1:10

Alexey

**This week, we'll talk about hacking your data career. We have a special Marijn. Marijn is a sociologist and a data scientist. He works at Capgemini. Did I pronounce it correctly? I had to look it up.**

1:24

Marijn

You pronounce my name correctly. Everything's fine.

1:28

Alexey

**I trained. [laughs] Marijn works at Capgemini as a data scientist manager. I think the last time we had you at our event, you were a senior data scientist.**

1:41

Marijn

What comes after “senior data scientist”? I don't know. So now I'm a managing data scientist, which is great because I get less recruitment spam now – clearly, everyone just Ctrl F's “senior data scientist” and that's it. I don't fall into that category anymore. So easy.

1:58

Alexey

**You have probably seen Marijn talking about doing good with data, for example, in our events. In general, you have probably seen him on LinkedIn posting stuff about data and other topics – cats and other things. Welcome to our event, Marijn.**

2:17

Marijn

Thank you for having me.

# Marijn’s background

2:19

Alexey

**Yeah, I had to train myself a little bit with your name, because you were already at our event last time. Before we go into our main topic of hacking a data career, let's start with your background. Can you tell us about your career journey so far?**

2:35

Marijn

Sure thing. It started like nine years ago – that's painful to admit – when I started out studying sociology. My dad told me, “Man, you're never going to get a job like that. Sociology? Who does that?” But what I learned during my studies was how to predict election outcomes, model Twitter data and discourse, how to mine social media – that kind of stuff. Because even though a lot of social science focuses on human behavior qualitatively – like interview surveys, etc – you can also do so quantitatively, which means I was modeling pandemics even before COVID was cool. I did Ebola. Good times. That one we managed to contain.

For my graduation, I was modeling how many people get stabbed in my hometown of Rotterdam and what main neighborhood circumstances could predict lower or higher numbers of stabbings and drug crime and such. Then I met Capgemini and the entire field of data science told me “Man, that's cool. That's data science and AI.” I thought it was just statistics and distinctively not sexy. I said, “Sure, now please give me a job.” And that's how I became a data scientist.

3:53

Alexey

**That was very short. [laughs] I got comfortable listening to your stories because you have a good way of telling stories.**

4:02

Marijn

I can go on for a few more if you want me too. But the main point here – and that's also something we're going to talk about a lot today, I think – is that I stand out in the field of data science because of my background. But that has little to do specifically with social science and more with that your background actually determines a lot of your competitive advantage in data science. It's not about *knowing the same stuff* as the next guy or girl, it's about what you know that they *don't* know. I see tons of kids these days shitting their pants as they look at the labor market and they try to know the deep learning algorithms better than their peers. That is a tough thing to compete on. Try to compete on the things that you know that *they don't know*, rather than the same things. That's how I got lucky.

4:58

Alexey

**So what do you know that they don't know? Sociology?**

5:03

Marijn

I did a mixture of sociology, criminology, and quite some statistics. I remember one of my very first projects we had – we were modeling the effects of gender and my computer science and AI peers (because we work in a team) were like, “Okay, how do we test that?” And I'm like, “Oh! Checking for gender bias! I know that! I studied that crap for four years!” So yes, I studied bias and discrimination and algorithms for four years and was told I'd never get a job. I hit the labor market and everyone's crying about bias in algorithms and I’m like, “Hey, what happened?”

In the same way, I have economists, like astronomists or astrologists – one of the two, I always mix them up and they get angry over that – physics people, there's one who did psychology, theoretical mathematicians – all part of the teams I lead these days, and we *all* see different problems. We also see different solutions. But by screaming and arguing at each other, we find *better* solutions. The opposite of that is what you see in most organizations, where their entire data science team is just 10 people who studied econometrics or 10 people who did computer science – and there's nothing wrong with either study – but if you do econometrics, you're kind of stuck soon as you have to do NLP, to give you an example.

