1:44

Alexey

**This week, we'll talk about hiring data scientists. We have a special yesterday, Olga. Olga is a delivery director – a delivery data science director at Microsoft. Olga has worked in AI for over 16 years. She has worked in different roles, from researcher to product manager to people manager. As a director, she's quite involved in hiring data scientists, so we invited her to talk more about this. Welcome.**

2:14

Olga

Thank you very much for having me.

# Olga’s career journey

2:17

Alexey

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

2:24

Olga

Right. I started really early with what you would refer to as data science. I was about 18 when I was studying at the university, and I was studying Applied Mathematics. And then out of all the subjects mathematics, statistics was basically the one that got me fascinated right away, because it made sense. It was something that was one of the most tangible. Then I discovered a lot of uses for statistics. I was also very, very lucky to have great lecturers and great professors along the way. So I think it's also important to have great teachers in your life. I was really lucky. And then I kind of stuck with this since then.

My first job was in 2006. I was working for a small consulting company – a boutique consulting company – doing forecasting of box office revenues of motion pictures. I think the model, or at least the approach, definitely evolved. But I think the approach is still in use by the subsequent derivation of that company today. Then I worked for a mogul in telecommunications doing the very classical use case template in communication, which is called churn prediction. Then I decided to pursue my BA – no, first, I think it was my Master’s, and then it was my PhD. All of the stuff that I've done was also kind of applied statistics, machine learning.

Then I worked in a bunch of places – five different countries, startup setting, freelance, more consulting. Now it's a very large, (I think the best in the world) software company. [chuckles]

4:15

Alexey

**Nameless company. [chuckles]**

4:17

Olga

I think I have always done things which have to do with financial services, mostly – or health. Basically, if you put them together, it's well-being with a small exception where I spent two and a half years working for an airline. It was a very romantic time. I was living in the airport. It was incredible. I do not regret it in the slightest. It was really kind of a separate experience. It was really great.

4:48

Alexey

**What did you do there? Also related to data science forecasting or something else?**

4:52

Olga

Yeah, it's a German airline, which you might know. Yeah. [chuckles] I was working there, first as the lead data scientist and then I worked there, basically, not as part of the airline, but in a consulting company, which was working with the airlines that was devoted to it. I was leading AI projects, initiatives, and helping create personalized flight experiences.

5:20

Alexey

**So you did the PhD, you worked in a startup, then you freelanced, then you joined the consultancy company, and now you work in a corporation. It’s like you tried everything, right?**

5:33

Olga

I also did a postdoc. It was also always a dream of mine. It's something that many PhDs want to do. Sometimes they want to do it for reasons that they cannot explain to themselves. Alexey, I also told you, I think, while we were talking about preparation of the schools, one of the podcasts that you have made really moved me and motivated me to speak, which was about the transition from academia to industry.

This is also a story of my life as well as many other lives. This is why, if you have done a PhD, you want to do a postdoc. You consider this and you do not really know why. You just want it and you go for it. So it was the story of my life. I don't regret it in the slightest. I'm still friends with all the people I used to work with. Yeah, it was a fantastic experience.

6:25

Alexey

**By the way, check out this episode that Olga was talking about. I think I was talking with CJ Jenkins. It was called “From Academia to Data Science”. I think it was a couple of seasons ago, check it out. That was indeed the cool one. Olga, what was your PhD about?**

6:48

Olga

It was about the application of machine learning to assess air pollution. It was in the metropolitan region of Barcelona, precisely. This is important, because you could have many diseases that could be linked to certain levels of air pollution. They're not directly related, but they could be potentially linked. I think it's a little bit common sense that if you have highly polluted areas, the propensity of somebody getting an ailment is higher.

The idea was to create a measurement method to be able to come up with a pollution surface, to be able to precisely estimate the level of air pollution at a given point and thus be able to say what would have been the estimate of air pollution of a certain place. And the strength of this was one of the interesting methods, which is called conformal predictions. Instead of giving you a point estimate, it would give you an interval, but it will be a very valid and very reliable estimate. The more data you have, obviously, the narrower your interval is.

8:09

Alexey

**Was it something related to Bayesian statistics, where you get a distribution… [cross-talk]**

8:13

Olga

You could do that as well, but not necessarily. Yes, there were some parts of underlying methods that use Bayesian inference – so I haven't used the word in 75 years. But it was frequentist as well, so it didn't really have to be.

8:30

Alexey

**Okay. And what do you do now? If you can tell us.**

8:34

Olga

What do I do now? I lead a team of people who are much smarter than I am – much more intelligent and much more advanced in modern technologies, in modern data science, they’re much better versed. So I try to support them on their way towards supporting our customers and making a positive impact, especially a positive business impact, working with our customers. I mean, the customers of Microsoft are obviously very important. We also work…

I'm kind of on this podcast today as myself, not as a representation of my company. So everything I say, I have to say that this is my own opinion, and does not necessarily reflect or does not necessarily somehow match the opinion of my employer. But I'm very proud to work there, I have to say. I mean, our motto is to empower every person or organization on the planet to achieve more. And I'm just fascinated by how we work with different sizes of organizations from moguls, corporations, Forbes’ 500 list of companies, up to startups, essentially.

