1:08

Alexey

**This week, we'll talk about teaching and mentoring in data analytics. We have a special guest today, Irina. Irina works as a teacher and curriculum developer for data analytics and data science at FrauenLoop, which is an NGO in Germany. Judging from the name, Frauen, it’s women.**

1:24

Irina

Yeah, and also AI build, actually as well.

1:28

Alexey

**Irina studied computer science in Romania and Germany. She has worked at a variety of tech companies in Berlin, Amsterdam, and the Bay Area. She's very active as a mentor in the area of data analytics and leadership. She is also an activist with a focus on supporting women in tech. So welcome, Irina. Welcome to our show.**

1:48

Irina

Thank you very much for having me, Alexey. I'm very happy to be here and to tell you all about these things. These are my moonlighting jobs. The nine to five is daylight and teaching and all of these things are my Moonlighting. But it's a great passion and important to me.

2:04

Alexey

**Okay. I also wanted to mention that the questions for today’s interview are prepared by Johanna Bayer. Thanks, Johanna, for your help. Let's get started.**

# Irina’s background

2:13

Alexey

**Before we go into our main topic of teaching and mentoring in data analytics, let's start with your background. Can you tell us about your career journey so far?**

2:22

Irina

Absolutely. How many minutes? [laughs]

2:26

Alexey

**Ten?**

2:27

Irina

[laughs] All right! That's more than enough.

2:29

Alexey

**Five? I don't know. Take your time. [chuckles]**

2:31

Irina

I come from Romania. I was born in ‘83 to a family of nerds. Luckily enough, I got to code already at age 12 because I got bored in my school and I wanted to see what the heck that computer is doing, which was Basic.

2:49

Alexey

**At what age?**

2:51

Irina

11- 12.

2:51

Alexey

**11. [nods]**

2:52

Irina

They were making us knit and I finished all the knitting and embroidery. I don't know what this is called, but you know what I mean. And I was like “I’m done with this. Now give that green screen computer. I want it.” Of course, we also grew up with mathematics books, and with our dad doing puzzles.

I always knew I wanted to go to mathematics and computer science after that, because I really liked it. Funny enough, my brother was always faster than me and he was like, “I'm not gonna study computer science.” “But I am. I have to be the quickest.” We had a very geeky childhood. My dad was a professor in AI and my mom was a database analyst.

3:38

Alexey

**That generation’s data analyst.**

3:41

Irina

Basically, yeah. Exactly. I always wanted to study computer science. This is what I chose. I did my Bachelor's in Romania and then I moved to Germany for my Master's. I wanted to do research and I wanted to do things that I knew could not be done in Romania at the time. The industry did not support that kind of direction. So I decided to do a Master's in research at Max Planck Germany, and then take it from there. This was my journey. I loved that.

I recommend studying at Max Planck to anyone – it's like a mini-Google with high impostor syndrome. I did one year and a half of PhD after that and I realized that this doesn't make me think. The teaching does and the students do – the classes do. But I like to work in something that has an impact in the real world and I felt like the PhD was not really connected to it. So that's how I landed in Berlin.

I started with a very generic role as a data consultant or engineer, and then moved slowly and slowly towards data – all in all four years, where I reached out the role of a BI manager. I was disappointed about the layoffs that year in DaWanda. And I remember, “Okay, let's change the city or the scene a little bit.” So that's why I took an offer from Google and I moved to Munich.

I was working in the advertising department in publishers, on a technical ladder, taking care of really large clients for Google. It was a normal job, nine to five [metal dropping] promoted so you have to find other projects that are aligned with your strengths and skills. I chose data, data logs, data models and forecasting – if you do this, then you get more money out of that. This is how I got into the whole workflows and pipelines. This was my most coolest project back then. Then I moved to San Francisco, because I really wanted to report to a female manager, who still works there in that org. There were not many female managers anyway.

6:14

Alexey

**That was at Google, right?**

6:16

Irina

Yeah. In the German offices, I don't remember female managers. But in the San Francisco office, there were. I have to say this because I think it is very important in terms of the percentage of women that are in tech – I reported to a woman manager for the first time at age 32. I was working in a team where they were something else than men at age 32. So we know that we have a gender gap present in Germany. I'm not going to continue more with Google. I took on other roles, did some research papers, and I moved from buy side and sell side – I wanted to know as much as possible that I could get from this particular org, add cloud on top of it, and then I left Google and kind of merged all my worlds into one.

From then on, I worked in the Netherlands as a consultant. Then I started, again, the managing path. Actually, the first team I managed was a team of data engineers. And I really liked them, because they did stand-ups early and I love a good early standup. A year now, I'm back in Berlin. I'm a manager of analytics, but I can also do data engineering, so I can wear very many hats. As I mentioned, I really liked teaching during my PhD. This is something that I kept investing in all the time after I left. That means going back to Romania, to my computer science high school and telling kids what they can do with it, encouraging women to study, or becoming part of the first lean-in group a couple of years ago in Berlin.

But when I got really serious about it, when my brother moved in with me, which was 10-11 years ago, and he saw that I was so happy with my job. He said, “What do you do, Irina? I don't know what you do.” He studied marketing. “Tell me. How do you come happy after work? What's that? I want to be part of it.” So yeah, he was my first student and it was pretty successful, I would say. That made me enroll myself as a mentor in other programs as well.

8:32

Alexey

**So your brother was doing marketing and he wanted to get… [cross-talk] Sorry?**

8:38

Irina

He was also doing a lot of things, like IT support. He's also a nerd, but in his own way. More on the hardware part.

