copenhagen. Ivo has a keen interest in optimizing strategies for societal impact of research projects, by using open science to support open innovation, and public engagement in the research process. I'm your host, Gretchen Repasky. And this is Postdocs Talking, connecting diabetes and metabolism research to society.

Welcome to you both Luke and Ivo. Thanks so much for joining us today.

Luke Johnston 01:45

Thank you.

Ivo Grigorov 01:46

Thank you. Glad to be here.

Gretchen Repasky 01:48

Luke, let's start with the big picture. There's really been a slow but kind of growing movement to make science more open. Tell us about that.

Luke Johnston 01:56

Yeah, so like, I guess openness has been something that's been kind of a key feature of science from the beginning. But the tools to be making it more open, were not as sophisticated in the past. So now, we have a lot more tools to make things way more open that we can see the full research lifecycle. And it also enables us to be able to have the research more transparent and accountable, or like with a record, and it allows us to be able to verify the studies that are being published a lot better.

Gretchen Repasky 02:38

Ivo, why is something like this beneficial?

Ivo Grigorov 02:43

Openness, has kind of, it's rooted in making journals and research accessible to patients, right. This is very much should resonate with both this podcast and the audience. And it has trickled down from just doctors and patients having access to basic research into having the entire research process more transparent and more open and more accessible, and more importantly for us, as scientists, independently verifable. So it's extremely important if we want quality work and somehow trust downstream into that work.

Gretchen Repasky 03:25

Let's step now a little bit inside this bigger picture. Luke, can you tell us about your experiences as a researcher and how that led you to take a focus on open science?

Luke Johnston 03:36

Well, I originally started doing my masters and PhD in nutritional sciences, doing diabetes epidemiology. And I would say that's probably where it largely got started, because when I first came on board, in my group, the data was a little bit of a mess. And it took me a long time to be able to figure out. It was a mess to me. And also it wasn't super structured, and like trying to figure out what variables there were in the dataset, and all that stuff, took a considerable amount of

time to get organized and be usable in a consistent way, not just for myself, but also for others, because previously what other students would do it would they would just take the variables that would they would need and just focus on that. But I didn't want it just to be me. I wanted like the later students, or like others in the lab, to be able to have an easier time too. And I think that kind of started the whole thing. That, and learning how to code to do the analysis. Like we didn't really there was no training on it. There was no like we kind of shared code a little bit, but there was no...it was just sort of *ad hoc,* as it goes, and it felt a little disjointed. And, I knew that there was better ways of doing things. And that's where I started getting more into, okay, like, 'why can't the code be like...why can't we have code more open? Why can't we see like what people are doing? Why is it that methods paragraphs on like the statistical analysis, it was only like a paragraph or two long?' So I got really a lot more into that, at that point. And then I started transitioning a lot more and more, as I saw that it wasn't just my group at the University of Toronto, it was, like other groups, in other places. And you consistently hear the same story from PhD students and postdocs. The exact same problems that are coming up, and you say, 'Okay, this is actually a really big problem that we can potentially solve, or like, try to work towards fixing.'

Gretchen Repasky 05:47

Once you recognized this sort of practical problem, where are there resources you could go to, to try to build something that that was better?

Luke Johnston 05:56

Not really. I mean, there was a little bit, but it was kind of...it was more difficult to find, like online. I think it's only really been relatively in the last 10 years or so that online resources have become more of a common thing. And before, I think it was just more. It was more difficult to find, like central resources that you could go to. You'd have to really like know how to Google the exact right thing to be able to find the resource that you need. Because I mean, science is a small subset of the entire internet. And Google like searches by the most common. So because most people aren't searching for those things, Google doesn't necessarily list them as the top result results. So you have to go down a little and then you have to do some digging. So it took a lot of work.

Gretchen Repasky 06:45

And what you're talking about there is kind of an area in the broader scheme of open data, open code. Ivo, can you help us put that into the context of the of the research lifecycle, overall?

