1:40

Alexey

**This week, we'll talk about collaborative data science in business. We have a special guest today, Ioannis. Ioannis is a lead data scientist at EasyJet, if you’ve heard about this airline – I certainly have because I used it a couple of times. In his role, he works on creating data products and solving business problems. He also leads the EasyJet MLOps team. Ioannis is also one of the graduates of our MLOps Zoomcamp. I was quite surprised that he actually took it – with his experience, he should have been one of the instructors. [Ioannis chuckles] But I'm pretty happy that you, Ioannis, did take the course because otherwise, we wouldn't be talking now otherwise. Welcome!**

2:30

Ioannis

Yeah. Thanks for having me and for the introduction. It's been a pleasure.

# Ioannis’ background

2:34

Alexey

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

2:45

Ioannis

Yeah, absolutely. Education-wise, I have a bachelor’s in mathematics and a postgraduate in data science from Essex University. It's been fun because I wasn't always planning to become a data scientist. Essentially, I'm Greek and this is important, because in Greece, usually when you have a bachelor’s in mathematics, there are not many things that you can do with this degree. You either become a teacher – which is, although exciting, wasn't something that I wanted to pursue – or you find a way to mix it with some other things. After I finished my Bachelor’s, I was thinking about financial mathematics or actuarial mathematics. I didn't know what to do.

Luckily, I got introduced to the notion of data science by watching Netflix – actually, the famous Sherlock Series. There was a moment when Sherlock and John Watson were on-screen, and John Watson was impressed by Sherlock’s decision-making skills. I remember he asked him, “How do you make decisions that fast and so accurately?” And Sherlock replied, “You see, but you do not observe.” So that was John's problem. That really sat well with me, and I was thinking, “I want to improve my decision-making skills.” And this is how I started Googling around “decision-making, inference” and all this kind of stuff. I came across data science as a profession. That was back in 2016, I think. So yeah, I did a master’s in data science from Essex University, followed by a three-month internship, where I was able to develop a machine learning model to predict children who are being abused in their current environment. That was great because it showed me the power that lies behind data science and machine learning in general. I knew that this was what I wanted to do.

After the internship, I had a four-month experience working as a data scientist consultant at a company named AKKA Technologies in Geneva, Switzerland. After four months, I decided to move back to the UK, where I started working as a data scientist for EasyJet, where I'm still working. I started as a graduate data scientist, got promoted to senior data scientist, and right now, I'm still a lead data scientist, working with business stakeholders and trying to transform Easy to become the world's most data-driven headline. Yeah, that's pretty much me.

5:21

Alexey

**Do you get a discount at EasyJet if you want to go somewhere?**

5:25

Ioannis

[chuckles] I think that's one of the best perks that we have. [chuckles] Yeah, the truth is that we do and it's an excellent discount. I use it all the time to travel to different European cities. It's been great.

5:40

Alexey

**Because EasyJet is… when it comes to Berlin, I don't know about the other cities and I'm based in Berlin – it's one of the airlines I usually use when I want to go somewhere.**

5:53

Ioannis

I'm happy to hear that we're doing something good, then. [chuckles]

5:56

Alexey

**Well, in terms of coverage, it's probably one of the best ones – at least going to Italy or some other countries. Funny that you… [cross-talk] It’s funny that you mentioned the Sherlock TV show. Have you seen…? There is another different TV show (an American one) called Numbers. Have you seen that one?**

6:21

Ioannis

Oh, that's interesting. Not really. But noted.

6:25

Alexey

**It's about a mathematician who uses his skills to solve crimes. They use statistics and data science. Well, I wouldn't call it “data science” in the sense that you and I mean it. But still, it's quite close.**

6:48

Ioannis

I'm always excited to hear about these use cases where data science is being used for good, like the project that you just mentioned – to solve crimes or the internship that I did. I think it's great to show how data science can serve the people or not be present to replace people’s jobs are some of the things that you hear from time to time.

7:11

Alexey

**Yeah, so it's called Numbers. And I think the E is spelled with a 3. So, it's like Numb3rs.**

7:20

Ioannis

I think it rings a bell.

# Ioannis’ role as Lead Data Scientist

7:23

Alexey

**Yeah. Anyways, what do you do as a lead data scientist?**

7:28

Ioannis

Currently, my role as a lead data scientist is a partnership with the business stakeholders from Digital Customer and Marketing. These are the departments that I oversee from the data science and analytics perspective. I try to understand their pain points and translate them into data products and data solutions that go into production and solve whatever problem we encounter at the time. You can think of my role as having accountability for the projects to ensure that they reach production and, of course, we meet the financial benefits that have been agreed upon at the beginning of every financial year.

8:09

Alexey

**In practice, what do you mean when you say that you “partner with business stakeholders from Digital Marketing”? What does it look like in practice? Is it you proactively reaching out to them saying, “Hey, can we talk?” Or do they reach out to you? Or is it a combination of both? What does this collaboration look like in your case?**

8:32

Ioannis

It's a great question. Usually, one of the things that I love about EasyJet is that it's a really friendly environment. You can think of it as me having a close collaboration in terms of meetings, sitting with them during the business days, and trying to understand what decisions they have to make on a daily basis and then trying to understand, from their perspective, what their strategies are and what their vision is for their department, and understand how data science can support reaching their vision. This is how it looks on a day-to-day basis – meetings and meetups, etc.