6:38

Alexey

**Does it happen because of the interview process? Let's say the person who does the interview asks just what they know, and then this way…**

6:49

Marijn

If you think about marketing and recruitment bottlenecks – this process is crazy, because we only grab the specific words that we think we need for our team. This is not a data science problem. This is an IT and the entire labor market-wide problem in that we recruit based on boards, not actual capability. The same reason tons of people only recruit data engineers these days, assuming they also have data science capability, because they previously just recruited everyone who said “data science” and then found out they couldn't do the engineering part. Meaning we're now doing the *same* faulty mechanism, just with another word in our recruiting and find data engineers who can't do the data science part.

7:34

Alexey

**Yeah, but how can you find them?**

7:36

Marijn

We keep depending on V lookup, basically, for our recruitment.

# Standing out in data science

7:42

Alexey

**You mentioned that people need to find something they know that others don't know. Since you mentioned that – this wasn't a part of the questions I prepared, but maybe you maybe you can help with that. A few days ago, I think it was yesterday, somebody in DataTalks.club asked this. The question is, “I'm trying to find the most complete curriculum for data analytics, data science, and machine learning. Each site has different things and so on.”**

**The question goes on to find the perfect curriculum. Here, you say, “Hey, forget about all this stuff.” Right? If you learn deep learning, then everyone else also knows deep learning. So what would you advise to this person? What would you suggest to this person? Their goal is to find a data role – what would be your suggestion?**

8:31

Marijn

Striving for perfection is something you do in academics and in science. As soon as you step outside of that, you realize there's no such thing. There is *no* perfect model, there is only “good enough.” Now do you realize that data science is basically three things: statistics, programming, and a field you apply it in. To my great surprise, the field and the statistics, I already studied. Just like a computer scientist has programming and statistical math activity. But you start with one of the three and then you learn the second one through a minor (or I didn't know) and the third you just kind of wing.

There is no perfect curriculum, just like there is no consensus within science. There shouldn't be. Science agrees that this is a very dangerous statement to make. In the same way, your “perfect curriculum” at the micro level (because you can have a local one) very much depends on the field you're coming from. Now, this relates to my previous statement of “Don't double down on the things you don't know. Double down on the things you know that give you a competitive advantage.” Sure, I try to improve my programming. I will never be as good a Pythonist as some of the computer scientists in my team – also because I still prefer R (I'm sorry). But I don't try to compete on that stuff.

No data scientist, except a few *unicorns* are complete experts on both programming and statistics, as well as – I don't know, finance or outlier detection or whatnot. You grab the part you're good at and you improve on that, because that makes you a valuable asset and a valuable addition to a data science team. The only alternative is an organization where you are the only data scientist and you have to do all three, which is very educational, but also makes you lose hair. It also means that you're working for an organization that really doesn't understand a damn about data science, because *one data scientist* is not a *data science team*.

10:51

Alexey

**Okay. So, if I summarize – don't strive for perfection and just double down on things you know. That will give you a competitive advantage. For you, the competitive advantage was your social sciences background, right? You already had some experience with – I think you mentioned quite a few things –election outcomes, modeling, and finding the district in Rotterdam with [cross-talk]**

11:16

Marijn

For example, yes. We did big time. We found some really big issues in a model when we were predicting burnouts. We had to interview a ton of people for that – to check if our findings for biggest predictors of burnouts were actually valid. Everyone was like, “How the hell do we interview people?” And I'm like, “Oh! I know! I studied this and I interviewed a ton of people.” They were mainly criminals. But as it turns out, interviewing people without bars in between is basically the same, just slightly more scary. Which, again, was a thing that I happened to know that made me a valuable addition to the team – as you know a greenie.