I have people in my team, me included – but I don't really spend enough time on this as much as I would want – to support startup founders and also female founders. We have a program, which is called Microsoft for Startups. This is also something which is quite cool. But going back to what I do, as well, essentially it’s supporting my team in driving the best and the most meaningful business outcomes for our customers with data science. I'm also responsible for business development in specific regions in EMEA, and to help drive technology innovation, essentially. And to do the right thing with AI.

# Hiring data scientists now vs 7 years ago

10:38

Alexey

**And, as I mentioned, you're involved in hiring, right? That's why we’re having this conversation today.**

10:44

Olga

Yes, actually, I have been involved in hiring for the last seven years, I would say. I’ve been hiring data scientists for about seven years’ time. It has not been a small journey, really.

11:00

Alexey

**Since you mentioned that, I'm really curious how the definition of a data scientist changed over these seven years. Was it the same seven years ago? Or now the data scientists you hire are completely different from the data scientists you hired seven years ago?**

11:16

Olga

It depends, Alexey. Because if you hire senior people, they are the people from before. I think they are mammoths, like you and I, [Alexey laughs] who started studying when there was no such thing as data science. So my newest hire, they graduated from data science programs which are formally called Data Science Programs. When I started, there was no such thing. I studied applied statistics, you know? What did you study, by the way?

11:41

Alexey

**Information technologies.**

11:44

Olga

You see? So there are so many different walks of life from where people come to become data scientists. Physics, all flavors of engineering, neuroscience – social sciences as well. I think that sometime, so basically seven years ago, there were a little bit less expectations, let's put it that way – less expectations around what a data scientist would do. But I guess what has stayed in the variant is pursuit of excellence, to be honest. Because a great data scientist back then is still a great scientist now. You could have been excellent seven years ago, and you could be excellent now, if that makes sense.

12:39

Alexey

**The thing I observed – maybe it's just that my sample size is a bit biased – but what I saw is that now there is more emphasis on the engineering aspect of data science. And seven, eight years ago, the focus was more on mathematics and the mathematics behind machine learning. This is at least what I observed. But again, I myself was moving more into the engineering direction. That's why maybe – just the things around me changed, but not the environment, let's say.**

13:16

Olga

Right, I would agree with that because I think that five, seven years ago, we didn't have the infrastructure, really, to deploy. For instance, we couldn't even speak about MLOps, because we didn't really have infrastructures to operationalize whatever. Actually, the notion of data science is really, really old. If you think about, for instance – let's leave academia alone, because they have been doing this forever, kind of analyzing data and coming up with inference based on that. But if you look into the industry, for instance, these really old industries such as insurance, or for instance, banking – they have had their quant desks, they have had their analytics desks, they have had their actuarial practitioners for a very, very long time. [cross-talk]

14:03

Alexey

**50 years, right? For 50 years or… 30 years, for sure.**

14:09

Olga

Yeah. Longer than that. The thing is, to be honest, were they doing data science? Absolutely. Were they putting it into prod? Not really. I remember seven years ago, there were a lot of use cases deployed on a local machine. This is true. But nowadays, it's definitely the growth of computational technologies and the rise of cloud which kind of goes to the whole thing. And I agree with you, there is now a lot of focus on engineering. Also, my honest opinion is that it takes a village to deliver AI sustainably. It’s not just one person who knows a bunch of algorithms, to be honest.

# The two qualities of an excellent data scientist

14:49

Alexey

**Now, when you're hiring data scientists – when you're looking for a data scientist – what are the qualities you're looking for, typically?**

14:58

Olga

There are two things which are very, very important. We always… and this prevails, it's true for me. It was true for me as a hiring manager in my current job, and it was true for me when I was working for another company. Two things, essentially, which are pivotal for me. One is technical skills – they have to be excellent. Really, really. This is something that you really need to bring your hard skills with you. You really have to work on them. If you are feeling that you're not quite there, there is no reason for you not to get there. It's all a matter of practice and studying.

15:35

Alexey

**Well, what do you mean by “excellent technical skills”? Can you maybe give an example? Great Python knowledge? Ability to solve different algorithmic challenges? Or?**

15:45

Olga

That's a great question. What I mean by “great technical skills” is essentially. [sighs] How do I say this? You know, if you're a professional chef, you know that you’re good. [Alexey laughs] You know how to chop an onion very, very well, but you also know how to create masterpieces with it. What I'm trying to say is that essentially, for me, excellence is something which cannot be cataloged, guys – you cannot really put a definition on that. But for me, it manifests in understanding very, very deeply what you do. For me, an excellent data scientist understands what the solution does on the level of underlying algorithms, one that can really plot it, write it, describe it with a mathematical formula on a piece of paper. And they could explain very, very simply what it actually does, and how it does it.

What would be the underlying assumptions on data? What would be the risk? One of my favorite interview questions, for instance, which I asked a lot of people is, “What does it mean? What does data mean?” It sounds very, very obvious, but try it. You will see so many interesting answers to that. It gives you a lot of background of how people actually feel what they do. Then another thing, which is very, very important for excellence is, well, essentially practice. You have to have a passion for this. Because if you don't have passion for this, you will not be able to practice this over a long time. How they say “practice makes perfect,” there is no other way.