8:45

Alexey

**So he was doing marketing and a lot of other stuff and he saw this world of data – and you helped him to get in data, right? So you mentored him.**

8:57

Irina

Yeah, teaching SQL, databases, queries…

8:59

Alexey

**And then he started working as a data analyst or data scientist?**

9:04

Irina

Yeah – a data analyst. I'm so proud of him because he said, “I have experience in IT support and I want to go into BI. You have two internships in both of these areas. Give them to me both and I will convince you to enroll me full time.” I’m so proud of his “sass”. [chuckles] It requires courage. He did a great job, so they hired him. Now, the rest is history because he's doing great.

# Irina as a mentor

9:30

Alexey

**Do you like the process of mentoring people?**

9:34

Irina

Yeah, because I understand how people think. For teaching him – it was a visual way of teaching him. I was trying to explain to him tables and databases, and how to do the connections between the models, and aggregation works – how to imagine that it's kind of collapsing when you have a window function.

9:53

Alexey

**That's the most difficult concept. I'm still struggling with that one.**

9:57

Irina

I drew. I like to go and then explain like, “If you do this, this is what happens.” Some people are not necessarily visual, so you have to find other ways. It is so interesting to understand them – how they understand the problem so that you can explain to them in a medium that is best for them. Some people like to study alone and work with you. This is kind of where it all started. Within Google, I was part of the official mentoring areas, as in like the Anita Borg Scholarship recipients from Eastern Europe, groups where we had, again, a really cool project, and Winning Ladies – they have actually created two NGOs that I've also volunteered for from time to time.

In the US, because it was easy due to the language to mentor, I’ve also worked with kids, and I worked with soon-to-be veterans going towards data, I’ve taught kids to code with Blockly and all the logic loops. Coming back to Europe, I discovered the Berlin Mentoring Club, which I thought was great. Both for me as a mentor and sometimes I also need advice, so I find people to help me there as well. That's when I'm focused more on teaching you how to be a data analyst – how to enter the data field. Of course, women prefer women, and there are not many women in data that are mentors. So it's a lot of requests.

Then I got into the bootcamps in Berlin, and I started analyzing the market a little bit, because I hired some analysts from two different schools and I saw some patterns. They learn this beautiful code, and they learn modeling, but they don't know some of the business – the gut feeling in business. So that's how I looked in, analyzed that part, and discovered FrauenLoop. And I'm absolutely in love with their community, because it's so positive. Also the fact that it's not for profit allows a different… more flexibility – more freedom in the curriculum, more tips, more stories from my own experience when somebody needs extra help. In the summer, I also designed the curriculum together with a couple of people at AI Guild, and put fraud detection and chargeback as well as a fancy class.

# Designing curriculum and program management at AI Guild

12:34

Alexey

**That's actually one of the questions I have about that. In your LinkedIn, you have this experience with the AI Guild that you mentioned. The title that you have for this job is Program Manager. The responsibilities, according to your profile, is designing the FinTech data science curriculum, collecting market standard practices, datasets, and exercises per topic, and finding teachers. I got really curious because in DataTalks.Club we also do courses and this is something I also need to do often. So I really wanted to ask him more about this.**

13:18

Irina

Right. My employer was AI Guild with Daniel and Chris – if you don’t know, it's a small community and we know each other. First I think I was asked to be a teacher and I said, I had a lot of ideas about the program itself and how one can update the curriculum a little bit to make it a bit more modern. And that's how I got to work together with Daniel.

Then I got inspired, “Maybe we should have some Google certification, cloud certifications. If we work in BigQuery and do anything data-related, we might as well get some certifications for those students.” So I pinged my connections with Google and got access to a learning platform, where if you are a Google customer, you can have access to it. So now, on top of everything, the students also get certification. So I think that's pretty unique in that sense. What are the questions you have around that?

14:21

Alexey

**The question I had was – you actually didn't answer that yet. You said that the AI Guild asked you to join them as a teacher, but you were more interested in updating the curriculum.**

14:35

Irina

They wanted to create a master class. I did both in the end.

14:37

Alexey

**Oh, you did both. Okay. I'm really curious, what does it mean to design a curriculum? What did you do there?**

14:47

Irina

Exactly. This I didn't do alone – it was also with Daniel – credit where credit is due.

14:53

Alexey

**So “you” plural, I guess. [chuckles]**

14:56

Irina

We…first of all, “I” work here, I was interviewing financial techies, so managers of data in FinTech companies, or in FinTech departments, or in fraud departments – what kind of methodologies they have implemented in their teams. This is an exhaustive list – I came up with a list like, money laundering, identity theft, identity management, and so on and so forth. Then I audited some Coursera classes to see what fraud classes they have, because there are classes on fraud – massive, massive classes, it could be a semester – but just to glean a couple of concepts from them.

What I decided specifically – the curriculum was fintech. After getting the high level topics, we decided to spend a couple of days on each topic, find the datasets to support it, think about complexity, try ourselves to actually – we coded the problems ourselves before putting out the curriculum, because “What if it's too hard?” As a manager, I need to understand the internships. I have to walk that path. For this curriculum, it was a company that wanted to have it – to hire these people. And we thought, “Well, since the company is paying for this pilot, we could have also added cloud certification on top, or added a little bit of business knowledge or notes related to why we are doing these types of models with lots of examples. What does it mean in commerce? What does it mean on dating? What does it mean in SaaS?” And so on.