Ivo Grigorov 06:58

Sure. But first of all, I'm completely excited. Young researchers - why do you care so much? Why do you care that you can't dig around the behind the paper and find out all the components and reproduce it? The majority of the - just playing devil's advocate - but majority of students don't,

right? Majority of early career researchers look for getting that trophy at the end of the three years or whatever their project is. And how do you get a step on the tenure track? And that, today, does not involve open science? Why do you care so much?

Luke Johnston 07:32

I thought about this also, quite a lot. I could have done a lot of my PhD work so much faster, if I had just done it the way that everybody else was doing it.

Ivo Grigorov 07:43

That's an important intrinsic motivation, right. And, unless we match that with external rewards in the current system, we're never going to harness the energy. So, I can see how your frustration, initial frustration, drives you a little bit to crawl down the rabbit hole and dig deeper and find out how, how you can make the way we publish research more open. But to put it in a global context. I mean, why care? If you know papers are more reproducible, we could dig down all to the individual components of them? Well, if we zoom out of the individual struggles, or the disciplinary issues of like, you mentioned that there's lacking tools. Yes, there are many disciplines where intense research, intensive disciplines, where tools are the key bottleneck. But, if we zoom out of a discipline and zoom out to a helicopter view of research as a whole, tools is one bottleneck. Attitude is another; incentive is another. If you combine the two we get a very dynamic picture of why reproducibility's stumbling. And if we appreciate that, and then zoom back in into the individual career progression and the research workflow, where should the average young researcher from the average discipline care about openness? One could argue, at least from my biased perspective, is that it is in the single step of the research process that is repetitive and that consistently feeds you resources and allows you to do the research that drives you, right? So in a brutal and very monetoristic way, that's the grant. That's the grant, because you're going to be doing grants for the rest of your career as a young researcher whether you like it or not, right? So if that is something so repetitive, if the beautiful concepts that you are struggling with are relevant to that repetitive task, that's where the intrinsic, pure motivation is, you're talking about earlier, that drives you really to get traction. That's where the tire hits the road. That's where it comes from a beautiful concept to something that is a useful tool makes you competitive, makes you more reproducible. Eventually, it will feed into tenure. And we can discuss a long time.

Gretchen Repasky 10:36

Yeah, and we can come back to this. You're hitting on the how to incentivize. And I think this is what we could come back to in a few minutes. Do you want to do you want to tell us about the research lifecycle?

Ivo Grigorov 10:49

One of the first or most referred to papers was back in 2011. And it kind of sketched a basic lifecycle that might be relevant to most disciplines, right, and it starts with the light bulb in the head. I have an idea. And then I'm going to write it down and conceptually, and when and if I get the resources, it will lead to some data generation be it laboratory based or social science methodology, and that may create negative results, surprise, or positive results, that we can cherry pick and publish.

Luke Johnston 11:32

or null results

Ivo Grigorov 11:33

Or null our results, right, which we unfortunately, don't give enough credit to. And, then depending on the discipline, you may require a complicated simple or conceptual model to make sense of that data that will lead to your desirable *Nature* papers that will get your professorship and then that will open hopefully more questions than you answered, which will continue to loop and, eventually result in questionable practices like salami-slicing papers, which will feed a beautiful H-index, which will lead to an amazing career and a professorship title. Right? So that's kind of the pessimistic view, if you like, that's not to say that all science operates this bad.

Gretchen Repasky 12:24

And if we apply some openness to that research cycle, and we think of each of the steps along the way, how can science be more transparent?

Ivo Grigorov 12:33

If the podcast dies, and you have to walk away with one single message, as a listener, well, just do a little bit better than yesterday. And just don't be frustrated by lack of tools. There are always lack of tools. There's always lack of resources. Your supervisor or your team leader, he will always question everything that you're spending time on other than the next *Nature* paper. So do what you need to survive, but just do something more open than yesterday, and you're probably on the right track.

Luke Johnston 13:08

I think I agree with that. Like, if we were to define more like what exactly would be open, like all the components that would be open? You'd have - you'd pre-register your study beforehand. So that people can see like, what your protocol was, then, preferably, I mean, depends on the fields. I think, Ivo, you're in a nice field where it's easy to do open data, because you don't have to worry

about privacy and all that stuff. So having data being open, and then the analysis - once you do the analysis, then you have the analysis code there, publicly accessible, or like available to be scrutinized. And then eventually, you have the open article, the article that is provided...