9:12

Alexey

**So they have their usual day-to-day meetings, and you’re like, “Hey, can I join you? I just want to observe what you do.”**

9:21

Ioannis

Kind of, yes. We have a recurring meeting where we discuss what they're doing, brainstorm together to have – let's call it a framework, where we discuss their day-to-day job and what they're trying to improve and see how I can support them with data science.

9:42

Alexey

**So you have a monthly meeting or something like that?**

9:45

Ioannis

Even more frequent – weekly, actually.

9:49

Alexey

**Weekly, okay. [Ioannis chuckles] There are some leaders from these departments, and you talk to them saying, “Hey, what’s up? What are the current problems you have? How's it going with the previous projects we implemented for you?” And things like that. Right?**

10:09

Ioannis

Absolutely. The way I frame it is – I think of the heads of the different departments, from Digital Customer and Marketing as being my best friends in the working environment and try to understand how I can be supportive and how I can help them.

10:24

Alexey

**So how can you be supportive?**

10:26

Ioannis

[chuckles] Exactly!

10:29

Alexey

**What does it look like?**

10:33

Ioannis

Usually, it involves me getting enough business knowledge. If we talk about the Digital [department], it involves me understanding how, let's say, how the PPC advertisements work or how the SEO organic results work, and trying to understand what their aim is – which metrics they're interested in and what they do on a day-to-day basis. Then I see, “You know what? If we had a predictive model that could do X, Y, and Z, would that benefit you?” And then we have this kind of discussion that would essentially create some clarity on the business problem that we will then try to tackle.

# The importance of having business knowledge

11:15

Alexey

**I’ve heard the term “digital department” [from you] many times but to be honest, I have no idea what it actually means. It probably means different things at different companies, right? [Ioannis agrees] because different companies need to do different things. In your case, you mentioned PPC advertisement – I don't know what PPC is – Pay Per Click, right?**

11:37

Ioannis

Exactly. Pay Per Click.

11:39

Alexey

**So the digital department is also some marketing stuff, right?**

11:43

Ioannis

Exactly. Pay per click, if you think about it, these are the sponsor ads that you see on Google. If you go on Google, and you type “flights from London Gatwick to Berlin,” let's say, and you press “enter,” you see the 10 results that appear on the first page of Google. What you can see there first are usually the sponsored ads. These are the pay-per-click ads, as they’re known. The reason they're called “pay per click” is because there is an incurred cost every time a person clicks on that specific ad. We're trying to, in a way, optimize sponsored ads that appear on top. And we do the same thing for SEO results – we tag the organic URLs that appear which are usually below the sponsor ads. In a way, it’s an optimization that we're trying to do, so that the flights that we want to promote always appear on top and then, hence we can improve the conversion rate.

12:47

Alexey

**The other day, I was checking the cost per click in Google for keywords like “MLOps,” or “MLOps courses”. [Ioannis chuckles] Sometimes, for more niche words, it's like three euros per click, and then for more broad ones, it's like four or five, which was like, “Wow, is it that expensive?”**

13:14

Ioannis

Yeah, yeah. [chuckles] You have to bid on the right keywords, and then become relevant and all this kind of stuff that is happening in Google behind the curtains.

13:22

Alexey

**For you, as a lead data scientist, you need to figure out what these people talk about, like, “What does PPC mean?” “What do people care about?” “What is optimization?” And then, with this knowledge that you can extract from them (learn from them) you then go and share this knowledge with the data science team and you say, “Okay, these are the problems that these departments are struggling with. Let's think about how we can help them.” Right? [Ioannis agrees] And then you translate the problems into the language of data science and then, together with the team, you work on solving this. Right?**

13:58

Ioannis

Exactly. Yeah, absolutely.

# Getting projects to production

14:00

Alexey

**In addition to communicating with stakeholders, I think you mentioned other things – you make sure that projects reach production. What does that mean for you? Okay, you first talked with the stakeholders, you understood that these are the pain points they have – what happens next? What do you do next as the lead data scientist?**

14:23

Ioannis

As soon as I have the problem statement defined, we have an operating model within EasyJet that really helps us to understand, first of all, what the different steps are that we have to take to ensure that this resolution of the problem will reach production, and then we make sure that we adhere to all these different steps. There's a sequence that we follow. As a lead data scientist, I am accountable for ensuring that all of these processes are being followed. We make sure that when the data product reaches production, it will have the impact that was expected. And yeah, that's pretty much it in terms of my role. I can talk a little bit more about the framework if you want me to.

15:17

Alexey

**That’s quite interesting. What are these steps and what is this operating model?**

15:23

Ioannis

Yes, the operating model that we have, I think is one of the best things that we have created in EasyJet. I had a speech about that at the MLOps Summit. The operating model consists of different stages – I think it's four phases, if you will, that highlight all the different steps that we need to take to ensure that the model will reach production. The first thing is to get clarity on the problem statement, and this is pretty much my role. We like to call this a “single front door,” where we take a business problem or an idea into the funnel.

As soon as we do this, we have a meeting where all the relevant stakeholders come together and discuss the idea a little bit more. In attendance, you would expect people such as the business analysts and the finance team to understand the financial benefits that might be involved with the project, a lead data scientist, data engineers – every single person that needs to be involved in that specific project. As soon as we do that and we understand, “You know what? There's a real possibility of something good in this project,” we can take this on. We prioritize based on different ideas that have been submitted over time. And then we create something like a priority, “You know what? This problem is the most crucial one, so let's try to work on that first.”