There is no degree of talent, in my experience, that will compensate for lack of practice. Something has to drive you forward. I mean, it's obviously discipline, but it's also practice. So yeah, through practice, an understanding of what the thing does, and constant learning, you eventually achieve excellence.

17:47

Alexey

**I guess this is pretty subjective. When hiring, you want to be objective, right? So how do you check if this person really has this deep understanding of what they do? I guess you just asked questions, like “Hey. What do you do? Can you explain it to me?” Right?**

18:03

Olga

Right. Do you mind if I answer about the second aspect of what I'm looking for? You asked me about what I look for in a data scientist.

18:11

Alexey

**Yeah, please. The first one was technical? Right? And the second?**

18:15

Olga

Attitude.

18:16

Alexey

**Attitude. Okay.**

18:17

Olga

Yeah. I think that's equally important – one of the most important things is the attitude. You have to be a constant learner, you have to know your worth, you have to be humble – you have to love it, really. Also know precisely, or at least be honest with yourself, where you’re going to and what you're going to do there. Try to find as much as possible. So are you going to work on a product? Now you go to work for the customers? Are you going to work deeply in research? And then have the right approach to that.

18:53

Alexey

**So how do you check love?**

18:55

Olga

How do I check what, sorry? [cross-talk]

18:57

Alexey

**You said attitude is important and that people should love what they do. How do you check if they actually love it and they are not there for the money? [cross-talk]**

19:05

Olga

It shows. It absolutely shows. I mean, passion is something that you cannot really conceal.

19:10

Alexey

**Can you maybe mention how it can manifest when you see that the candidate *doesn't* have this attitude or *doesn't* have this love that you talked about?**

19:23

Olga

Well, I think, again, you are absolutely right – it's very subjective. I'm not a guidebook on hiring, to be honest. I'm a bunch of [cross-talk]

19:31

Alexey

**I’m curious, but you don’t have to answer that.**

19:34

Olga

I promise you, a fair share of those experiences and views will potentially be corrected over time because I'm wrong very often, like every human person. Also it's a collection of mistakes that I have made, but what is true for me nowadays – “How do you check for attitude?” You ask about what I would do. –I would ask about motivation. “Why do you do data science?” Potentially not so directly, but you can ask a bunch of questions around this to understand what the person's motivation is to go to the field.

20:09

Alexey

**So it would be behavioral questions like, “Tell me about a time you did something.” Right? These kinds of questions?**

# What makes Alexey do this podcast

20:16

Olga

Let me try it. So, Alexey, why are you into data science? What makes you do this podcast?

20:23

Alexey

**Okay. [laughs] I was not prepared for that. I guess meeting people, if we're talking about this specific podcast. Meeting people and learning from them – learning from you. Right now, I'm actually taking notes. For those who are watching it on video, not listening – I'm showing now the notes I make. Then after this episode, there will be a transcription. I can go through this transcription and really learn this again and again. This is super useful – with every episode, with every guest, I learn something new. I guess that's one of the motivations why I keep doing it.**

21:02

Olga

Perfect. That sounds really good.

21:03

Alexey

**But your question was about data science, right?**

21:05

Olga

I’m still not done. Do you mind if I dig a little bit deeper? Of the episodes that you have conducted with your guests, what resonates most? I mean it doesn't necessarily have to apply to people – the topics or things. If I would ask you, potentially, wake you up in the middle of the night and say, “Alexey, what have been the most interesting topics? Topics that have been the most important to you that have been discussed here – what have those been?

21:34

Alexey

**Usually those are around career, more soft topics, rather than hard topics. Not topics about things like how GPT-3 works, but more like how you progress in your career, why are you doing what you do, or how you became what you are right now – so the career trajectory, different roles – this is the kind of stuff I really like. One of the recent ones was, for example, yesterday I interviewed a person about his experience of becoming a company founder. He was creating open source and became a founder. These are the kinds of episodes I really love.**

22:19

Olga

Awesome. That sounds really, really good. And the last question around this – what have you learned doing this exercise, this podcast? What is that you have learned about people, about data scientists, about the human aspect?

22:35

Alexey

**That is a pretty broad question. I don't know how to answer that. [laughs]**

22:39

Olga

The most important thing – that kind of sticks.

22:43

Alexey

**People are interesting. Everyone is different. And everyone has a story to tell. It's funny, when I reach out to people and say, “Hey, do you want to be on the podcast?” Some of them say “Yeah, but there’s nothing I can talk about.” But then it's like the best podcast episode.**

# How Alexey get the latest information on data science

23:01

Olga

Right. And then another, almost last question – I also know that you have a day job as a principal data scientist. On your reading list – the stuff that you're reading right now – what is the most prominent source of where you get the news about the latest updates on data science?