# Doing the opposite of what people tell you

12:05

Alexey

**Okay. [laughs] I came across your post on LinkedIn. I reached out to you multiple times inviting you for the podcast. Then I saw this post, and I thought, “Okay. This is what I want to talk to Marijn about.” And the post was about “how to hack your career” and this is actually the name of the event we have.**

**First you wrote “How do you hack your career?” and the answer was “Do the opposite of what people tell you.” So can you tell us more about this post that you made and what exactly did you mean by that? Maybe you can also give some examples.**

12:44

Marijn

I think my post started (or ended) with the quote, “Why try to fit in when you were born to stand out?” As far as career advice goes – and this is more than just data science – if you're, I don't know, a florist or you're into accounting, (maybe not accounting, but you get what I mean). Everyone wants to make you “fit in”. It is natural behavior of us upright monkeys – to do the same as the rest does. But if you *want* to get that job, or that assignment or that project or you want to excel – if you want to stand out – then why the hell would you do the same thing as the rest? Oh, let me guess. You're going to try to do the same thing as the rest who stand out? That still means you're still mimicking other people.

I started out the whole LinkedIn game a few years ago, trying to do whatever all the popular people were doing. More than half of it really didn't work because I wasn't doing anything authentic. I wasn't doing anything “me.” I was just mimicking some other monkeys rather than that even other monkeys. Now, I'm pretty sure there are a lot of students here listening in on this in addition to data scientists. If you're in a team and everyone's everyone delegates tasks – because we have hierarchies and we have top-down management and everyone shits down – I learned that rather than waiting passively to get your assignment and your tasks, it really helps *you* and *them* to be proactive in that. If you are the *first* in the room to say “No, I'll do this” and you pick up the task – first, you get to choose which task you get to do. So not the sucky one – that helps you. Secondly, you're standing out. You're the one they'll go to next time because you were *proactive*. And that is scary, especially the very first time when you are fresh out of college, on the work floor, but it gets easier the more you do it.

I have a story about my hair, I think you mentioned it previously. I am known as the guy with, you know – beautiful hair. On LinkedIn as well. But one day I had to call the boss. I'm not gonna mention names, but I have to mention that I had to call the boss of the company I work for in the Netherlands. It was an escalation, it was an important thing. His phone number was just in the system, but you have damn good reason to call him. Because he'll pick up. And I said, “Hi. This is Marijn Marcus speaking.” And he said, “Who?!” I'm like, “[whispers] oh, crap. Then I said, “The long haired guy who talks AI.” And he said, “Oh, you! How’s it goin’?” That's when I realized that my long hair is my trademark. It has *nothing* to do with data science/AI, but it makes me stand out. Now, that doesn't mean y'all need to grow out your hair. But it *doe*s mean you can own and should *rock* the things that make you unique, rather than just trying to fit in. Because that never worked for me and it won't work for you either. You can spend a whole lot of time, money, and effort on being just like everyone else. But I don't think that makes you happy and I don't think that makes *them* happy unless they really like matching ties or something like that. That goes for the rest of the LinkedIn post as well, I think.

16:36

Alexey

**But maybe about this “delegate task” instead, your suggestion is be proactive. First, pick the task, which gives you an opportunity to choose things. So it's like, there is a meeting and then somebody says, “Okay, we need to do this thing.” And for that thing, I imagine there is a senior data scientist who decomposes the task, the product manager says, “We need to do this.” Then the senior data scientist says, “Okay, for this, we need to do this, this and this.” And you can say, “Me, me! I want to do that!” Is that right? How does it happen usually?**

17:09

Marijn

This also relates to our previous topic of the “perfect curriculum,” because practice is better than any amount of training. Because I was proactive back in the day – in my first few projects – “Okay, sure. I'll do that. Then maybe *you* can do *that*.” Half a year later, I was being asked “So what management skills do you have?” And I realized, “Oh, hell. I was basically doing the scrum master thing over there at some point.” *Everything* is a learning opportunity. Working in teams as well. This is in part about being proactive and it's also about picking up opportunities to learn, like managing teams, distributing tasks, being a product owner, being the go-to person for a specific topic.

Especially for the young people – also the old people in the field – you want to keep learning. If you don't want to keep learning then why the hell are you in this field that keeps changing so much from year to year? And every new project that I do, I see as an opportunity to do something I haven't done yet. That's always scary at first. It gets less scary, the more you do it. But it's a waste of you if you don't grab that opportunity to learn. Because if you're not learning, then why are you doing it? Just for the money? Then you should quit your job and become an accountant. Though, they'll all be automated away in 20 years, I'm pretty sure. Don't be an accountant.