As soon as we pick up a project, we will create the so-called “Definition of ‘Done,’” which is at the business understanding phase, where we try to understand a little bit more about the requirements that we need pick to make this project a success, which business KPIs we need to influence, improve, or increase or decrease, and how we can measure the benefits. For the latter, it means, let's say, I give you random numbers as an outcome, how do you know whether these random numbers are good or not? So we make sure that we create a document (the Definition of Done document) that highlights, “This is the data product. This is what production looks like. These are the benefits that are going to come about based on this calculation methodology.”

17:37

Alexey

**A large document?**

17:40

Ioannis

Not that large. Usually it's a single document – we have a template. You can think about two to three pages, tops.

17:49

Alexey

**Two or three, okay.**

17:51

Ioannis

Yeah. It's not that bad, I think. It outlines on a high level what things we need to make sure to deliver at the end of the day so that we don't have really much of a moving target, if you will.

18:05

Alexey

**I assume you have some sort of a template, right? A Google Document or maybe a Confluence page, and then you just copy this page and fill in all the things.**

18:16

Ioannis

Fill in the information. Absolutely.

18:18

Alexey

**And you do this?**

18:20

Ioannis

Not me, at this stage. I oversee the entire procedure, but usually, we would have a business analyst having workshops with the business stakeholders who are going to be the business accountable for the project. We try to capture every single requirement in this Definition of Done document.

18:39

Alexey

**Here, you don't talk about machine learning yet? It’s more about, “Okay, this is the project and this is the impact that we expect this project to achieve. This is how we measure this impact.” Things like that, right? You don't talk about machine learning at all at this stage. Right?**

19:02

Ioannis

Nothing at all. It just captures the definition of “done”. It captures just the “what” of the product, not the “how”.

19:11

Alexey

**There’s no discussion of the solution at all, right?**

19:19

Ioannis

Nothing whatsoever.

19:23

Alexey

**Okay.**

19:22

Ioannis

Because at the end of the day, we may have a document and we may realize down the line that it's not something feasible. We may know what we need to do, but after we have established all the requirements, we may realize, “You know what, the data is not actually there, which means that this is a no-go.” When that happens, although it doesn't happen frequently, this is a “fail fast” scenario. Then we say, “You know what, we cannot proceed with that. Let's take the second in line.”

19:48

Alexey

**But this happens later, right? [Ioannis agrees] At the business understanding step you come up with this Definition of Done document for a project, which is like two or three pages long, and then I guess you proceed to the next step, which is, as you mentioned, checking data and things like that.**

20:03

Ioannis

Exactly. As soon as everybody has signed off on this document – the business stakeholders, data scientist (which is me, in this case) , the data engineer, and every single person involved – then we proceed to the next phase. This is where the data science-y involvement starts to kick in – inception. You can think of it as the EDA (exploratory data analysis) where we try to ensure that we have everything that we need. That includes access to the data, if the data is already present, any GDPR concerns that we might encounter, exploring the data sources as in different distributions and these kinds of constraints that we might have. Yeah, that's pretty much it.

# The inception phase

20:54

Alexey

**At which stage do you actually…? You said that this is when data science kicks in. Is this the stage when you think, “Do I even need machine learning here or is it more like an analytical project?”**

21:09

Ioannis

Absolutely.

21:11

Alexey

**Okay.**

21:12

Ioannis

As soon as we kick off the inception phase, this is where the data scientists and analysts come together, and we brainstorm about the solution – we discuss the “how”. At this point, we understand whether this is a data science project that would involve machine learning or data analytics, or whether it's a hybrid between the two different sub-teams (data science and analytics).

To be honest, we do have some idea, when the business stakeholders discuss the problem, and we may have already decided at this point that this is a data science project or a data analytics one. But at the inception phase, we’re absolutely certain that, “You know what? This is 100% a data science project,” for instance. It’s just the confirmation that we have of when we started.

22:09

Alexey

**And depending on whether it is a data science project or not, I guess the next step would be different, right?**

22:15

Ioannis

Absolutely, yeah.

22:18

Alexey

**Then if it’s not a data science project, you say, “Okay, I'm a data scientist, I cannot help you,” and then somebody else takes this over, right?**

22:27

Ioannis

Not really. I’m accountable for both the data science and analytics projects. The only difference is that if it's an analytics project, the technical lead who will work on the project is going to be a data analyst instead of a data scientist. I still hold the accountability for making sure that the product is delivered end-to-end.

# Agile practices

22:48

Alexey

**So what's the next step? Or is it different for different projects?**

22:55

Ioannis

Not really. As soon as you have an idea and you have defined the “how” of solving the problem statement, this is where we move into the research and development phase. These are the hardcore modeling steps in data science, where we follow all the different design methodologies – sprint planning, stand-ups, retrospective – all the usual suspects are usually there, where we discuss all the different stories that we have defined in a Kanban board, for instance. We define sprints, “This is the goal for sprint one, sprint two.” This is where we start building whatever that solution might look like. We also make sure that the stakeholders are closely working with us because you have to make sure that… It's a common problem that we're trying to tackle so you want to make sure that the business stakeholders are part of the team and they're not just sitting around waiting for a delivery in three to six months’ time, depending on the complexity. So we make sure that we tackle that as a single team.