23:17

Alexey

**Yeah. Twitter, I guess. I check Twitter and then if something gets a lot of exposure, I look into this. But usually these are topics more around the engineering aspects of data science, rather than “the latest trends in AI”. I, of course, played with DALL-E, it was fun. I tried different prompts. You know what DALL-E is, right? You give it a prompt and then it generates nine pictures. I was trying weird stuff like, the guy from KFC is Colonel Sanders, right? He's on the logo of KFC, the fast food chain. So I was trying things like, “Colonel Sanders is having a Zoom call with puppies to take over McDonald's,” and then it will generate random pictures. So that was fun. But yeah, I'm just playing with this kind of stuff. But I am reading more about engineering stuff.**

24:15

Olga

Yeah, I see. And I infer that this is why, probably because you have more passion for the engineering aspect of data science rather than the algorithmic one. Like you said, everybody's different. Some people, like myself, for instance, will be more driven to potentially… I mean, I'm very much into technological development. The first thing when I read about new technology, the first thing that I tried to infer for myself is how it works under the hood, so “What is the underlying method?” Today, when you said “Bayesian,” I was like, “Oh, my God,” because the level of conversations that I normally have in my daily job, I do not often get asked about it. And that's a shame. I should probably work on learning it myself. I probably should have more such conversations because I really, really love that.

25:03

Alexey

**I just remember it from my university days – when you have point estimates, it's a frequentist approach, but when you have a distribution, that's Bayesian. That's why I thought, “Okay, let me show I’m smart too.” [chuckles]**

25:16

Olga

You are. [laughs] You officially are. Stamp of approval. [laughs]

# How Olga checks a candidate’s technical skills

25:21

Alexey

**I think I asked you a question, which you kind of diverted from. We were talking about two things that you look at – technical skills and attitude. I think both of these things are kind of hard to measure. They're more subjective, but there must be some objective way to measure this, right? Attitude is more like what you just tried to me by asking these questions. [chuckles] What about technical skills? Do you also do this by asking questions like, “Tell me about a project you did.”?**

25:51

Olga

Yeah. You know, the thing about that – I have to be honest with you – I have had many, many, many, many, many, many, many, many, many, many, many interviews in my life. So this is something that you eventually get – it's like muscle training. You get a lot of intuition, just based on the experience that you would have had. Like you, I very much like getting to know people. For me, the interview process is always an opportunity for me to learn from another person – where they come from, what they have learned. Every time I have an interview, I learn something new. And that's absolutely fascinating. I think it's the most rewarding activity that you can do. [sighs]

But I would be very cautious answering your question. Because you definitely do have objective criteria – and you *have to* have objective criteria. It's absolutely mandatory. I will elaborate on that in a second. However, you must not forget that you’re dealing with people and people are not necessarily always objective. So I think one of the fallacies of the data scientists (generally, any technical people, I think), who move into more kind of a soft-ish discipline, is not to forget that you’re actually dealing with people and not to expect them to follow a blueprint or kind of “if *this,* then *that*” approach. With human beings, it’s utterly wrong. At least I think so. *But*, on objective criteria.

For instance, for me, it has always been very important to conduct a technical interview. You could have different shapes and forms of that. It could have some kind of code exercise, it could have some kind of analytical exercise as well, and then you would ask a bunch of questions. And one of them is, “What does it mean?” [laughs] Then it will also help you understand very many things about the depth of somebody's knowledge. For specific things, you also…

I think one of the things that you need to also remind yourself of (or at least I do, when I interview) is to stay humble, and not *dig* into places where *I* know stuff. Because I can ask somebody, like, “Are there any kinds of distributions that wouldn't have a mean?” And then get a person completely lost and frustrated during the interview. But I also have in mind why I'm asking these questions – what do I want to test? What do I expect this person to be able to do in order to be successful at this job? So this is important as well, to also to prepare for your technical interviews and be very meaningful about what you're asking. Also, to be respectful all the time.

28:32

Alexey

**So are there distributions that do not have a mean?**

28:34

Olga

Yeah. Sorry, I probably shouldn't have asked that. [laughs]

28:37

Alexey

**[laughs] What are those distributions?**

28:43

Olga

I invite you to check Wikipedia. [laughs]

# How to make an answer stand out (showing your depth of knowledge)

28:48

Alexey

**[chuckles] What was the most interesting answer you got when you asked about the mean?**

28:55

Olga

The most interesting answer?

28:58

Alexey

**Yeah. Because I would just answer “It's average.” Right? Just “.mean” you know?**

29:04

Olga

Do you really want to go there? [laughs]

29:07

Alexey

**I’m just curious.**

29:08

Olga

[laughs] All right. No, but I mean – interesting answers. I think when you say “interesting,” you mean something that has stuck with me – their reply – right?

29:16

Alexey

**Something unusual, something you remember, yes.**

29:19

Olga

Yeah. So I remember a person who we hired answered that question and they were very specific in their answer, because they also have given kind of a textbook definition of a mean, or an average, of the data. They have provided an example of potentially how you measure this in a sample, they have provided an example of how it extrapolates to whatever population you will take the sample from, and then they have provided a practical example of how this applies to life. This was really, really good. It was really impressive for me.

30:02

Alexey

**Yeah, that's pretty extensive.**

30:03

Olga

The depths of knowledge. You know, for instance, one of the ways you could ask, if you're still not convinced with my favorite question. You could ask somebody, “Let's assume you want to move to a new country, which you know *nothing* about. Their currency also says nothing to you.” I mean, you can obviously check it at places like xt.com, but you will not live that and it will not be something that would be natural to you to operate in – in this environment. “And then you would want to buy or rent a house. And then you would definitely want to check the neighborhood and stuff like that – you would want to know the average level of prices in this neighborhood or in this country. Basically, what would be the main statistic? What will be the mean, essentially, and what would be the deviations of that and stuff like that.” So, on a human level, that makes a lot of sense to know why you want to know this. And this gives you an understanding, essentially, why you want to know this is what this means, and how it translates.