Gretchen Repasky 13:59

open access.

Luke Johnston 14:00

Yeah, open access. Thank you. That's exactly. And as you were saying, like we want, like anything is better than nothing. So any little step is better than nothing. And at least for myself, I would say that one of the easiest things is having your code be open because it's like there's literally no barriers for us in biomedical field. Like we don't have to worry about privacy stuff because you know with data we it's more difficult to have stuff open. Pre-registering studies can be a little bit difficult to do. Because, like not always are studies, like within epidemiology, not always do the studies have a set, defined you know hypothesis or like testing environments like a clinical trial might. But sharing code is one of the easiest things you can possibly do, and it can help. It can start there and spreading at that point.

Gretchen Repasky 15:00

In these days, much of science, regardless of field of study is done on computers. We're generating and processing data. We're running statistical analyses on that data, and making figures of the results, writing up a report to present the results, and eventually disseminating the findings. However, scientists rarely have any formal training in some of these essential steps, such as writing code for analyses, and structural incentives for correct and transparent analyses are often lacking. Mistakes can be made and inadvertently leading researchers to even draw incorrect conclusions. A key pillar of science is the ability to verify claims made. And to do so, two things are essential - reproducibility, which means using the same data and same analysis, and replicability, which is using the same study design and analysis but collecting new data. Processes to assess or measure the reproducibility of studies don't yet exist widely. So we don't actually know the extent of the problem of non-reproducible research. Large projects, though, that check replicability, like the Many Labs Project, find that most studies are not replicable. So how does all this relate to society? Well, if we're producing non-verifiable research, we could be wasting public funds or time. Our society is driven by advances in science. Shouldn't that be science that can be verified?

Gretchen Repasky 16:36

How are you, in particular, working to change the system, for example, in sharing code?

Luke Johnston 16:41

I talk about it a lot. I talk about it a lot. Almost all of my work is, like, almost everything I do is on GitHub or GitLab. So you can like easily just go see, like, what I'm doing basically.

Gretchen Repasky 16:56

Tell, tell our listeners a bit about GitHub or GitLab.

Luke Johnston 17:00

They're both websites that host Git repositories. And Git repositories are a formal way of tracking your changes to files and folders. So like, if you make a change in texts, like for an analysis, and then you put that into the history of what was changed to the file, and then you can put that up online through GitHub and GitLab. And, it shows you basically the full history of what was done to all the files in a particular project. So you can really track, and I have like a been quite transparent about what was done to the project. It all started with the software engineering, largely. But now like people write books on Gits. With Gits they publish websites. You can do so many powerful things that are outside of software. I also teach R courses through the Danish Diabetes Academy. They're pretty popular, I would say. A lot of people really enjoy coming to them. So through there, like, I push not just teaching R but also like how to be more reproducible in your work, and then also how to share, eventually, because at least in the introduction course, we teach about Git and GitHub. And even though like sometimes in the feedback, we get people being like, oh, there wasn't enough, like, R being taught. And I'm like, well, that's, that's not the point. It's not just about R. It's about the reproducibility side of things, and it's difficult to be reproducible if you're not open. So it's kind of like, these are the tools that we're teaching you.

Gretchen Repasky 18:35

And so these kind of go sort of hand in hand,

Luke Johnston 18:37

Exactly.

Gretchen Repasky 18:38

with one another. If we have listeners who don't use R, or don't know what it is, tell us just briefly what it is and what people are using it for.

Luke Johnston 18:47

It's basically designed to do data analysis. It's largely used by researchers. The users - they're largely people who are doing research or in academia, but it's becoming more and more common in like, general data science. So data science is getting like quite popular lately, especially in industry. And in industry people are using either Python or R now. So like, basically, those two languages are dominant in data science. So, and the other thing I was gonna say is...it's an open source language. So that means that if you want to see how a linear regression was done, which is a statistical test or a statistical analysis, you can go in and look exactly how it was implemented. And that's part of the point of being verifiable is to be able to see - verifiable and transparent - is to be able to see how things were actually done.