24:08

Alexey

**So that's why you have regular (at least weekly) meetings with them, right? You want to keep them updated on, “What is the progress? What is being solved right now? What stage are each of the projects?” Things like that?**

24:22

Ioannis

Absolutely. Also, at the end of every sprint, which is usually bi-weekly, we have a demo where we show, “These are the things that we have delivered.” And, if possible, we have an actual demo where they can get a sense of what we're building and influence some of the steps that we might take on the future sprint. They oversee the project from the beginning all the way to the end so they make sure that what gets delivered at the end of the day is something that they will end up using.

24:59

Alexey

**So I guess you also give them some sort of demo – a Streamlit App or something like this – that they can play around with so they see, “Okay, this is not what I meant.” Or “Yeah, this is what I need.**

25:13

Ioannis

Absolutely, yeah.

# The pilot phase

25:17

Alexey

**After the R&D phase, is there anything else?**

25:22

Ioannis

Yes. Then we have the pilot phase. In the Definition of Done, we have already defined the KPIs and the baseline that we're trying to beat. Usually, there's an existing “as-is” process that we're trying to beat with a new solution. Then we move into the pilot phase, which usually looks like A/B testing, where we test the “as-is” process compared to the “to be” process and ensure that the product that we have built improves the KPI of interest.

During that time, we also collect feedback from the business stakeholders because that can influence a second iteration of the product if needed. After the creation of the model, usually, it's the pilot phase, to ensure that we get the benefits that we were expecting. If that succeeds, then, I guess, it's deployment.

26:15

Alexey

**I’m just trying to come up with a joke about the “pilot phase”. [Ioannis and Alexey laugh] I’m not creative enough. [chuckles]**

26:24

Ioannis

[laughs] I know what you mean.

26:26

Alexey

**So okay – the steps are (the phases are): first, it's the business understanding phase, when we come up with this Definition of Done for a project. Then it’s the inception phase, where people actually… In the first step, you talk about the “what” and not the “how” but in the second step, you discuss their actual solution and you also decide if it's a data science project or more like an analytical project. Then, during the R&D phase, you work on the development – the research and development of the project. Then you also talked about how exactly you do this – all these agile techniques. At the end, there is the pilot phase, where you take what you developed and you see if the KPIs you defined in the Definition of Done are actually met. Right?**

27:22

Ioannis

Absolutely. Yeah, that's correct.

27:25

Alexey

**So those are the four steps that you mentioned. Is there a fifth one after the pilot? Like, the production part?**

27:32

Ioannis

It's usually the production. As you probably already know, “production” is a spectrum. Production might mean surfacing some insights into a Tableau dashboard, for instance. It can be some predictions being surfaced into an external tool. That can be all sorts of different things. Depending on what this means, we have the appropriate, let's say, production framework, which is still being developed at the moment. Of course, MLOps is certainly still at the beginning. But yeah, after we see that the benefits are already there and we beat the baseline, we roll this out to the entire market, depending on the project, of course.

# Other departments at EasyJet and competitors’ business models

28:18

Alexey

**The use cases you deal with are mostly related to marketing and similar cases – all these campaigns.**

28:27

Ioannis

Нes. Mostly Digital and Marketing.

28:30

Alexey

**So you don't try to work with the actual planes and the schedules?**

28:39

Ioannis

Not myself. But that's an excellent question because, as a data scientist, I look after Digital Customer and Marketing, but actually we have two or three more lead data scientists, where every single one looks after a different division of the business. So we have a lead data scientist who looks after Scheduling and Network, and another lead data scientist who looks after the Ops when needed, and, of course, Pricing and Revenue.

29:12

Alexey

**I noticed that tickets became more expensive after COVID. [Ioannis laughs]**

29:17

Ioannis

I have no idea about this. [laughs] No comments.

29:21

Alexey

**Well, you have a discount, right? [chuckles]**

29:24

Ioannis

Yeah. [chuckles]

29:25

Alexey

**I remember that a trip to Italy, before COVID, cost… Sometimes it was actually more expensive to get the bus that goes from the airport to the city than the actual ticket. These days are gone. Now it's more expensive to travel.**

29:46

Ioannis

Yeah, I guess inflation as well. Yep.

29:51

Alexey

**I was always wondering how companies like RyanAir can keep their costs that low – when it's like 10 euros for a ticket. But they probably cannot anymore because now it's different.**

30:01

Ioannis

Exactly. I think it's because of the different business models that different airlines operate under. There's a specific mindset that allows, let's say, RyanAir to operate with tickets that have an X price compared to EasyJet or Wizz Air – different competitors, of course.