31:15

Alexey

**Interesting.**

31:16

Olga

One other thing, it’s also maybe it's because I'm a statistician – one of the things is that people very much talk about predictive statistics and prescriptive – this is very, very important. Definitely. But one should not underestimate the power of descriptive statistics. When I studied mathematics, they told me it’s elemental not because it's simple – it's elemental because there are elements there from which other more complex things are constructed.

31:46

Alexey

**So descriptive statistics is all this mean, median and all these kinds of characteristics of data, of the sample you have, right?**

31:56

Olga

Descriptive statistics are the statistics that describe the data.

32:00

Alexey

**My statistics is a bit rusty. I know how to fit and predict in SciKit Learn… [chuckles] but if you ask me… [cross-talk]**

32:10

Olga

You should be able to understand what the data tells before…

32:16

Alexey

**I’ll take a refresher, I guess. [chuckles] After this conversation.**

32:17

Olga

Yeah, honestly, I take it all the time because I forget stuff. And then when I have the time, I will just read old books like, for instance, Hastie/Tibshirani, “The Elements of Statistical Learning, I think [cross-talk]

32:28

Alexey

**That's a good one. I have it.**

32:30

Olga

It's a very good one.

# A strong mathematical background vs a strong engineering background

32:32

Alexey

**Okay. Since we talked a bit about statistics and math – what do you think? What is more important, a strong mathematical background, or a strong engineering background when we hire people? Or do we just hire both? How do we decide?**

32:50

Olga

That’s a very interesting question. I think it depends on what you want to achieve.

32:56

Alexey

**But let's say you have a team, and you want to hire a data scientist for the team. You have some objective –there is a product – you want to add artificial intelligence/machine learning to this product. Now you have a task to hire three data scientists. So how do you decide if you want to get people that have a stronger mathematical background or those who are stronger engineers?**

33:24

Olga

When I hire my data scientists, I need to be able to, very, very disambiguously know what I want them to do. Just today, we have had a team meeting and we discusses things about the scope of data science work recently, because it's an emerging discipline *still*. It’s been very broad and it doesn't always play as a favor for data scientists. You know, in German, there is a concept of Eierlegende Wollmilchsau.

33:52

Alexey

**I have no idea what that is. [chuckles]**

33:54

Olga

It’s a very interesting animal. It's essentially a pig, which also gives wool, and also gives milk, and also gives eggs. So we should avoid having a data… I mean, a more polite way to call this is a Swiss army knife, or another way to call it is jack of all trades. But being that person is really no fun. Honestly. And it is also difficult to retain somebody like this in the long run. Because if you expect the person to be able to do everything, like “Yeah, whatever. You can, so you have to.” It's not necessarily fun to be that person.

34:34

Alexey

**Okay.**

34:35

Olga

So you have to be very specific. When you want to hire data scientists, you have to be very, very specific. What do you expect them to actually do?

34:43

Alexey

**If what you expect is integrating with the backend of whatever the existing team is doing, then it's probably more engineering, but if you expect them to do more analytics, more talking to business, then it's probably more on the other side of the spectrum, right?**

35:00

Olga

Yeah. Let's assume you are a startup and your budget is really limited. You want to deploy a very specific use case. And then you already have a couple of people. Let's say, with no lack of generality, your language is Python and you have everything. Then, for instance, you would want somebody who will have a deep understanding of the underlying algorithm, and will be able to develop a custom solution, and who will be able to put it into prod and ensure that it's not just kind of something that works in the research environment, that it runs in real time, its realizable, and so on, and so forth.

You would expect this person to understand how this works, how this could be parallelizable. But potentially, if you already have a couple of very, very good software engineers, you will also be able and you will be willing to potentially forgive a lot of (forgive us the wrong word) but be kind of *accepting* of this person's potential coding skills or on par with those of the professional software engineer. And actually, generally, for data scientists an expectation should not be that they should be an *awesome* software engineer, because it's not their job. I mean, they *can* be. I have met people who are. I have respect for them, to be honest. I'm not one of them, unfortunately.

But en masse, because you would also have expectations of a very deep understanding of algorithms – what the underlying thing *does* and how it basically stems from the data landscape that you have – what kind of they do have to have. You would not expect this person to have tremendous programming skills. They can write what allows them to research and that's okay. While it's commented. [chuckles]

36:41

Alexey

**So if you need somebody with strong engineer skills, just hire a developer, not a data scientist? Right?**

36:48

Olga

Depends. Because, for instance, if you have a solution that you need to be able to productionize – if you need a solution where, essentially, the engineering aspect is more important for you, then you will look for somebody who has an understanding of underlying data science methods, but not necessarily so well-versed in different underlying algorithms. One who has an understanding that is potentially a little bit basic of how the mathematics under the hood works, but then you would need them to really, really be good in programming. It's also fine. It depends on what the business needs.