Thinking about technical skills related to ML in production could be some parts of the curriculum and then alternated with the data science part. We also thought that it would probably be really good to give the students more business understanding skills. This is something that is so forgotten in our data-driven world – data analysts need soft skills: presentation skills, storytelling, slide design, all of that.

17:33

Alexey

**Design? Even that?**

17:36

Irina

You know what? I think at the end of the day, all of us, even if we are techies, we're also salespeople. We sell and we negotiate in every single interaction. We need to do that better in our tech world. In my Master's at Max Planck, I had a female coordinator (the program coordinator) who taught us technical writing and soft skills as well.

She used to film us, she used to tell us (actually, this was part of the pitch to get a PhD) “Tell the topic of your thesis in an elevator pitch. If you can convince me, then you get it.” [laughs] Then you get to film yourself, to see your body language – What does your face look like? Do you look at the audience? Do you not? Do you time yourself? We thought it would be a really cool idea to bring that.

18:27

Alexey

**You did this with the AI Guild as well, for the course. Did you?**

18:32

Irina

Yeah, I found somebody through – yeah.

18:37

Alexey

**That's cool.**

18:38

Irina

It was finding teachers – who could teach? Also, the Max Planck network has all of these AI teachers, Hasso Plattner, you can also ask them if they're available, they also have teaching industry classes. If you look at data storytelling and what kind of consultants offer your services on that – there are not very many. There are very few people.

I think the most prominent of them was one in the Netherlands, and a couple of them in the US, but not here in Germany. So that's why it's really hard to find teachers for that. The students said that they liked the class very much. They never expected to get such training.

19:22

Alexey

**I guess it's not what you expect from a data analytics course by default – or *any* data course. You expect technical skills. You don’t expect to learn how to design a slide. It's important, of course.**

19:35

Irina

Of course. How do you pitch your idea? How do you convince people? And the cloud part was an idea. I follow this guy from Microsoft – he designs ML exercises, writes ML books, and is very active. He was commenting on how current universities are a little bit behind on technology and that you're not going to find a class in cloud in a specific university (I don't know, MIT maybe. Maybe you can at MIT, so Stanford? I don't know.)

But in general, you don't find the practices of the real world being taught in universities. I think he posted something about an American bootcamp that said, like, “We have the trifecta – cloud, data science, and ML in production.” I was like, “Yep, that's where we want to go.”

20:34

Alexey

**Well, I guess if you're at MIT and you need to design a course, it will take a couple of years and then by the time you do this, AWS or Google Cloud are already different. Maybe they don't have the services that they used to have two years ago and they have something else. Then it means you need to redo the whole thing.**

20:55

Irina

And the UI has changed or the place where the information is. Yes.

20:59

Alexey

**That's why I guess the universities (like more “classical” universities) are always behind with these things.**

21:06

Irina

Because they're more reactive. They catch up to the market. How do you convince somebody who can teach you those things? Or for a university it might be expensive to come and teach those things. Maybe.

21:19

Alexey

**So your role as a program manager was to design a curriculum. And what that means is, coming up with a course plan, right? What are the modules? What are the units there?**

21:34

Irina

What are the topics underneath? What datasets could we have? What could we use as an exercise? So you have insurance, that means you have scoring. How do you do scoring? How do you find the data? Kaggle is amazing. God bless Kaggle. [laughs] There’s a lot of very good data.

# Other things Irina taught at AI Guild

21:48

Alexey

**So first, you talked to managers from different fintech companies who work with data, you collected some ideas from them to form the curriculum – a list of topics you need to cover – then you understood what exactly you need to cover there, what kind of datasets you could use, and then you found teachers for each of their modules. That's the process. Did you teach anything yourself in any of these modules?**

22:14

Irina

Yeah – fraud, and chargeback. Fraud detection as a rule-based was one example. And the other was neural networks for chargeback protection. I was trying to explain to the students that for fraud, you have a lot of signals and that's why you can use something else. You can even use rule-based – you have a lot of information. We can also discuss fraud changes. Fraud methodologies change every day, so if you have an ML model, we have to train it every week. That's another kettle of fish. On chargeback, you don't have as much information.

There were conversations with the students like “Think. How does it happen? What is the process? How many signals? How many signals does e-commerce see? How many signals do the bank applications see? Well, then if you don't have signals, and you didn't get anything that triggers it that’s within your power or your tracking, then yeah – neural networks. Because I don’t know where else to go.” I'm sure there are other models as well that could do better. Of course, we're not only talking about datasets, but how do you create your feature vectors? What are common standard practices to create that? From experience? From theory?

23:31

Alexey

**It was more data science, then.**

23:33

Irina

Yeah, I taught data science. BigQuery and data science.

23:36

Alexey

**Sometimes the lines between what data analytics and what data science are – the lines are a bit blurry.**

23:45

Irina

I’m happy to differentiate those definitions, because I was also teaching them a little bit.

23:47

Alexey

**But what you described, to me, sounds like data science – application of machine learning to a particular domain.**

23:52

Irina

Yes. Data science.

23:56

Alexey

**Okay. Data sets – so you talked to these domain experts (managers) right? So you… [cross-talk]**

24:03

Irina

Because I can look into a class, I can look into a book, I can look online on fraud and the categories for it. But that's not a direct view of how it is actually used in industry. If you have friends in industry, invite them for a coffee and ask them, “Have you implemented that?” It's good to know what's possible to do and what the standards are.

24:24

Alexey

**In your description of the job you were doing, it was “collecting market standard practices.” This is what you mean by that. Talking to people… [cross-talk]**

24:32

Irina

I kind of wrote it myself, but this is what I did. Yeah.