# Utilizing Scrum practices in data science (the importance of MVPs)

30:21

Alexey

**You already talked a little bit about Agile methodologies that you use during the R&D phase and I was wondering if maybe you can talk more about this? How do you structure your day-to-day work when it comes to working on data science projects? In my experience, I remember… It was some time ago, and we tried Scrum. Maybe I'll take a step back. My background was originally a Java developer, and Scrum works well for well-defined developed software engineering projects.**

**But when it comes to data science, it's a little bit more ambiguous, because you don't know whether what you will have at the end (the thing you build) will work or not. In software engineering, it's usually less nondeterministic, let's say. Usually, you know that you will eventually build the thing that solves the problem, you just don't always know how long it will take.**

**When it comes to data science, you not only don't know how long it will take, but you also don't know whether it will actually work in the end. [Ioannis agrees] How do you structure your processes around this problem? You mentioned agile sprint planning and Kanban – so I'm curious to know in more detail how exactly you structure the work.**

31:51

Ioannis

Yes, absolutely. Of course, I was working as a technical lead (as a senior data scientist) which means that, now, as a lead data scientist, I don't schedule all the agile ceremonies. But as a technical lead, when I was a senior, I did have that experience. What I was following was all the different agile methodologies that have been introduced – I was making sure to stick with them. What you said about being ambiguous is actually true. Because in data science, you don't really know what you're building until you go and actually build it. This is when you realize whether it works or not.

So what we try to do to make the process a little bit simpler – to ensure that it's working – is we have the notion of MVPs (minimum viable products) which means that, in the Definition of Done document, we have the list of all the requirements that we know we have to build, which means that we kind of already have a sense of what we're building and which direction that we'll be taking. And because we know what we're building, it's a bit easier to estimate the time that it might take for us to deliver a single requirement or a single feature. That doesn't mean that we're always following Scrum – personally, I'm an advocate of Kanban, because of the complexities that have to do with data science and machine learning. But usually, we’re pretty good at estimating whether a specific feature is going to take, let's say, a week and a half. Even though we may not strictly follow the Scrum methodology, we actually have a Kanban board, and we try to put some timelines into our schedule to ensure that, “You know what? We'll have something built by the end of this two-week sprint.”

Of course, we do this with all the different agile ceremonies that we mentioned – we have sprint planning, which ensures that we have the different complexities allocated to the different stories. Of course, there are many ways to do that. At the end of the day, we do have some sense of how long something is going to take because of the notion of MVP, and we try to stick to these two-week sprints.

34:24

Alexey

**So you group all your work into these two-week sprints and at the beginning of each sprint, you do some sort of planning where you decide, “Okay, for these two weeks (for this sprint) we take this, this, and this. It will take probably the entire two weeks to do.” Right? And then during the week…**

34:51

Ioannis

Exactly, depending on the resources.

34:54

Alexey

**The resources are the people who work on this, right?**

35:00

Ioannis

Yeah. Something to add here, which also helps us estimate the different stores and how much they're going to take, also comes at the inception phase. At the inception phase, we dive into the data and try to understand a little bit about the quality of the data, how much preprocessing we might have to do, or how much time a specific implementation might take depending on the complexity of the project. The inception phase also gives us an understanding of how much time this specific implementation is going to take. That helps us estimate the timing a bit.

# A typical sprint at EasyJet and other Agile practices

35:38

Alexey

**Can you maybe walk us through the entire sprint? So, the sprint starts with planning and I think it ends with a demo – what happens in between?**

35:47

Ioannis

Yes. In between, we have daily stand-ups. Of course, it can be a written stand-up, or an actual 15-minute stand-up, usually in the morning, where the entire team comes together and we say, “I've been working on this story. This is the progress I’ve made so far. This is the plan that I'm going to work on today (or for the next couple of days). These are the blockers (if any) that I'm encountering at the moment.” Usually, when this happens, you have a senior member jump in to support – we make sure that all the blockers are removed so we can deliver the project or the feature on time.

Of course, depending on the complexity of the project, that can be an everyday stand-up or every other day – it really depends. But I think what works the best, according to my experience, is having two stand-ups per week so that it gives time for the people to work on the different stories. And, of course, if something goes wrong, you can always reach out to a teammate to ask for support. That's pretty much it in terms of stand-up. And of course… [cross-talk]

37:02

Alexey

**It’s not a very heavy process, right? What I understood is that you have this estimate – the start of the sprint where you estimate. Then you have some stand-up meetings during the week. Then, at the end, you have the demo. Right? That's basically the process. So it's not very heavy. [Ioannis agrees] Because I know in Scrum, there are all sorts of other things like grooming. I don't even remember what else, but I remember that the backlog grooming can get quite heavy if you follow the book and try to implement everything.**

37:42

Ioannis

That's true. But I think the notion of Agile is actually being agile and seeing what works for your team and what doesn't. We have tried with different meetings, according to what has been proposed over time. But we have identified that this framework that we have works great for our team and we follow this specific framework. One of the things that Ben Diaz, who is the Director of the Data Science and Analytics team, says is, “We have to be agile at being agile.” I think that summarizes everything. [chuckles]

38:17

Alexey

**What does estimating look like for you? Do you use something like PlanningPoker or things like that?**

38:26

Ioannis

It depends. Different teams use different techniques. We have T-shirt sizing, sometimes we follow the Fibonacci sequence to allocate points. We also have Scrum masters who support us in that way. We make sure that we don't use days as a way of estimation. So, whatever has worked for the different team members over time, it's usually the technical leader of the project who decides which method they want to use.

38:57

Alexey

**Yeah, interesting. So you do some sort of planning poker, right? Or?**

39:04

Ioannis

Yeah, yeah.