Gretchen Repasky 19:38

Ivo though you mentioned a few minutes ago, about how important it is to do 10% more than what you might have done the previous day in regards to openness, and listening to what Luke is saying about talking about taking up practices and teaching others about how to take up practices...do you think it really is possible to change the system? Is it possible to do 10% more than you did yesterday?

Ivo Grigorov 20:07

It's definitely it's not a given. I mean, you got to find the 10% time right? Time is never enough. And early career researchers are, or you could argue, overworked, overburdened to pushed always to perform. But the only way you're going to create that extra 10%, or whatever you decide, is if you do a mind switch from "what you are asking me to do in terms of open is extra time and work", to the mind set where the 10% you're asking me to do is what I need to do to enhance what I do. And that is, that's literally mind yoga. And once you go for it, then the requirements on openness, whether they come from funders, or from your progressive team lead at an institution, or whether you're just gone to a conference looked around your peers, and they're doing it but you're not, then it's a peer pressure, external pressure doesn't feel like extra work, but it actually it looks like a pointer to where we should be going, right. And, really I mean, we can over state this: take baby steps. Don't break your neck going out to make the entire research workflow of your discipline, open in single go. Take the baby step. If yesterday, you only published an article and nothing underpinning it, tomorrow, publish an article and make one or two of the underlying outputs a little bit more open than just storing them on your laptop. And that's an important...that is you're entering the rabbit hole, and then you're never coming back.

Luke Johnston 21:54

If actually, if I can jump onto that, about the taking a little bit of time to become better, like the 10% time...there's actually examples where you take those time, like so there's a I can't remember the exact project that this was, but it was a group that does frequent measuring of ocean measures...

Ivo Grigorov 22:22

It's a beautiful piece of work. And actually, it stems from something that's very complicated, and it's called the Ocean Health Index.

Luke Johnston 22:34

That's it!

Ivo Grigorov 22:35

And, the lead author on that is - I highly encourage all listeners to follow her up.

Luke Johnston 22:39

Definitely agree.

Ivo Grigorov 22:40

Julie Lowndes, if I pronounced the name correctly. But her team is amazing. So what they tried to do something extremely complicated Ocean Health Index. Look at the entire ocean and devise a set of indicators that tell you what is the state of that patient. And that data was extremely, if I remember correctly, heterogenous. And it required a lot of data. So hard data science to crack, right? And even still now it's controversial in the ocean community. However, it's an important step, and all the lessons learned around that piece of research, around the original paper. I mean, the original paper was about how you crack this index about this complicated issue. Right? And that could be the ocean that could be a biomedical patient. I could replace your research question here, right. However, once you crack that paper, then that trickled down into a number of other papers, which I'm sure the, I haven't spoken to the authors, but I'm sure that team did not necessarily plan or anticipate, right. So now by using an approach which may have appeared time-consuming or not aligned with incentives, but because it was the right approach, then that trickled down into a number of other publications, which have gathered attention and caught your attention. And so that's a good example of how, if you bet on the right horse, you're gonna get a lot more windfall than you think.

Gretchen Repasky 24:14

Okay, so we've talked about mindset, the success of mindset, the success of taking the time, and how taking the time may not actually even amount to more time, but there must be challenges in changing the system. So can you talk a little bit about that? And maybe in particular, from the perspective of early career researchers, what challenges are they facing to change their mindset, to take the time?

Ivo Grigorov 24:42

Oh, we're back into carrots and sticks?

Gretchen Repasky 24:43

Yeah.

Luke Johnston 24:44

Yes, there's so many challenges. And sometimes it feels like overwhelming, like, you're just like, is that even worth it? And I think for many PhDs and postdocs, I think they, they see these challenges and they're like, it's not worth it. And then they leave for industry where they usually get paid a little bit better, probably quite a bit better. They probably have less like stress because they, you know, can finish their day and go back home. And they don't have to always worry about getting funding and getting like the pressure to be publishing more and more. So there's huge, huge barriers, and we're losing people because of those barriers, because of those challenges. And it's difficult because changing the system is so huge. It's literally a global thing, like universities or university rankings are based on some of these metrics that we have to then pursue, because then University metrics or university rankings are also a way for the universities to get students to come to them. And then so it's such a huge problem that it sometimes feels quite disheartening to try to tackle. And then that's like, at a global scale. But then at the smaller scale, like, you know, as Ivo had said, he, you know, we're overloaded, we're overwhelmed, and trying to find that time to like, do like, learn better tools, and then actually try to like push for change takes up a lot of mental energy, mental and emotional energy. When you're just like, you know, what, I'm just going to focus on, you know, doing what I do here and writing the stuff and, you know, appeasing my supervisor or research committee or whatever it is, it can sometimes be easier, and there's definitely nothing wrong with that. It's a very challenging thing for sure.