37:30

Alexey

**We have quite a few questions. From the questions I prepared, we covered only two out of I don’t know how many. [laughs] But yeah, I think it’s time we also took a look at the questions from the audience.**

37:43

Olga

Sure.

# When Auto ML will replace the need to have data scientists

37:44

Alexey

**The first question is from Ilya. Ilya is asking, “When will Auto ML replace humans and hiring data scientists will not be necessary anymore?”**

37:51

Olga

Never.

37:52

Alexey

**Never. [chuckles] Why?**

37:58

Olga

[laughs] Well, because of the human aspect. You know, there is a fantastic book called “Human and AI,” which is written by Accenture’s chief technology officer, I think in around 2015. And it talks about the professionals of the (so-called) future. Definitely AI is developing – Auto ML is tremendously powerful and this power grows day by day. And this is good, because we also want it to. Also, some no-code environments such as, for instance, from different designers. They're also fantastic because they enabled so-called “citizen data scientists”.

This is the very, very right way to go, because you will have some of the solutions, really robustly tested and safe for a person with no degree in advanced mathematics to be able to safely use it to meet the goals of the business. But you would never… This is what we discussed in the beginning. Data scientists are first and foremost people. You will never be able to replace the human at the end of it. Because you do things for humans, essentially, so this is why I think it's never. I’m not a futurist, but I think you will need a human in the loop.

39:21

Alexey

**I think one of the most common reasons that data science and machine learning projects fail is misalignment – miscommunication between different groups – like between the stakeholders, the product people, the users, and the data scientists who implement the solution. You have to have this conversation.**

**You have to have this communication with others in order to do what is the right thing to do. Right? Now if you replace data scientists who are okay with communication (maybe not the best ones, but okay) with machines who do not communicate at all, who expect clean data then, of course, it will not work. We still need communication. AI will not magically solve all the problems of your users.**

40:09

Olga

You know, there is certain automation possible for data quality checks as well. So it's not about that. It's what you said – about the delivery, or the deployment of an AI project (or any project). The most important thing is vitamin C, which is Communication. And this is not my finding. I read it in a book. There is a woman from Spain, Araceli Segarra, and she is the first woman who has climbed Mount Everest and she wrote a book about this. I love this book. I think one of the things which resonated with me most was – essentially, she provided examples where the lack of vitamin C was leading to outcomes where people literally died, going up. So this is pivotal that people communicate, really. And this is why I think that you will always, always need people.

41:11

Alexey

**Yeah, you can automate some things, but Auto ML aims at simplifying the work of data scientists, not replacing them.**

41:18

Olga

Exactly. Essentially, you have hit the nail on the head. It's simplifying the work that you're doing. Because before, you would spend a lot of time trying and testing different algorithms, but now you have the opportunity to kind of have the machine do this for you very, very quickly, and then save a lot of your time and energy. But it's the same, like, before we all used to wash the dishes by hand and now we all have dishwashers.

41:46

Alexey

**I recently discovered the beauty of having a dishwasher. I didn't know how to start it. We had a dishwasher for like, eight years? I just didn't know how to turn it on. My life will never be the same now, because I know how to turn it on now. [laughs]**

# Should data scientists transition into management? (the importance of communication in an organization)

42:09

Alexey

**Anyways, an interesting question, “Personally, I am good at data science, but my manager wants me to do a course in management, which I feel is not relevant for me since my passion is in data science.” Do you have any suggestions for people in this situation? What can they do? Let's say you don't want to become a manager, but your boss is kind of trying to push you a little bit in this direction, saying, “Hey, we need managers. We want to grow the team.”**

42:40

Olga

I think it's a vitamin C thing. I think you need to have an honest conversation with the manager and understand, basically, where they are coming from. To understand what their motivation is and why they're pushing you to do that – what they want you to do. Do they want you to really lead the team? Do they want you to take more responsibility? And essentially, why you? Because, essentially, if you don't want to, you don't have to. That's the thing. The career path of a data scientist is a very, very hot and acute topic. People expect some kind of blueprints that will be robustly tested in this. And it's not possible, because discipline, especially in the industry that exists for about seven years. A little bit more maybe.

But you still don't have enough examples of people who would reach C-level positions. I mean, there are, but there are not as many as you would have, potentially, in engineering – to be able to understand what your options would be. Because in many environments, people assume automatically that the only way up is through management. When you start managing, how my boss likes to say, which I really, really like – he says “Everybody can be a leader, but not everybody can be a manager.” Extending this quote, I will say that everybody can, and it's absolutely fine. You don't have to manage people to be a data science leader.

Because you have to understand one thing – it's physics. It's the law of conservation of energy, right? So in order for you to give more to something, you will have to take this energy from somewhere else. So if you manage people or if you spend time in business meetings, it means you're less hands- on and you lose touch with the technology. I mean, you practice your craft much less, if at all. It's also something that you have to take into consideration. So it's essentially a matter what you do in your life, you know?

Going back to this example, I would say to have an honest conversation and try to understand where the manager is coming from, and then provide an alternative solution. Say, “Look, potentially, if I do this, you will lose a great data scientist because I will kind of lose my practice. Are you ready to do this? Are you willing to do this? Does it go well with the goals of the organization?” Because maybe there is a better solution – you can hire somebody else to be a team lead, and let this person be a kind of advanced, as a manager/specialist or however you call them in your organization – a kind of a senior responsible person for technology or something like this.