24:37

Alexey

**Okay. [cross-talk] Yeah, that makes perfect sense. It's very concise, but dense. We can talk about each of these words or phrases for the entire interview, I guess. [chuckles] That's quite interesting. So you worked as a data analyst, you worked as a BI manager, data activation specialist, BI manager again… [cross-talk] Sorry?**

25:02

Irina

At Google, I was a technical account manager.

25:05

Alexey

**Technical account manager. [nods]**

25:06

Irina

It’s an umbrella term actually, like” solution consultant”. It's an umbrella term that actually could fit every single title that I had.

# Why Irina likes teaching

25:18

Alexey

**And now you're a teacher. Right?**

25:23

Irina

I moonlight as a teacher, but I'm also manager of analytics during the day. Yeah.

25:27

Alexey

**So you have a full time job.**

25:31

Irina

I also have a full-time job, yes. Well, I've been interviewing in the last two months. So that's why you didn't see anything.

25:35

Alexey

**Because what I wanted to ask you is whether you're teaching full-time. But I guess the answer is no, or you have two full-time jobs?**

25:43

Irina

I don't have kids, right? [laughs] I’m free! [laughs] I think it depends very much. What I feel is that I can support the community in many ways, but I'm also being very mindful of my own energy as well. So for example, if I started a new job, I'm not going to teach in the first month or two, because it's too much. It's overwhelming. I will not be a full-time teacher, because I really love building things. I really like the idea of business and moving the needle and how engaging it is. I like the friction in a normal management job and teaching people there as well.

I also can say that sometimes I'm so positively surprised by the mentees or the students that I'm teaching (that might be even better than my own team) of one-two analysts. It's fascinating to see just how different the personalities are and the different speeds of desire to acquire new knowledge. I can explain that as well, like FrauenLoop, I started as well this year – first as a more of an SQL teacher, but we also looked into how we could update the curriculum with a few more modern concepts. For example, I had students or I had analysts reporting to me coming out of bootcamp, they didn't know what cohort analysis was, or churn, or retention – so the business KPIs.

I would also do that at FrauenLoop. And then, because you're an NGO, it's really hard to own tech – a tech stack on which you can teach. And I offered the opportunity to some of the students, if they want to, to work in my cloud instance for FrauenLoop, on my datasets. They can practice whenever they want their SQL or other things, or some BigQuery ML. So why not? You can write it directly in BigQuery. So yeah, I kind of also brought FrauenLoop. [laughs] People get scared of it, like, “Oh, my God. It must be something so big and complex (Google Cloud or any other cloud solution).” Then you show them that, actually, it’s really not. I'm only showing one part which is easy. I’m not going to complexify things.

# Students’ reluctance to learn cloud

28:11

Alexey

**To be frank, if I look at the Google Cloud interface, it's so much simpler than some of the other clouds. I mean, I use AWS at work.**

28:22

Irina

Everybody who is an engineer uses AWS.

28:25

Alexey

**Yeah. Then if I look at Google Cloud, some things are much more straightforward there. You kind of know how to find stuff, which is rare. [laughs] I mean, not everything. But sometimes, some things are a bit easier there. I also see that in students this… I don't know if “fear” is the right word.**

28:52

Irina

I would say reluctance.

28:54

Alexey

**Maybe to learn cloud, because this is something big and scary. [cross-talk] You don't need to learn the entire… I don't know if you can say “entire cloud” but – *all* of AWS, *all* of Google Cloud. There is one piece of that that you just can focus on and that's enough. Like BigQuery, right?**

29:13

Irina

At the beginning of the year, I was doing a Greenfield assignment. I learned from scratch and built a data warehouse from scratch. I mean, as long as the data sources are pretty simple, and you don't have a complex data ecosystem, heck, it's very easy to build pipelines, really. It's not rocket science. That was also really cool. I like building things. I don't think I could do just one job. I need to have a mix of things that always kind of excite me and bring new ideas.

# Irina as a manager

29:51

Alexey

**Which, for you, is a manager, right? You're a manager, and if you manage analysts you need to get into analytics – if you manage data engineers, you need to go to get into pipeline building and engineering. And then if you manage both, you kind of need to get into both, right?**

30:09

Irina

And product analytics and marketing both fall under BI. Architect maybe less – and ML engineer as well. Yeah, it's hard!

30:23

Alexey

**And then you also teach.**

30:25

Irina

Yes. But for example, I had an ML engineer, bless his heart – really talented. I think one day I read about him in the news. I cannot understand everything he does. I read a little bit, but I feel like I'm behind in some areas. That doesn't mean that it makes you a better or worse manager. If you conceptually understand things… if you conceptually understand an A/B test and how to do it, you're not going to do the distribution of the results as well. [chuckles] But again, to each their own. Yeah, you need to – because otherwise you cannot, especially with engineers, and especially as a woman, I feel that I need to prove that I’m technically savvy, maybe more than usual. That's my own [uncompleted]

# Cohort analysis in a nutshell

31:18

Alexey

**You mentioned one thing – you were working with somebody who didn't know what cohort analytics is or what churn is. I'm really wondering – I must admit, I also don't know what exactly cohort analytics is, maybe you can give it in a nutshell?**

31:35

Irina

It’s super visual. Imagine that you… Let's talk about a game – online gaming – Pokemon, Diablo, etc. I play Diablo. I love Diablo. [chuckles]

31:48

Alexey

**Which one? First? Second?**

31:50

Irina

The first and second I've exhausted them to the end together with my brother in the private LAN in our block of flats in Romania. But now I'm playing the immortal – the online version. Good memories. But imagine that you're launching a game and you know that there are parts of the game that are in different sets of beta, or maybe you don't have as many journeys. And depending on what users find in that game, they're more or less active – if they find the things interesting. So actually, as a product manager of a game, or owner of a game, what you want is that your users spend as much time on your platform as possible. You're interested that they return, and that they return after the first month and second month, and you want a long relationship with that client. I'm trying to explain how the cohort analysis will look.