Gretchen Repasky 26:45

Is there a fear of making mistakes?

Luke Johnston 26:49

Oh yeah. For sure. I think there's a huge fear of, of making mistakes. But like many fears, they're not often justified. And I think that's the key thing is no, very, very rarely will you actually have somebody go in and look at the code and be like, Oh, you made a mistake there. If they if they actually went in and looked at the code and understood it, that's a huge bonus. I mean, that's something that is meaningful on its own. But yeah, I would say the fear of making mistakes is pretty big. Yeah, yeah, it's quite real.

Ivo Grigorov 27:20

And I bet that there is also another layer of fear of actually correcting some of those mistakes because of how the environment which we swim in might actually preceded the immediate supervision the collaborators why disclosing more than you should, when we could have squeezed in another paper out of it, right, or someone will steal our idea, the wind under us and those are probably fears that are more difficult to detect, measure and somehow talk about because the direct factor of people being just frank, as that senior researcher you mentioned. But they're an underlying driver against open advocacy. And we need to really capture somehow the good faith of senior researchers who mentor early career researchers in a discipline. And if the older generation doesn't believe in openness so much, then that is fine. But they shouldn't proactively put barriers on young researchers to find their own way of doing research which will underpin their career.

Gretchen Repasky 28:40

So we need the juniors and the seniors on board.

Ivo Grigorov 28:43

Absolutely, absolutely. The seniors can put important brakes on making research more open in an invisible way, they are in a position of power. Their approval is important. If you are uncertain, a young researcher, if you're looking to establish yourself, the opinion of the seniors does matter. So we need to get them on board. In a way they hold the young, the early career researchers' career progression in their hands, and some things they should proactively mentor and for some things, they should leave an open door for interpretation.

Gretchen Repasky 29:25

And if we flip that around, we could say that the seniors could promote in a very visible way.

Ivo Grigorov 29:33

Well, role model is great, but it doesn't pay the bills. But if we sidestepped from maybe the discipline that we talked about, now to psychology, there is a good trend there in at least a dozen universities from Canada to Australia to Germany in the EU where mentoring the next generation in open processes, open research workflows, open science as a whole, and it is actually a requirement to getting tenure.

Gretchen Repasky 30:16

If possible, I still wanted to bring us back to some of the sort of concrete tools or the different ways, especially Luke, that you're having your eye on. Because you mentioned GitHub. We

haven't talked yet, though, about reproducibility networks or repro hacks. Could you tell us a little bit about this?

Luke Johnston 30:35

The UK has been kind of pushing quite quickly ahead on this reproducibility topic. They have the UK Reproducibility Network. And then they have - I think it came from them - where they have these repro hacks where you can submit your paper or your analysis, and then they try to reproduce it. Like, it's kind of like a journal club style. Like people get together and they try to reproduce the study using the exact same data and the same code. So that's one way of like, independently checking whether your study is reproducible or not. So you can do that before publishing or something like that. I would say that's, like, probably the biggest thing that you could, you could potentially do, but even just publishing your code, no matter how ugly, no matter whether or not it's reproducible, it doesn't matter. Just getting that first step of getting comfortable with having your code shared, and seen and looked at is a huge first step. And then after that, everything sort of becomes a lot easier, because then as soon as you have the code shared, then other people can come on board and then, you know, look through it and check it, make changes or make suggestions or whatever it is. So that's a very first step, I would say.