39:06

Alexey

**And what does it look like? I imagine that there's a meeting, and in this meeting, you have different people –you, a scrum master, project lead, data scientists can implement this, and then somebody (for example, you, as the project lead) says, “Now, let's talk about this task (this story) that we are going to take in this sprint, which is about changing the color or changing the chart on this dashboard (or whatever).” Right?**

39:37

Ioannis

Yeah, whatever that may be.

39:39

Alexey

**Everyone says, “Okay, I think this is a very easy task.” Right?**

39:44

Ioannis

Exactly, that you put that number on top. Depending on which one you think is the most complex, you put the corresponding numbers. Yeah, this is pretty much it. Every single team member… Of course, there are always outliers, but usually, you have all the different stories and you say, “Okay, which one do we think is the most complex one?” This gets allocated with that specific number, and then we increase the complexity depending on the methodology that we use.

40:10

Alexey

**Yeah, interesting. In your experience, does it work well?**

40:15

Ioannis

I think so. There have been examples where it has worked out perfectly and, of course, there are always [chuckles] the bad examples where you can see that you're quite tough when it comes to timelines. But I think the bottom line is that you have to adjust and be mindful of the fact that not everything is expected to go well on every single project. As soon as you manage your expectations, I think you're good.

40:49

Alexey

**When it comes to business stakeholders, I assume you don't invite them to your stand-ups, but you probably invite them to demos, right?**

41:01

Ioannis

Yes, that's correct. I think that's a great way for the business stakeholders to get a sense of what we're building because they can get an early interaction with the tool and the direction that we're taking. They also feel like a part of the team and that makes them more engaged in what we're building and quickly sense that we're a team and we're trying to tackle this problem together instead of us acting like consultants, “This is what we're building for you. Just use it.”

# Explaining results to non-technical people (the importance of soft skills)

41:33

Alexey

**I also imagine that the business stakeholders – it could be the Head of Marketing or Head of Digital, or some other Head – don't necessarily know what every C-curve means or precision-recall and things like that. [Ioannis agrees] When it comes to demos that are maybe a little bit more technical, they sit there and are just like, “Okay, I don't understand this, but I trust that you’re doing your work.” How do you deal with this – when stakeholders do not necessarily understand what the team is talking about? Or do you maybe educate the stakeholders, educate the team, or both? What helps?**

42:15

Ioannis

I think, in cases like that, you really have to be a chameleon and this is where soft skills come into place. When we have a demo session at the end of every sprint, we have to make sure that we never use technical language with them, because you have to adjust your context for a non-technical audience. I don't think there's been a single project where we have thrown some technical jargon, if you will, at all.

42:52

Alexey

**You educate the team members. You can say, “Look, if you say ‘ROC curve,’ they will be like, ‘Okay, what is that?’” So you teach them how they can present findings, the projects, and the demos, in a way that stakeholders will understand.**

43:14

Ioannis

Exactly. We never use any technical language with them. And if there's something that you need to explain that might require some technical knowledge, we always make sure that we use examples that can be easily interpretable compared to a technical implementation that you have seen. For instance, if you think about recommender systems and you want to understand how a specific person is closely related to another, you wouldn’t say, “As a measure of understanding how close two individuals are, we use the Keegan distance.”

Instead, you put forward two examples where you say, “You see that these two people look similar?” And you don't really need to define similar in this context, because they can see that all the different roles, for instance, look the same, compared to another individual that is completely on a different cluster. So when you want to explain these kinds of technical details, you can always use an example that would make sense for a non-technical audience.

44:19

Alexey

**Well, I assume that this is also a skill – presenting your findings in a way that non-technical people can understand. [Ioannis agrees] It can be even more difficult to learn this skill, to master this skill – let's say, even more difficult than learning machine learning, at least for technical people.**

44:39

Ioannis

Potentially, yes. [laughs]

44:40

Alexey

**People who are used to terminals and notebooks and all this stuff – going in and presenting something to business stakeholders might not be something that they're used to doing. So how do you educate people? How do you help them learn this skill or master this skill?**

45:10

Ioannis

I don't think there's an easy way. I think this comes with experience and just making sure that you always enhance your soft skills. One of the things that usually helps is thinking about all the different inner sentences that people usually say, “Pitch it to me like I'm a five-year-old.” Or I think Einstein had said, “If you can’t explain something in simple terms, you don't know it that well.” So, I guess it's just a matter of reminding people that the people that we have on the other side of the call don't have the technical experience that you have, so try to speak their language and explain what you're doing like you're speaking to a five-year-old. I guess there's no easy way to do this, it just comes with experience and constant feedback, of course.

46:04

Alexey

**And I guess having a five-year-old helps. [chuckles]**

46:15

Ioannis

Yeah. [laughs] I can only imagine.

46:17

Alexey

**Maybe if you don't have a kid who's five years old, you have no idea how much knowledge they actually have. [Ioannis agrees, chuckles] I have a son. He's seven years old. He sometimes asks me things like how GPS works. And I have no idea. Let's say if I go on the internet and type, “How does GPS work?” then the explanation would be super technical. Then I think, “Okay, how do you explain this to my son?” So it's a skill. Well, one hack I found quite useful is just asking ChatGPT. I guess everyone uses this now.**

46:58

Ioannis

Oh, yeah, of course. Absolutely. I still remember the days when ChatGPT wasn't out – I remember, I was a graduate data scientist at the time. I got the opportunity to present something to business stakeholders. I think this is when he found out, not in a nice way, that my ways of presenting and soft skills are not as good [as I thought]. I remember there was a really cringe moment where I was trying to explain why having 99% accuracy as a wider term doesn't mean anything unless you know about the balance with the labels. Yeah, I think it didn't go well. I think this pushed me a little bit to try to understand how I can present to someone who doesn't have technical expertise. I think it comes with experience at the end of the day.