Imagine that in the first month of launch, you have people being super active because it just launched and everyone wants to try it and the servers are super busy. The engineers are focused on that and not on new maps and new challenges. As a result, let's say that after the initial hype of two, three months, they don't return because there's not enough complexity in the game. Now, the PM says, “Well, now I have more time to think about what extra things I can bring into the game. Let's make it longer. Let's make it more complex. Let’s make it more social, with more challenges.” The next set of users that sign up and start playing the game have a different experience than the first ones. As a result, they have different behaviors. And as a result, they have a different way of coming back. Or they come back for things that the first group did not come back for. So what you do is count the users who joined in the first month and then the second month. And then you calculate what percentage of that first group returns the second, fourth, and so on. If you do it with a very nice visual, you will see a gradation of color and how that retention cohort analysis looks – the retention rates.

34:03

Alexey

**The cohort here is not a cluster, it's something else.**

34:07

Irina

Exactly.

34:08

Alexey

**Okay. Because what I thought that cohort analysis is when you have different clusters of users and you have different profiles – different personas.**

34:17

Irina

You can also do that. You can use it for the same month, different users – like an A/B test. You want to see if the test is impacting something. Or maybe you have users who see or engage with ads to get credits. And users who play should know that – users like me. [chuckles]

You can also create two cohorts and look at, let's say, not at the retention rate, but maybe the monetization rate or something else related to that behavior – being exposed to something. It can be chronological like month after month after month, or it can be at the same time and more like group comparison – persona comparison.

# How Irina started teaching formally

35:00

Alexey

**Yeah, thanks. I'm really curious, how did you start with teaching? You said you mentored your brother and that you liked it. But how did you start doing this more formally? You also did this with the AI Guild. You were doing this with FrauenLoop. How did you find these companies? How did you start mentoring with them? How did it happen to you?**

35:28

Irina

Yeah, it happened. [chuckles]

35:30

Alexey

**It just happened? Was it intentional from your side? Or it just happened?**

35:34

Irina

With AI Guild it really just happened. I got an email from Chris, like, “Hey, we saw your profile. You are FinTech, you are data science, you’re in AI Guild. Come teach for us?”

35:42

Alexey

**Well, it didn't just happen. They reached out.**

35:45

Irina

They reached out. [cross-talk]

35:48

Alexey

**[chuckles] You didn't actively do anything. [Irina agrees] You just did your stuff and then they reached out.**

35:57

Irina

Yeah, I think they were reaching out to data specialists who were working in FinTech at the time, or had long, extensive experience of marketing data – all sorts of data. I told Chris, “Hey I actually have more time than just this, if you want. We could build more.” That worked. With FrauenLoop, I kept hearing about the NG from so many people. And then one of my hires in January of this year was a graduate of FrauenLoop. She kept saying so many good things about it that I decided to join, like, “Hey. Can I offer you my services as a mentor or teacher? I would love to. I have some extra time if you need that.” Apparently, I came at the right time because there was a need for mentors. So I started right away.

I think this is something that for your gender is harder to understand. I've been in conferences where there are only women – I’ve been in classes where there are only women. It has a different vibe. It's a bit more jolly. It's a bit more relaxed. It's really nice. You can discuss more honestly, as well, like “Hey, you are selling yourself short. You have apologized for some things 10 times and I'm going to put an apology jar – I'm going to make you buy coffee or chocolate. Because you don't need to apologize this much. This is not a reason to apologize. This is just tech work.” So it allows for more connection and more honest conversations as well, like “Don't do that. Don't do that when you interview. Trust me. That not It's not how you should approach things.” Yeah. And I hope for a long collaboration with FrauenLoop. We're always thinking of things.

# Irina’s diversity project in the works

38:07

Alexey

**From what I know, mostly from Johanna, who talked to you before this interview – she mentioned that you work on some projects where you are looking into gender diversity in big companies.**

38:25

Irina

It's not just me. It's a group of people.

38:27

Alexey

**Yeah. “You” as a part of a larger group.**

38:33

Irina

It’s really at the beginning. We're just really excited.

38:39

Alexey

**“At the beginning,” does that mean that you are not ready to talk about the initiatives you have because they are not concrete yet?**

38:49

Irina

Yeah, they're not. The signatures are official. But I would like to keep things under wraps just a bit longer until we get the first [inaudible]. But it's not some crazy new idea at all. It’s all about the data engineering question, because I've been thinking about grabbing LinkedIn data or Glass Door data, and analyzing it to see if I find things. As a data engineer, I was looking into how I can get this data and what are the right things – I need a developer account, I need to pay for this, I need to have IPs to scrape things. So it took a long time to find efficient solutions. I mean, I'm a data manager and I'm a data teacher. There's also the concept of data governance and legalities.