45:27

Alexey

**So the risk for the company here is to lose a great data scientist, but maybe get an average manager?**

45:35

Olga

Precisely.

# Switching from a data analyst role to a data scientist

45:37

Alexey

**So the company should really think about this. Okay, fair enough. How hard would it be to switch from a data analyst role to a data scientist?**

45:53

Olga

Depends.

45:57

Alexey

**[chuckles] Depends on the company where you work? Or depends on what you call a data scientist and the definition you have in your head?**

46:06

Olga

I guess on your background. I mean, also what kind of data scientist do you want to be? Depends.

46:12

Alexey

**Yeah, there are different one, like data scientists who are more leaning towards the business side of things. And then from analytics, I think the switch is not that difficult, right? And there are data scientists who are more into MLOps and engineering, and then it would require taking coding classes and getting practice.**

46:35

Olga

Yes, that as well, but also understand the underlying mathematics.

46:39

Alexey

**Yeah, right.**

46:42

Olga

So it depends on how well-versed people are already in this. You can learn anything, because we all learn. This is something which is very, very important. What makes a great data scientist? A great education, and then how many hours of practice they have. You know how they say that in order to become an expert in something, you have to put in 10,000 hours of practice – give or take.

# Attracting female talent in data science

47:06

Alexey

**One thing I also wanted to talk to you about was one of the things you mentioned when we just started this conversation – about inviting you to this podcast. You mentioned that you managed to attract a significant share of excellent female candidates with excellent female talents to your team. I thought about this, and I'm also part of the hiring process – I'm a hiring manager, I also take part in hiring as the interviewer – and I don't see many females in the candidates. So maybe the ratio is 9 to 1, so 90% males and only one, one and only 10% females.**

47:50

Olga

Don’t see them where? What do you mean when you say that you don't see them?

47:54

Alexey

**When people apply, mostly it's males. So I'm really curious how you manage to get a significant share of female candidates and female talent.**

48:07

Olga

Okay. Now I'm going to speak, I guess, on this because it's a very, very good topic, which is very dear to my heart for obvious reasons. Probably I will open the gates of hell, but you know what? I guess it's the truth and it has to be spoken. It's very, very hard. I mean, it's much easier for us now, but it's generally very, very hard for us women in STEM fields, because the bias against us is systemic. It's more pronounced in some societies and less pronounced in other societies. But if you focus on this a little bit more, you would see that.

For instance, when boys start developing and growing up, they are almost encouraged and incentivized to go into things such as engineering, mathematics, or physics. Whereas girls are more expected in social pressure to prioritize stuff like liberal arts (which is a very honorable discipline) and care. But there are not too many girls who would choose mathematics. I will be absolutely honest, some women just don't care. They're not interested in it, really. And it's all right, because some people are interested in literature.

Some people are interested in knowing history. Some people are interested in music. It's all right to not care about mathematics or engineering. But it has nothing to do with gender, but a lot with the social kind of pressure and the expectations. Also on what people intake from their families, which have also been brought up by generations of people thinking things a certain way. So it's not potentially some one person to be [cross-talk]

49:46

Alexey

**I think I know what you're talking about. In my group, when I was studying at university, there was only one girl – in a group of 25 people.**

49:55

Olga

When I started, we had 25% females and I think that was quite a lot. But then later on, something that happened also is that girls are a little bit overseen, oftentimes – not *all* the time, but oftentimes – when they study, because they're not expected to have a certain level of ability or dedication or whatnot. I have not encountered this myself, but I had friends who told me that they personally have encountered this discrimination, I would say, based on the agenda. They would say, “You will never attain certain heights because you're a woman,” directly or indirectly.

Also, it becomes increasingly difficult for women when they enter the workforce, because there are many biases such as, for instance, like-like bias. So you would normally unconsciously be prone to like people who are more similar to you because you feel safer around (let's say, broadly) people who look like you, who are like you. And this is why if you go to an assessment center, or to a company, and it's populated with males – that's it. They're going to basically say, “Yeah, this is a safer choice,” unconsciously, with *all* the best intentions in the world.

If you want to be able to attract diverse talent – basically, people who are different from you – you should potentially shift your mindset a little bit and also be more open and be more willing to understand where other people are coming from and what is important to them. That also applies to women, essentially. Another thing that is important to do is knowing – *knowing* that there are more males than females and the system is skewed (there’s kind of a male prevalence in this industry) and to apply more effort. Because I have seen a lot of people kind of expecting that somebody, all out of the blue, will source them outstanding female talent and just say, “Look, *you* choose.” It doesn't really work like this.

You have to be able to go out on a limb and go and search for these people. You go search LinkedIn and you search your weak ties – this is actually one of the best ways because application through reference or hiring through reference, in my experience, works really well. Because if you know people who are excellent, chances are, the people from their surrounding are also excellent. Then you can just basically ask and say, “Hey, do you know any excellent female data scientists who are now looking for a job?”

The last very important thing, which is also very important to mention, is that there is research (and I don't want to sound that I'm against men or anything – obviously, I'm not) *but* one thing also to say that there is research which shows that a woman would feel secure to apply for a job only when she ticks 100% of the requirement boxes.