Ivo Grigorov 31:58

And it's important to overcome that vulnerability or insecurity and, you know, 99 out of 100 cases, you will find that that rewards you and the fears that you had about it, disproportionate. But you know, just do a quick example, that even the best get it wrong in exactly that context, when you say, okay, my research is based on code. Do I hide it so I look good? Or do I expose it and risk embarrassing critique by my peers? Well, at least 10 years ago, there was a good case by NASA. NASA has a code for Earth system science, which basically, I might get the exact detail, right, if there's a system scientists listening, but stick with me for the for the general narrative. So if I remember correctly, the code is about modeling the surface temperature of the whole planet, right. And that's important in the climate change context and written by smart guys, you know, real coders. But coder scientists, there's only so much time they will spend to perfect a code. The code was functional, it was reproducing, living up to the reproducibility standards at the time. What happened is that in Cambridge, there was a engineering company full of competent coders who at some point had enough time to actually download the open NASA code and see if they can make it better. They were that competent, and that bored. And actually, what happened is that they took the open code. They looked at it line by line. They found that the even the experts have this insecurity, the metadata, the description, the information in that code wasn't always allowing somebody who didn't design it to run it, right. But being engineers, they stuck up to it. And they rewrote the original NASA code, which reproduced the NASA results 100%. I mean, if you plot the two curves on top of each other, you can't actually see the second one. But what they did find is that being nerdy software engineers, they actually reduced the code lines by 40%. Now, for those of you coding out there, you appreciate the number. That makes the code executable faster, less CPU time, in the modern of today's research, when you

have to write a proposal, you have to put significant funding into that kind of costs. So if your software running costs are 40% down, you're laughing, because that's 40% more salary for you, or going to the next conference. So that number is important while the reproducibility was not reduced, right. So if you think that in your area, you are some kind of a unicorn where you're not writing perfect research, well wake up. Even the experts get it wrong, and they get corrected. But the take home message is that if NASA did not make that code open, they wouldn't have harnessed the benefits. And I believe - let's check and point the listeners to the right source information - but I believe actually NASA did take up this rewritten code or important parts of it, because it was so more performance. So this is if you like, I'm going to zoom out with this example into something that's called engineering serendipity. Right? Everybody would love in their field, to use serendipity as a wave to surf and improve discoveries they make. Well, being open about everything in your next publication is one way, one first important step towards engineering serendipity in your collaborative network.

Gretchen Repasky 36:20

I think this really nicely leads us into thinking about the connection between science and society. Because if we think about a 40% savings of time, and the financial benefit of that, I think we can start to link actions in sharing code specifically, or then open science generally, into really concrete benefits for society. So Luke, would you like to expand on that a little bit? How these benefits can come to play for the broader community beyond scientists?

Luke Johnston 36:57

Yeah, beyond scientists, you know, at first glance, it kind of seems like it's more of a scientist problem, or issue that we have to face. But actually, because, you know, as I said, at the very beginning, about science is about verification, if a lot of the processes that we have aren't even really checking, verifying whether the science is being done, or whatever results the science is producing, then we're wasting a lot of time. And that's also a lot of time, like, so we're publishing stuff that might not actually be meaningful or real. So then it gets, you know, public hype. You know, you browse on Facebook, or Twitter or anything, and then you see, like, oh, this study says this, or this study says that. So if we had more verification, there would be less misinformation, potentially, or at least confusing information for people when they try to understand the science. And it also saves a lot of money if we can make things better from the beginning. So there's less burden on the taxpayers, less burden on society, and that way, at least in terms of money. And it also means that we get better science in less, in a shorter amount of time, which means that we have, you know, for instance, interventions like for public health interventions, coming sooner, which means that we have benefits of society sooner than what we might have originally. I think that's like the very basic level. Yeah.

Gretchen Repasky 38:40

What about any relationship to how the public trusts science?

Ivo Grigorov 38:44

I don't know what the latest literature or wisdom on the public perception of science is. But one thing is for sure, if you have a spark of public mistrust in science, and at the same time, you are less than 100% open and verifiable about your science, that only feeds the spark, right? So is openness, is open science, is everything we're trying to sell today the antidote to public trust in science? I don't know. But that maybe is not ultimately the question that is relevant to the individual early career researcher trying to get a foothold, right. The reverse is more interesting - that by making everything open, you're not feeding this poor trend of mistrust. And guess what, there'll be other surprises for you, too.