For research, you cannot use information that is not publicly available or you have manipulated to get. This also was a problem to overcome. Otherwise everyone can say “I’ve developed something.” It has to be respectful and legal as well. I mean, I would love to do research on par with lean and stuff. I would really like that because I think there are so many pitfalls and we don't know, as women, every time we take a job, whether we're gonna work in a team with other women or not. Maybe we're going to be the only woman in the team. Sometimes being the only woman in the team is not easy. Hopefully, with such work, we can see what companies are open to diversity, and that diversity is not at the lower level, only at the higher level, which shows you that, “I have a model there who can stand up for me.”

40:43

Alexey

**Like managers, right? You mentioned the first female manager when you were how old?**

40:48

Irina

I was 32. That's insane. [inaudible] I don't know,

40:55

Alexey

**You had to move to another country to have this experience, right?**

40:59

Irina

I did some internships in Romania. But I genuinely do not remember. This was long ago. I don't know if I had a male or female manager. [chuckles] I only remember the project. That's all. [chuckles]

# How DataTalks.Club can attract more female students to the Zoomcamps

41:16

Alexey

**This project that you mentioned is at the inception stage now – there is nothing concrete. But we have a course – at DataTalks.Club, we have a course about data engineering and it looks to me that roughly 95% of the students (it feels like) they're males not females. Only a small fraction of the students are females.**

**Maybe you can give us a few suggestions on how we can make it more interesting, more attractive, for females to also join our course. This is a free course. How can we attract more females, more students, to join the course? The marketing we did so far was just some posts on Reddit, some posts on Twitter, LinkedIn and people come. That's pretty much the extent.**

42:10

Irina

How many women hang out on Reddit?

42:12

Alexey

**Zero, I guess? [laughs]**

42:14

Irina

I mean, I do. But who else? [laughs]

42:16

Alexey

**[laughs] It's like a super toxic website, right? Because it's anonymous.**

42:22

Irina

It is weird. It's a trash fire sometimes. It's a fun trash fire sometimes. As a parenthesis, I love the “am I the asshole” section because I always laugh when I read things from there. Yeah. How else can you advertise? You can put posters in universities. You can…

42:45

Alexey

**But maybe not general advertising, but like how to make it more attractive for females? Because I guess if we go and put this in universities, then again, it’ll be the same…**

42:55

Irina

Just saying. No, it has to be more targeted.

42:59

Alexey

**Like maybe in groups like PyLadies. That could be useful?**

43:04

Irina

PyLadies would be one. I think there's a lot of fear of data engineering, to be honest. That's a problem with signups. Give me 30 seconds to brainstorm.

43:20

Alexey

**Yeah, you were not prepared for that. That was not a part of the questions. [chuckles]**

43:29

Irina

That’s okay. Happens all the time. There are these great organizations, like CorrelAid, Citizens for Europe – they also offer classes. Maybe you can offer classes through the volunteers so that they can upskill. I've been to a couple of hackathons where we have a hard time on the data engineering part, not from the statistical point of view. There were very few people with data engineering experience. I can think of Ellen Konig, who was the only female and the only data engineer expert in the room. So it can be training volunteers, going to school things and hackathons.

Also, these organizations, as far as I've seen, have a slightly equal representation of genders, surprisingly. I saw men hacking for diversity at CorrelAid, which is cool. Women have a bit of a different obstacle than guys, in terms of who your target audience is. Do you want to target the moms? It’s the hardest for moms to find time to learn something, so you have to do classes in the evening. If you want to target PhD students, then you can advertise this class through alumni or Max Planck or programs or… I'm sure for example, it can be complementary to refugees, as well. This could also be things to pair. Just let me know.

45:10

Alexey

**Great ideas! [chuckles]**

45:11

Irina

Do you know the Gallup StrengthsFinder Test? [Alexey agrees]. My number one is “activator”.

# How to get technical feedback at work

45:24

Alexey

**If you come up with great ideas, please do share them with me. But I realized that we have six questions in Slido and we should cover them too. [Irina agrees]. The first question – the most important one – is, “How can I get technical feedback on my deliverables at work? My manager either doesn't have the technical skills or the time for code review or discussing data modeling.”**

45:50

Irina

Right. I would say, in that case, you have to make your manager aware that both your work quality and happiness would benefit from this. And if the team doesn't have implemented processes for code reviews, or ticket reviews… [cross-talk]

46:07

Alexey

**Would you just say it like that? “My happiness depends on feedback.” Would you be upfront about that without trying to say it in any other way?**

46:16

Irina

I’m giving large advice. The first thing I would do if somebody is very sad about this and really wants to have code reviews is find things internally. Don't escalate, God forbid. You don't need to for this. The purpose of getting technical feedback or reviews of ML models is very vital for the company. So yes, somebody should take a look at your code. As your manager, I would not let your code go live without somebody else seeing it, because it's important.

If you do not have peers in the analytics team, there can be peers in the engineering team that can give you feedback on the engineering part of the work. Or there can be a financial analyst who can give you some feedback on the logic that you have approached. If you are lucky and work for a stakeholder who is technical, they can also maybe replicate the work, or the stakeholder could be the reviewer as well. For example, the ML engineer cannot go with a code review to an analyst, they still need to go to another engineer for that setup.

So I highly encourage that, first of all, that you convince the managers – something that they can be solving by themselves. Find the right person, talk to more people. Engineers love to help with these things. I don't have any other experience. This will be appreciated and any flexibility around these approaches will also be appreciated. If you don't find any resources within your company, then it's tougher. Because ML models or other information can be very proprietary to the company and you cannot share outside – please don't. But then you can find other data and kind of have a similar approach and kind of have similar code, but please don't share it because people get fired if they do. Don't share company code on Reddit. Just code it again and ask the community that you're part of.