# Don’t shoot the messenger (carefully sharing your findings)

18:51

Alexey

**Yeah – some accountants are very difficult to automate, I think. Anyways, we already talked about everyone delegating tasks. Then you also mentioned being proactive, right? What was next on the list?**

19:06

Marijn

Oh, you have to list? Well, there was stuff like…

19:09

Alexey

**[laughs] Or just, what else was on your mind?**

19:12

Marijn

Oh I was going to give a terrible example – a lot of projects, especially in data science/AI, focus on forecasting, because people want to know the future. Let's be honest. But to know the future, you need to know the past. That's why, at first with feature engineering, you're basically usually researching what happened in the past and why. Some of the biggest causes of events in the past – I'm gonna use the burnout model again – I found a lot of people were having burnouts because of a few specific managers. Because you don't quit your job, quit your manager. So we found that some of the biggest predictors were a few sets of individuals. And I announced that.

People said “No, no. You shouldn't mention that. You shouldn't mention it.” I'm like, “I *should* mention it, because if we're going to try to predict who gets the burnout next, we need to be able to explain how the model works. Otherwise, they won't trust it and they won't use it.” Explainable AI is actually just all about regression and decision trees, and explaining how your model works and modeling the past. Again, I happen to have studied that stuff. To me, it was new that you can turn the model around to predict the future. But everyone told me not to mention that. Because then I was presenting to management that *those* four managers were actually causing a lot of attrition and a lot of burnouts. Those managers were in the crowd. So I got my ass kicked on that project, clearly. It all blew up in my face. *But* the rest of management was listening. Upper management loved it. Middle management hated it, because I made everyone laugh at them. I pulled down their pants. But *because* I mentioned that, later on, upper management actually acted upon those insights. I said what anybody else dared to say.

Now, there is a high chance of being the messenger that gets shot in data science, because you're the first to stumble upon bias, discrimination, all kinds of stuff. You know the Amazon recruitment case, where humans are actually I think, discriminating. But then they trained an algorithm to recruit people and found that it grabbed only white males. Then they called the algorithm evil and they went back to human recruiting, which is probably just as biased, but it's not expressed in data – so it's “okay”. We encounter a ton of this stuff, which is fun to me. This is why the field is fun to me. But it's also why the field is so valuable. If you don't say this stuff, you're not doing your job right. You should say it in a careful way, like “No, I anonymized the data.”

That's what I did a few months ago when I was analyzing a large company regarding which parts of the organization had pay gaps – like gender pay gaps, that kind of stuff. I didn't mention anyone by name. I just gave them the hashes. “Oh yeah, I found out that *these* bastards are the most sexist in pay gaps.” Okay. “Oh, that's so bad. That's terrible. Who is it?” “Well, I don't know – because it's all anonymized. But here is the hashing list, so you can look it up.” And then they looked it up and they got angry at them. So then it wasn't me that was the messenger who got shot. You get better at telling this stuff in a way that doesn't kill *you*. But it is so valuable and so important to mention this stuff. Because working in a data-driven way is still new to so many companies – so many organizations. And it's a fun part. It would be a waste not to. Same as it's a waste to keep your head down.

23:09

Alexey

**I do have a list actually. I have a post in front of me. One of the things I have is “Instead of getting told not to ask questions, ask anyway.” Was this something you already covered, or no?**

23:25

Marijn

I was told not to question by everyone why certain managers relieving – regarding burnout. So you should.

# Advising the seniors

23:37

Alexey

**Yeah, the next one. Or one of the next ones – “Don’t be advised by a senior with 20 years of experience. Instead, talk back and advise the seniors.” Do you have a story about that?**

23:48

Marijn

That's an easy one. That's just hierarchies – pyramid models of organizations – where the older you are, the more you know. I think IT struggles with that really badly because there's also a ton of people who are woefully outdated in their knowledge and a lot of them don't want to admit it. But a lot of them also like it when you teach *them* stuff. It doesn't mean they don't want to learn. But you also have legions of people who just nod and say “Oh, yes, sir. That's great. We should continue with this horrible legacy application idea. Yeah, landscape. Yeah, sure. Sure. By all means. It's the safest.”