And males are found to be able to kind of go with it, when they tick around 60% of boxes. So you would really need to empower and encourage, potentially sometimes convince and help somebody and also boost their self-confidence. Because guess what? Self confidence has been *unboosted* over generations for years, while the technical knowledge is actually brilliant. Women are also excellent in many things. I mean, every person is a different world. You cannot just attribute some qualities to a man or a woman. And I say this was my shallow knowledge of psychology. [laughs] But en masse, what I have seen is that women are really, really good at organizing work – excellent.

53:53

Alexey

**One of the questions I wanted to ask is how to make job descriptions more attractive to female candidates. I think you partly answered that question by citing this research that women apply when they tick 100% of the boxes, but when men apply when it's 60%. So maybe one of the things that we can do is, when coming up with a job description, list only the skills that are absolutely necessary and omit all the ‘nice to have’ ones, right?**

54:25

Olga

Yeah, I think that's always a good approach. But it's my personal opinion, because at the end of the day, regardless of gender or background, you want to hire a person who can do the job really well. There is no such thing as a perfect candidate. You should also be very, very frank with yourself – what do you absolutely need the person to bring and what can they learn on the job? And you have to facilitate the resources for them to learn on the job as well. But I guess, yes – you don't want to end up hiring Eierlegende Wollmilchsau because it's not sustainable and they're not [cross-talk]

54:56

Alexey

**Can you remind me what this is? It’s like a Swiss army knife, right?**

55:00

Olga

A jack of all trades, yes.

55:02

Alexey

**[laughs] Yes, okay.**

55:03

Olga

Yeah, so you would want to basically – my main suggestion is always to boil it down to what you actually *need* this person to be able to do.

# Changing a job description to find talent

55:13

Alexey

**I came across a startup, or a company, where they offer you to upload a job description, and in this job description, they will highlight things that say, “Okay, if you include this word, it will discourage female candidates. Will you consider replacing it with a different word – a synonym.” Do you think this works? Is this how it actually works? I was quite surprised when I saw this. It was like replacing one synonym by another and they just say, “Okay, if you do this, then you will increase chances of getting more female candidates.”**

55:54

Olga

I think it does. I think generally, it does. Yes. Because I believe that there are some wordings which feel more appealing, potentially. I would not say to just try this, and if you rewrite your job description, you would also find it more appealing to men. Because normally, it will suggest that you focus on things such as collaboration, team spirit, psychological safety. I don't know a person of any gender who would not like that.

# Long gaps in the CV

56:31

Alexey

**I see that we don't have a lot of time left, and I know you need to go. So maybe, I don't know if you can answer that quickly, but maybe you can try. “Will you consider a person who has a gap of six years while hiring?”**

56:46

Olga

Yes, why not? Why not?

56:49

Alexey

**Well, because of the gap. Maybe you think, “Okay, they don't have the skills that are needed.” Let's say, if I have a gap, I might think this way – that the hiring manager would not think I have the same level of skills as the person who just graduated from a data science boot camp or something. Or not a boot camp, but a university or whatever.**

57:12

Olga

You know, the reasons people take gaps in their career are very, very different. And oftentimes, they have *nothing* to do with their ability to do things. [cross-talk]

57:22

Alexey

**Like kids, for example, for women?**

57:26

Olga

Not only for women. For instance, I know people who are men who have taken gaps just to take care of their families, because they had. This is normal. I mean, life happens. But I don't think that it in any way reflects their skills. Absolutely not.

57:47

Alexey

**Do you maybe have a tip for them to feel better, or more secure, when applying? When you have a gap, how can you feel better about it and just keep applying?**

58:03

Olga

I guess, it’s what you said – keep applying and believe in yourself. The thing about that is – I can tell you a lot of encouraging things – but the thing is that it's on everybody to believe in themselves. Yeah, it's scary out there. Absolutely. You know, I feel insecure or kind of anxious or vulnerable *many times* – often and frequently – because I'm a human being and it's part of the human experience. But at the end of the day, you just think, “Okay, I already have a no as a baseline. What do I have to lose? That somebody basically makes a judgment about my gap. It's on them, really. I don't have to inherit this. I don't have to absorb that.”

# Eierlegende Wollmilchsau

58:52

Alexey

**That is a small request for you. Can you type the German worth your said?**

59:00

Olga

I can't. Because I don’t know [cross-talk]

59:04

Alexey

**Can you say it slowly?**

59:06

Olga

No. You know what I'm going to do? I'm going to do this in the YouTube comments, if that’s alright.

59:13

Alexey

**Yes, that is alright. So please check the comments section after the call. The word will be there. Okay. I think that's all. Thanks a lot for joining us. Thanks a lot for sharing your expertise with us. We didn't cover most of the questions I prepared, but that's okay, I guess, because the conversation we had was really nice. Thanks for asking me questions. This is not something that happens quite often. Thanks. And thanks, everyone, for joining us today.**

59:41

Olga

Many thanks, Alexey, for having me. And many thanks, everyone, for the questions and the comments. I will definitely look at them again. And feel free to post more questions. I will try to answer them offline individually.

59:55

Alexey

**Okay, awesome.**

59:56

Olga

Thank you very much. Have a great rest.