47:45

Alexey

**Actually, we can think of ourselves as five-year-old kids too, when it comes to learning new things. For example, when I read this article about how GPS works, I'm clueless. Okay, there are a bunch of us that try to explain it, but I don't really understand what's happening there. So the explanation that ChatGPT gave to my son was actually helpful for me to also understand that. I don't know if I should say that, but maybe we can think of stakeholders as kids. [chuckles]**

48:23

Ioannis

[laughs] Yeah, I think I know what you mean. I'm really happy that all the stakeholders that we have at EasyJet are really literate in terms of data science and mathematics. That makes our work really, really easy. So I'm so thankful for that.

# Ioannis’ experience with the MLOps Zoomcamp

48:38

Alexey

**Yeah. Great. Also, I actually wanted to spend a bit of time talking about the MLOps Zoomcamp course, because I was…**

48:47

Ioannis

Yeah, of course!

48:49

Alexey

**I was really surprised when I looked at your background – I thought, “Why would Ioannis even consider it?” Because with your experience – you're already doing all the things you talked about right now – I'm wondering, what inspired you to take our course? Why did you decide to take it?**

49:10

Ioannis

Yeah, absolutely. The thing is, as a lead data scientist, my role has become a little bit more managerial compared to the amount of time that I have to spend doing technical stuff. And if you ask me, having a bachelor of mathematics, I'm a geek at heart, which means that every opportunity I get to get my hands dirty with some data and build something myself – I always take it. MLOps specifically is, from my experience – I'm usually involved in, let's say, building the models and I didn't get much exposure to the productionization side of things. I was just intrigued by the course and the content. Of course, I was using MLflow, but then we had Prefect – the data engineering team – and we have been using airflow. And I'm like, “Let me get into that engineering side of things a little bit more and also get the opportunity to get my hands dirty.” I think this is what clicked for me. And I'm like, “Yeah, let me go for it.”

50:22

Alexey

**Well, as somebody who was a lead data scientist in the past, one problem for me was always time. [Ioannis chuckles] With all this stakeholder management, how do I actually find time to still be hands-on and experiment with things? [Ioannis agrees] And then sometimes, I wanted to take a course, but then I didn't have time, because there’s only 40 hours that you spend at work. How did you solve this problem?**

50:53

Ioannis

Yeah, that's a great question. I think one of the good things about my decision to become a data scientist is that I genuinely love the profession. I would be a data scientist as a hobby if my day job was something different. This means that even when I finish my work, I don't feel drained from all the information that I had to go through throughout the day.

I genuinely enjoy working as a data scientist, which means that I consider that as an activity rather than, let's say, something that will consume my time. So yeah, it was just great. I had my morning cup of coffee, and during the weekends, I took my laptop, went to a nice coffee place and just watched your courses and tried to do the assignments. It's been fun. And I got a little experience out of it, to be honest. So yeah, it was just great.

51:52

Alexey

**So instead of watching Netflix, you watched the courses.**

51:57

Ioannis

What was that?

51:58

Alexey

**Instead of watching Netflix, you watched the courses. Or… Maybe in addition to.**

52:01

Ioannis

Yes! [laughs] Absolutely.

52:04

Alexey

**Okay. Well, it sounded like the course was useful for you, right? Was it mostly like… I don't know if I should call it that – entertainment? Or more like self-educating? Or did you also get something out of this course and apply it at work?**

52:24

Ioannis

It was a little bit of both. It was entertainment in the sense that I got confirmation that what I'm doing is correct. But also, I got the opportunity to play with technologies that I otherwise wouldn't have time to. One of the examples is Prefect, for instance. Because as a lead data scientist, I’m not that involved in the engineering side of things, so I wouldn’t get the opportunity to play with Airflow or Prefect. So I think it had a good balance of both – getting the confirmation that what I'm doing is correct, but also learning something new. This is really important because as you mentioned in the beginning, I'm leading the MLOps team within EasyJet. Even though I give the guidance and have an influence on where we're going as a data science and analytics team with our MLOps journey, it was great for me to understand a little bit about the technical landscape. I feel that that's the best way to influence a specific direction. So that really worked well.

# On Evidently

53:33

Alexey

**Actually, before our conversation (before our interview) I had a chat with Elena from Evidently and she said, “Oh, Ioannis is coming to your podcast? Make sure to ask about Evidently!” [chuckles]**

53:48

Ioannis

Absolutely. Evidently, I think – and I'm not afraid to say this, but I think Evidently is the best Python library out there for model monitoring. This is something… the final assignment that I did for the MLOps Zoomcamp also gave me the opportunity to play with the Evidently library a little bit more. I had the time to play with Evidently, I think, two years ago, when it was still, in a way, the dev version. I remember the first time that I reached out to them, because I said, “You know what? I have implemented that and it doesn't look correct.” There was actually a bug and this is how the networking kicked in. But yeah, Evidently – absolutely the best Python library for model monitoring.