Then the entire damn thing burns down because it was all legacy. Don't tell people they're stupid. I learned that. I do that way too often. But give them an opportunity to learn. This is how I became friends with CTOs of multiple organizations. It was an accidental reply-all mail, I'll admit this. I hope you're not listening. “Data science, AI blah, blah.” And I said, “Well, actually no. You got that and that and that part wrong.” It was a reply-all mail. Oops. But it did get me coffee with the guy in the end.

25:09

Alexey

**So what did you talk about over the coffee?**

25:13

Marijn

Well, we slightly skimmed over the fact that he misunderstood some stuff and then we just talked about what else we were doing. A lot of people – let's be honest, we're all nerds here – we're very much interested in the stuff people do. “Oh, how did you do this? How did you do that?” In intakes. I am barely ever mentioning what perfect curriculum I followed or what degrees and certifications I had. I mainly just talk about “Oh, yeah. I encountered that problem before.” “Well, how did you deal with it?” “Well, the data quality sucked. Like this and that and I fixed it like that and that.” “Okay, and then what?” “Yeah, then I insulted the managers to their face. I learned from that. And then okay.” You're telling stories. We all tell stories using data because stories and experience are so much more valuable than Pokémon badges?

26:03

Alexey

**Than what?**

26:05

Marijn

Than Pokémon badges.

26:07

Alexey

**What is that?**

26:08

Marijn

You never played Pokémon? Come on.

26:09

Alexey

**I've watched the cartoon, but not…**

26:14

Marijn

Oh, man, you missed out. I'm so sorry for you. In Pokémon the game you collect badges. You collect certifications.

26:22

Alexey

**Is this a card game, or like a video game?**

26:24

Marijn

It's the video game. Never mind. I think the audience will get it. Comment, like, and subscribe if you do. Or if you don't.

26:33

Alexey

**Maybe I'm too old for that. But I think yeah, it was like five, six years ago that people were obsessed with getting Pokémon on the street. It was like some virtual reality thing.**

26:43

Marijn

It was like 15-20 years ago, but let's not go there. I already feel old enough as it is.

26:49

Alexey

**[laughs] Okay. Actually, coming back to your advice about teaching them stuff instead of getting advice from more senior people – when I work with juniors, I always admire their enthusiasm. They come with a fresh look, they're very enthusiastic, and they really want to learn. I don't have that kind of energy, right? I've been in the company for almost four years, now – three and a half. I don't have a fresh look.**

**This is very valuable for me when somebody comes to me and says, “Hey, what do you have here?” People may use different words, some of them may say “Okay, this is crap.” Or maybe “There are other ways of doing this.” But this is really great when people can tell me, “Hey, let's try this thing.” Or “Okay, this is what I did in my master thesis. This is a cool thing. How about trying that?”**

27:44

Marijn

It’s an interaction, right? You learn from them and they learn from you. And that is how it should be. But we're sooo scared. I was so scared at the start, because it's like talking to a professor. You don't dare to disagree with them. Well, it's exactly just like in academics, actually. The professor *also* likes it when you discuss stuff with them. That's the *science* in data science. Come on. It's research. It's discussing. And it's trying to prove each other wrong. We just use a lot of data to do it.

# Bite off more than you can chew, then chew

28:16

Alexey

**Was there anything else in that post? I think we covered pretty much everything, but I'm sure you have a couple of more stories.**

28:23

Marijn

My favorite last one is the “bite off more than you can chew, then chew” because otherwise, how the hell are you going to learn how much you can chew? I have a poster like that in my house somewhere with a wolf. It’s great.