48:30

Alexey

**Speaking of discussing data modeling, I guess in situations when your manager doesn't have the skills, then it's tricky. But maybe some other people do have time. For example, in our company, when somebody wants to discuss data modeling or architecture or something and they come to me, I'm super happy. I will be glad to put everything I work on.**

**Maybe I won’t help right there on the spot, but I will wrap up for the day, and then tomorrow I’ll be ready to discuss this thing. Because it's much more exciting than what I usually do in my day-to-day work. Data modeling means that you're probably starting something new and you need to come up with the right model for that or right architecture – right something.**

49:20

Irina

When somebody starts explaining this to you, like how they approach this knowing they have a lightbulb moment – the process of explaining something and taking it out of your head and putting it in words or on paper helps you clarify things too. Actually, this is one of the benefits of mentoring as well.

# Antipatterns and overrated/overhyped topics in data analytics

49:39

Alexey

**Okay. Another question, “What are the top five overrated things (antipatterns) and waste of productivity for people who learn data analytics?”**

49:51

Irina

Don't learn models if you don't apply anything.

49:54

Alexey

**By “models” you mean…**

49:57

Irina

Predictive ML. Don't learn ML if you don't have a project for it. This is something I see a lot. Also – overengineering, I call it. A little bit of overengineering. [pause] I'm trying to kind of bullet point my answers. On one hand there’s too much ML hype without being fully aware if that is necessary or that model is applicable to the data.

50:33

Alexey

**Data analysts don't necessarily need machine learning, right? If you're a data scientist, maybe you do need it, because otherwise you will not pass the interview. They will ask you about ML. But for data analysts, it’s not a must. Is it?**

50:45

Irina

Exactly. Then some other things can be antipatterns and things that are not recommended – overengineering is the second answer. This is something I've learned at Google. Ship it. Ship it and iterate. Find the first MVP that works for you and then iterate on it. Don't work in a waterfall model, like for a PhD and wait for failure after five years. [chuckles] Iterate quickly. If it does good – great.

If it's a model out of the box, if it’s a solution out of the box – and then later improve it to something else – great, fine! I've seen this kind of mentality in a couple of Dutch companies, where you take a model out of the box, like a recommendation system, and say “I can do better.” “Okay, well. First we have this. It makes money – and then we build on top.”

51:33

Alexey

**I think all the engineers are guilty of that.**

51:40

Irina

I'm also guilty of that in some parts. I know. I think there's a lot of focus on modeling instead of the business. There are low-hanging fruit with low tech too. If you think about it.

51:54

Alexey

**I guess if you want to get into analytics because you really like these shiny things – you might like machine learning, you might like these things – and you're saying now that “You know what? Actually, these things are not needed there.” Not always, right? Not all the time.**

52:14

Irina

Not all the time. Yeah. For example, I built a model that would estimate how much money you would lose in the tool AdExchanger if you do not block certain categories of keywords. And it was still a little bit like this. I mean, we had a lot of data to do the prediction on, but it was not a complex model. It was not. I think it was a simple regression or some clustering with something. Super light.

52:40

Alexey

**That’s still machine learning – regression. [chuckles]**

52:44

Irina

Yeah, but I don't really consider it that level, right? In the end, you know what? It got incorporated into the product. And when they found another analyst to build them a better model for that, they could incorporate that one. It’s okay. “Best of”. Other hype-ey things are tools. That's why you rarely see me talk about tools. Because I think it's like, you learn to fly a plane – it doesn't matter. Or you learn to drive a car – it doesn't matter which car, if you know how to drive it.

53:23

Alexey

**Maybe with planes, it's a little bit different. [chuckles]**

53:26

Irina

Yeah, I know. That's why I switched to cars as an example.

53:28

Alexey

**Maybe bicycles are an even better example. [chuckles]**

53:32

Irina

Tools come and go – the solutions that existed 12 years ago for BI, for ETL, and for visualizations are not the same.

53:44

Alexey

**I think Microsoft Integration Studio is still there. [cross-talk]**

53:49

Irina

[inaudible]is not, I think. [laughs]

53:51

Alexey

**Yeah. Most of them, I think. Right?**

53:53

Irina

I used to work with Pentaho, which used to be open source and for free.

53:38

Alexey

**Is it still around?**

54:00

Irina

It's still around, but it has been rebranded as Sisense.

54:07

Alexey

**I think I remember when I studied BI as my Master’s and Pentaho seemed like a big thing back then. But when I graduated, nobody seemed to care about that, actually. [chuckles]**

54:21

Irina

I mean, that was like one or two years when it was big and it was a tool of choice.

54:30

Alexey

**Basically, if we talk about all this Pentaho and other ETLs [cross-talk] you have a concept (which is ETL) and all these tools, they're kind of secondary, right? That's what you're saying?**

54:40

Irina

Yes. I don't see the point of using DBT if I can do my orchestration myself.

# Advice for young women who want to get into data science/engineering

54:46

Alexey

**We have a few more questions. You probably have a lot more antipatterns, but the next one is interesting. “I am making a career transition for data science. The background is in energy sales (GIS) teaching assistant. I was told that I won’t become an engineer. Do you have any advice for young women?”**

55:09

Irina

Oh boy. [laughs] Tell me who told you that. I can come myself [shakes fists] [laughs] No such thing. First of all, I would say that this person who chooses data science from a bit of a less technical role has a longer road of learning to go through than the classic data analyst because you still have to understand a little bit on the business end. It's not just the model, it's also the business model – how the money is made.