54:40

Alexey

**Do you use it at EasyJet as well?**

54:43

Ioannis

Absolutely. We will use it to their sense of embedding that within our MLOps framework. It's still a work in progress but we have made tremendous progress throughout all these years. I think, especially now that we're trying to define our MLOps capabilities, Evidently is the best thing that could have happened to me and to EasyJet to that extent.

55:11

Alexey

**Just curious – I know Evidently, right now, has its own dashboard, but what you do is probably based on some sort of other monitoring framework, like Grafana or something like that, right?**

55:25

Ioannis

Yeah, I mean, right now we're thinking about using the Tableau dashboard and I have a proof of concept that I'm about to present to the EasyJet MLOps team. But before that, because I had already implemented a proof of concept, we weren't using Grafana – we didn't have the UI. To be honest, I had implemented a custom function that would trigger an email alert to the technical lead of the project in case there was data drift or model drift detected. It was, I think, two to three years ago.

56:01

Alexey

**You mentioned Tableau, and it's interesting how versatile this tool is. [Ioannis chuckles and agrees] It's not just a dashboard, you can even build simple, rudimentary monitoring in Tableau. I remember we had problems with data quality and then our analyst quickly came up with a dashboard that shows how many records there are each day in the important tables. Then, what he did next was configure Tableau to send an alert if the number for one of the days was less than expected. He did that in like 30 minutes or something. That was amazing.**

56:44

Ioannis

Okay. That's great. It indeed sounds amazing. Goodness.

56:47

Alexey

**I mean, at the end, it's just a bunch of SQL queries and then knowing where to put these queries and which button to click to create an alert, he knew how to do this. Not everyone knows that. But it was a quick and dirty solution that worked pretty well. It's amazing.**

57:03

Ioannis

Yeah, that's good. It's always exciting when someone delivers something that fancy in such a short period of time.

# Ioannis’ resource recommendations

57:09

Alexey

**Yeah, I think we should be finishing soon. So maybe I'll ask you one thing. We talked a lot about communicating with business stakeholders, we also talked about Agile processes. We talked a little bit about MLOps. Are there any good resources that you can recommend to our listeners who want to learn more about these topics?**

57:36

Ioannis

About which topic specifically?

57:39

Alexey

**Well, about any of those that we discussed – let's say, about processes, about communicating with business stakeholders? When you were learning how to do your job well, maybe you came across some books or courses that helped you.**

57:58

Ioannis

There is a single resource that I would recommend to every single aspiring data scientist/data analyst to watch out for. I'm not sure if you know Cassie Kozyrkov – she’s the Decision Intelligence Advocate for Google, at least she used to be – she resigned. But Cassie Kozyrkov and her course on YouTube, Making Friends with Machine Learning, I think, is the best resource out there, in order to understand how you can communicate technical details to a non-technical audience. I think the way she speaks and expresses these kinds of technical details in such a nice and direct way, is one of the best skills that someone can get. And I think, watching her YouTube videos helped me to really understand “What would be the best way to explain a technical term to someone that is not familiar with my world and data science in general?”

I spent, I think, countless hours watching her videos, trying to analyze the way that she approaches things, terms, or explains how linear regression works. So if you want, Cassie Kozyrkov from Decision Intelligence from Google – her YouTube videos, Making Friends with Machine Learning. At least this is how to communicate to a non-technical audience. When it comes to technical details, I think different books like, Pattern Recognition from Gibson is one of the best books that you can go with. It's really heavy, so you have to make sure that you're comfortable with mathematics.

59:44

Alexey

**In many senses – because I remember we used this book for my machine learning classes and it was heavy for the class too. [chuckles]**

59:54

Ioannis

It was heavy, indeed. But I'm telling you, if you spend time and you actually focus – let's say you have a two-hour block of time and you go through that, it's one of the best things that you read to understand the mathematics behind machine learning and how it really works. Of course, LinkedIn helps a lot with different posts and resources that are being recommended. I think on a day-to-day basis, LinkedIn is my go-to resource website.

60:22

Alexey

**Cassie… I think this is how I know her – from LinkedIn. I don't know if she's active anymore, but she used to be quite active on LinkedIn and this is where I went to see her content.**

60:36

Ioannis

She is amazing, yeah – podcast, YouTube, LinkedIn, of course. I think she was all over the place. I think now she's building something on her own. This is why she left Google. And I'm really interested to see what this is going to be. I know this is about decision-making and decision intelligence, which is something she has established on her own. So yeah, I'm really looking forward to seeing her content.

61:00

Alexey

**Yeah. Thanks, Ioannis, for joining us today, and for sharing all that you shared with us today. Yeah, it was amazing. Thanks for finding time. And thanks, everyone else, too, for joining us and being active here. I think… I actually forgot – we had only one question that I accidentally forgot to mention. Is it okay, Ioannis, if Dave reaches out to you on LinkedIn and asks this question?**

61:34

Ioannis

Yeah, absolutely. I'm always open. I'm super active on LinkedIn. Any question, whatever that may be – feel free to reach out on LinkedIn and I’ll make sure to get back to you.

61:47

Alexey

**Okay, thanks. And with that, I guess we’re finished.**

61:51

Ioannis

Amazing. Thanks for having me!

61:54

Alexey

**Yeah. Thanks. Bye, everyone.**