28:40

Alexey

**So if I translate it to plain English – does it mean “take as many tasks as possible and then figure out if you can handle them or not? And then back off?”**

28:52

Marijn

Dare to pick up tasks and dare to try stuff. You will only find out what you're good at, what you like to do (two very different things) where the coffee is okay, and where your limits are, by trying stuff and seeking out those limits. You're not going to do that by calmly sitting in the corner and waiting for them to give you a Vlookup task. This happens to a lot of data analysts as well. They get hired somewhere and go “Yeah, go do this thing in Excel,” or “Go do this very simplistic analysis.” The cool ones then go like “Okay, I can do that. We can also run a time series model to automate that stuff.” Because they don't know that – you have all the knowledge. They have the business intel, blah, blah. And you're here to teach them as much as they are there to teach you.

29:49

Alexey

**I'm looking at the live chat and three people said “We got it. I get it.” That's probably for this Pokémon metaphor. I'm really wondering what that is now. After the episode, I’ll need to look it up.**

30:00

Marijn

Oh, what? I can't see the chat. Is that on..?

30:03

Alexey

**It's on YouTube.**

30:07

Marijn

Oh, YouTube. Okay. I’ll look it up right now.

# Marijn’s side projects (finding value in doing things you find interesting)

30:10

Alexey

**[laughs] Three people say that they got it. While you're looking it up, one thing I know you for is that you always find these little pet projects that you do and then you tell the world about them. And these projects are so cool.**

**Can you share a few ideas or a few projects that you did and then maybe some project ideas with the audience? Maybe not everyone knows you, so can you tell first about the projects you did – something awesome. You probably have something right now. Do you?**

30:47

Marijn

Yeah, it's loading right now. So I'm going to show you my house, sort of. This is my Home Assistant. It's running on Raspberry Pi 4 in my closet. I can control my entire house through it. I'm sorry for those listening in because you can't see it. You see a dashboard on my phone with all my lights, my coffee machine, my front door – the Christmas tree is offline as you can see, because I kind of disconnected it.

31:15

Alexey

**Because it’s not Christmas time, right?**

31:18

Marijn

Yeah, give it a few more months. It's all smart LED lights. I'm like really proud of it. You can see it on LinkedIn. My gas is obviously off right now and it's not going on for a few more months. Coffee Machine, stereo blah, blah, blah, amount of COVID cases in the Netherlands, my balcony lights. More importantly, these are my plants. I add the sensors to my plants – to check their humidity, light, temperature, the poopy icon on screen shows if they need fertilizer. I think this one definitely needs fertilizer. There’s also the battery icon because it's just sensors in the plants broadcasting over Bluetooth and the Raspberry Pi picks up those, writes it to a database and this dashboard then picks up the latest numbers and actually gives me audio cues (my plants scream at me) when the threshold of “oh, you need to water your plants” is reached.

I remember being in a big conversation with a big agri firm. They were like “Oh, what can you do with sensors and AI?” And everyone was just talking out there ass about “Oh yeah, sensors, AI, IoT!” And I just literally pulled this thing out of my pants and showed them and they were convinced, “Okay, Marijn knows plans.” Also, this one is tracking the local police radio to show me where the stabbings are in my neighborhood (also fire) so I can go watch… in case you know, “fun fires” are happening near my house. And yeah. That’s stuff I do at home. But then – here comes the important parts everyone – then I'm working and it turns out this is valuable experience. In the same way that I was asked “Hey, what do you know about social media analysis?” and I just started talking about an e-purse I wrote during my studies in 2013.

Yes, I know it's a long time ago. But I was scraping Twitter to research the refugee crisis discourse. That counts as valuable work experience as well in the same way hobby projects do. So please, youngsters over there, don't think you have no work experience because you just came out of university. All your internships, your thesis, your hobby projects, the LED lights you glued to your piano (cuz I know I do). They. Count. Too.

33:56

Alexey

**What happened to your plants that you had to make a system to water them? Did they just die one day and then you thought “Okay. How can I make sure that they don't die again?”**

34:11

Marijn

Basically. Some were overwatered, others were underwatered – because they all have different thresholds. On the backend, it knows what the type of plant is and what the threshold is, so they also scream at me if I *over*water them. But believe me, that occurs less often – because it's a rare event.