Data scientist – you can totally become a data scientist. Even if, let's say, that you have a role in sales at this point, or in finance – these are departments where forecasting is used, where time series complexity is used. You can start small – with smaller things that prove value and add to your portfolio while still working in the same domain and you gain knowledge in the other.

56:08

Alexey

**In sales, there are so many applications, actually. One of the things is lead scoring, right? So who is the most promising lead from the backlog of leads that you have, right? [Irina agrees] It doesn't have to be very complex. Maybe simple logistic regression.**

56:25

Irina

I was looking at how we could better forecast organic registrations for a dating app and I was thinking, “Okay, just plain forecasting like they do in Excel – this already exists. But that doesn't include things like density, things like population age, or the exchange of dollar.” So if you have those time series, and you enrich your time series with this, you can have a better forecast. And again, this can be done pretty easily by someone in this role. I'm happy to give her plenty of ideas to that person as well on how she can do this. She can book a session with me, she can email me. Not a problem.

# Finding Irina online

57:14

Alexey

**What's your email? Or what's the best way to contact you? I guess it's LinkedIn, isn't it?**

57:20

Irina

It’s LinkedIn. But you can also write first name dot last name, at Gmail.

57:26

Alexey

**So it’s Irina.Brudar@gmail.com [Irina agrees]– because you have two first names.**

57:31

Irina

Yeah. My mom and my grandma couldn't agree, so my dad said, “Why don't you choose both?” [laughs]

57:39

Alexey

**You usually go by Irina?**

57:40

Irina

Irina, yes.

57:42

Alexey

**So you don't use both?**

57:44

Irina

No.

57:45

Alexey

**Okay. [chuckles] So even though they didn't agree which name to give you, at the end you still use Irina more.**

57:52

Irina

My dad said “I'm Switzerland,” let the women choose. [laughs]

# Fundamentals for data analysts

58:01

Alexey

**[chuckles] There's another question about the best technologies to practice. But you said that you're not a tool person.**

58:08

Irina

Practices SQL, because it's the bread and butter for 20 years for me and for many people. Yeah.

58:15

Alexey

**Anything else that is as fundamental as SQL?**

58:20

Irina

For which field?

58:21

Alexey

**Data analytics, I guess?**

58:23

Irina

Data visualization, slides, soft skills, communication. [chuckles]

58:30

Alexey

**Since the question is about tools, then I guess for data visualization, you need to pick a tool. Does it matter which tool? Just pick one?**

58:38

Irina

It doesn’t matter which tool. And all of these tools, you can also find Tableau for free or they have a public version of it. Play with it and build your portfolio. But if you want to go into data analytics, SQL is the bread and butter. If you can add some coding on the side, that’s great. But if not, I don't think it's really required. I know some analysts have an area of focus, be it data visualization, or troubleshooting.

It depends, again, on the business model. If it's an online business model of ecommerce, then I highly recommend understanding “How does the page work? How does the app work? What is the normal user flow?” Probably understanding a little bit about tracking and product analytics is very helpful. Cookies are also. Because then you understand how things and data flows through the app itself in order to come to your database. And if you understand that, it's easier to work.

59:39

Alexey

**Don't get stuck choosing a tool, right?**

59:45

Irina

No, don’t.

# Suggestions for DataTalks.club collaborations

59:46

Alexey

**Pick any. I see that there are a few comments, so I want to mention them because they're interesting before we wrap up. One of the comments is “I was one of the few women doing the DataTalks.Club Zoomcamp. It was hard being one of the few. So the suggestion is join the effort with women who code and other women's organizations like Women in Data Science.” Thanks a lot. Great suggestion. I will, definitely. I am in touch with some of them. So we’ll definitely talk.**

60:16

Irina

Yeah, there are groups like Women in AI, Girls Who Code. Yeah.

60:20

Alexey

**Then a comment from Adonis is “Targeted advertising sounds like a good idea.” If anyone who is listening to this conversation right now knows how to do campaigns like Facebook.**

60:32

Irina

I can!

60:34

Alexey

**Oh, you can? Okay, good. [chuckles]**

60:37

Irina

Facebook very little, but Google I know.

60:41

Alexey

**Yeah. Okay. Because you’ve spent six years doing that stuff, right? [chuckles]**

60:45

Irina

Actually, last year, I had a mismatch in a role. I was supposed to be head of analytics, but I ended up doing tag management recommendation, troubleshooting, Google shopping ads, brand campaigns, conversion tracking – I can also do that. I can moonlight as a marketing analyst any day.

# Conclusion

61:06

Alexey

**We have a bit of a budget. We have sponsors, so we have a bit of budget that we can spend on promoting too. Yeah, I'll be in touch. [Irina agrees] Okay. Thanks, everyone, for joining us today. Thanks, Irina, for being here with us today, for sharing all this knowledge, all this expertise with us – for giving all this advice. Thanks for asking all these questions and sorry, we didn't cover all of them.**

61:36

Irina

I’m happy to answer on LinkedIn, or in email, I am very responsive and I'm very active and always open.

61:43

Alexey

[**Irina.Brudar@gmail.com**](mailto:Irina.Brudar@gmail.com) **right?**

61:47

Irina

Or just google me and you will find ways to get in touch with me.

61:52

Alexey

**Okay, then. Have a great weekend**

61:54

Irina

Thanks, Alexey. It was a pleasure. Thank you very much.