Luke Johnston 39:45

And actually, to just add on a little bit to that: when you're more open, you never know what your tools or whatever it is that you're putting out there will be used for. You have no idea. So if you put stuff out there, and then it's used in something that you would never have imagined in your wildest dreams or whatever it is, and that ends up being like, having a bigger impact on society or some like, niche segment of society, or whatever it is, that's a benefit on its own. And you can't necessarily predict something like that. But it wouldn't have happened if you didn't have this stuff out there.

Gretchen Repasky 40:29

So Luke and Ivo, let's draw our conversation to an end by actually trying to look forward. Where are we in 10 years? Is open science the future?

Luke Johnston 40:39

I definitely think it's the future. I mean, it's hard not to see. If you look at the programming just as an example, and software itself, the world essentially runs on open source software. And it has won for many reasons. And science is always like, in some ways we're at the cutting edge of what we're doing. But at the same time, we lag behind society in some ways, at least in technology. And it is going to come and it's like it's not a question of if, it's a question of when. I would hope that by 10 years, like it's way more open. I'm guessing that there's a high likelihood that like, at least when it comes to code, or at least more effort, when it comes to like data, I think there's definitely a likelihood that that's going to be a lot more open by 10 years, and then the tools around all of this whole ecosystem are just going to get better. So like, you know, one of the things that like, I try to think about when I'm like some of the projects that I'm on about making science more open is that open science will be the default, if it's easier. So the easiest it is, the more likely it'll go there. And I'm very certain that that's going to be the case with the tools that are coming up.

Gretchen Repasky 42:00

And it sounds like that's what you're working to make possible.

Luke Johnston

Yes, exactly.

Ivo Grigorov 42:03

Yeah. That's great. Great to have unicorns like this. We need to replicate you. But yeah, I mean, you know, from my point of view, I've now been advocating for this since I think 2006 is the first time I came across it. And for me, it's a no brainer. And open science is just science done right, as a friend of mine, coined a few years ago. It's just that somewhere along the way, because of skewed incentivized system, or pressure or environment, we kind of forgot that. It was never rewarded explicitly. It was we left it to the individual to drive in that direction if the individual chooses. And we lost track of that. And we'll have tracking that. And, I think the reproducibility debate is a consequence of that. And 5, 10 years from now, I think the major European funder is in place, the incentive system is being worked on at the moment by a major international consortium, DORA, who are getting funded also in Europe. And the bottom up individuals like you are driving the system from intrinsic point of view. So I think the ingredients are all in place, that five years from now, open science will be science, just science done right. And it will be closer to the incentive structure of having a career in research. And if you the listener, or a young PhD, just finishing I would argue that by the time you reach mid-career, you should, you will be asked to perform on these openness criteria. Just like the example in psychology departments today, right. So if there was one message to walk away with is don't wait. Don't wait. Don't stress out that you can't open everything in your research workflow. But don't wait. Just do one thing different today from what you did yesterday, and you're on the right path.

Gretchen Repasky 44:42

Thank you both very much for spending time with us today, for sharing these insights on the practices and the challenges, even, of sharing data and code or by extension open science, more broadly.

Luke Johnston 44:54

Thank you. Yeah, that's great.

Ivo Grigorov 44:57

Thank you very much

Gretchen Repasky 44:58

I wish you both all the best as you continue to promote these practices that foster both the community in science and also value in reproducibility. And, Ivo, your friend coined it beautifully "doing science right." Thousands of postdocs around the world are dedicating their careers to better understanding and improving diabetes prevention, care and treatment. You can learn more about Luke and Ivo, the guests of our show today, at our website, Danish Diabetes Academy.dk/podcasts, where we have short bios, additional information about open science and photos of them at work. Our show today was produced by the podcast agency Kontekst & Lyd. If you like what you've heard, please leave us a review. Reviews help new listeners to find us, and more people to take in all the great diabetes research our guests are doing. A very warm thank you to the Danish Diabetes Academy for keeping Postdocs Talking. Thanks for listening!