34:31

Alexey

**So you have some sensors in each plant, right? These sensors send data over Bluetooth to the Raspberry Pi in your closet, right?**

34:40

Marijn

Yeah, the Raspberry Pi stores it because they scream their new data – it refreshes every minute. It stores the latest values and I can connect to that and trigger events happen if thresholds are reached.

34:53

Alexey

**And the app you showed? Did you write it yourself?**

34:56

Marijn

I got a question from the audience, “Are you not worried about cyber attacks?” I'm not worried about them cyber attacking my plants because the sensors only send data – they can't kill my plants through it. The rest of the rest of it all run Zigbee, so it's all localized. It's not Wi-Fi, it's not Bluetooth and it all runs through Iceland. So I'm pretty secure on that end. Like really, if you're building IoT, go with Zigbee – way safer.

35:21

Alexey

**That's a long trip – from the plant to Iceland and then to your computer.**

35:27

Marijn

All the internet connections – like the ones I just showed – go through Iceland, yes. Through a VPN. Somebody asked, “Does my cat complain?” Of course the cat complains. The cat complains about everything. Mainly food, which is the one thing I haven't automated yet.

35:43

Alexey

**[laughs] Yeah, next project. Right?**

35:45

Marijn

Don't tell her.

35:47

Alexey

**Yeah. So what about the LED lights? You said you put some LED lights on your piano?**

35:54

Marijn

Oh, no. I'm still working on that. I put smart LED lights on my piano – now I'm working on the microphone that still isn't running – so that when I play, the lights move based on the tune and rhythm.

36:09

Alexey

**It's a great project. It's an amazing project. But somebody might be thinking, “Okay, this is fun. But this is not serious. How can I talk about this to a potential employer?”**

36:21

Marijn

Oh, easy example. I soldered some wires to my coffee machine so I can remote control it – so I don't have to walk over every time. But as it turns out, as I can now monitor my coffee use through the smart plug that measures the power usage and me pushing the button, I write that away to a database. Then suddenly, after two months, I had data to run a time series model on my coffee addiction.

36:47

Alexey

**[laughs] That's cool.**

36:48

Marijn

I have experience in time series modeling. I think the thing is currently offline. I'm so sorry. Clearly, my coffee consumption got messed up ever since we're going back to the office. So sad. Total trends breach. But suddenly, I have stuff that I can show. And it counts *more* than the next guy that says “Oh, I did a training on that.”

# Building a project portfolio

37:13

Alexey

**This is a question I get quite often. “How do you get inspiration for these projects?” I imagine it might be difficult. Let's say you are looking for a job because you're switching careers or you're just graduated or for whatever reason – you want to have a portfolio. Then the question is, “How do I build this portfolio?” Okay, you can go to Kaggle. You can pick a competition. You can take part in this competition. Then you can train a model, and then maybe if you're lucky, you can get a silver medal. Then you put it on your CV. But so do the rest of the people there, right?**

37:49

Marijn

Don’t get me wrong, I *highly* respect the Kaggle grandmasters who can optimize the heck out of stuff. I can't do that to that extent. But that brings me back to my original statement, “Why fit in when you're born to stand out?” Because it takes *so much effort* to be the top 0.5% who beats all the rest in the same Kaggle competition. It's much easier to create your own competition like I did with my coffee machine. How did I get the idea? Because I'm lazy. Because I didn't want to walk to my kitchen and back just to turn it on in between meetings, and then walk again one minute later. It is pure, unadulterated laziness. That is the main force behind automation in the modern world – and makes for really good developers. Laziness that got me to build that stuff – also sheer curiosity to see if it would work.

38:48

Alexey

**But what about other – let's say “more serious” – projects? You mentioned studying in Rotterdam and other things. I remember the story that you told me the other day about helping farmers in India.**

39:03

Marijn

Oh, yeah. Some NGO projects back in the day.

39:06

Alexey

**Can you tell us about that? How do you find projects that are not related to the plants in your flat, but other ones?**

39:16

Marijn

I think this is a broader data science thing in that you shouldn't just build what they ask, but instead you should build and do what organizations need. Many of you will join organizations and like “Yeah, we need something with data. We need something with AI.” But you're there to solve problems and help them make better decisions. There's a reason why Google calls it “chief decision scientist” these days. If you're just building what they ask – which is usually Excel++ stuff – you are not profitably applying all the ideas, all the knowledge, all the skills that you have.

This brings me back to the “be proactive” part and “suggests stuff.” This is the important part. This is why they call me a consultant, because I come in there because of my expertise. But most data scientists are consultants within their own organization because they're the guy or girl who knows that stuff that the rest doesn't know.

# Marijn’s NGO project

40:25

Alexey

**Yeah. Can you tell us about this helping farmers project?**

40:30

Marijn

Yeah. Yeah, sorry. To get back to that – it's an NGO. Two, three NGOs that we're working with in India and Kenya. They aid smallholder farmers in developing countries – help them achieve higher yields because between 30 to 50% (depending on your math) of the food in the world is produced by smallholder farmers, according to UN statistics. So if you want to help world hunger – and holy moly, we're going to get another world hunger crisis this summer – you need to help the smallholder farmers produce more. So we give them advice on how to achieve better yields, what their yield will be at the end of season, how they can apply crop rotation, planting what would actually give them the highest earning and stuff.

We use weather satellite imagery, all kinds of feeds for that. I'm hopefully going to India in two months to help out over there – to speak to locals, to interview them, to figure out what they use and what they don't use. Because here we have part of the ask, like “What is my yield going to be at the end of season? I would like a forecast.” But the optimization part they are not asking for, but we are going to provide that anyway. Because if we’re building a prediction model, we will see what the biggest predictors are for you. And we can advise you on that. We can talk to you about that. Some of those predictors, they can change – you live in *this* part of the West Bengal region, it's *that* fertile over there. So that's the yield you're going to get. Some other things, like what they plant, when they plant it, how they plant it – I am by no means a farming expert, I have sensors to keep my own plants alive – but this is stuff that they do not ask, but that is of *immense value* to them when you give it to them.

That, again, is data science and also my niche in the field, I guess. Everyone screams about “explainable AI,” but I think there's a very small threshold between explainable AI and just using statistics and models to explain what the hell is happening to help people make better decisions. And I do that for my job. A lot of Dutch and international organizations also want to help smallholder farmers in God-knows-where, because it's the same damn thing – just different data and like 15 different ways of spelling cow dung.

# The importance of a team

43:08

Alexey

**If you are not an expert in farming – or whatever other domain, whatever other field – and there are some things that domain experts, people who are your customers, don't ask for. How do you know that they are going to need this if you're not an expert in that yourself?**

43:30

Marijn

Alexey, this refers to our earlier topic. You don't need to know everything. I'm part of a team. I'm bragging about it, but I'm just one of five glorious bastards who work on this project. We all have different skills – that one knows much more about farming and that one can make PowerPoints that I never could. It's collaborating with people who know a lot of stuff that you don't know. That gives you an opportunity to learn.

If you're stuck in uni (like I was stuck in uni) surrounded by people who know the same stuff, you will feel very inadequate because they know more about your specific field. But here, I'm working with a guy who actually was a farmer years ago. A lot of tractor jokes, I swear. But he can advise me, and us, on stuff in ways I can't. So I don't need to stress out over knowing everything because we work in a team for that. I just stress out over knowing the things I should know. Believe me, that’s stress enough.

44:39

Alexey

**I was thinking you would say something about “You have to ask questions and (maybe not grill experts) but ask them a lot of questions. Try to really understand the domain.” But having a domain expert in the team is easier, I think. Right?**

44:55

Marijn

You need them. You need to make use of them. Of course, you should ask all the questions because people explain the problem, but they usually don't explain the problem *behind* the problem. You get what I mean?

45:06

Alexey

**So this guy with the tractor jokes and farming experience – this is why he got the job – because of his farming experience? Was it the way for him to get on the team?**

45:16

Marijn

He's got quite a lot of other skills as well – but it does make him stand out. And that is exactly what you want. You can stand out with your hair, or you can stand out with having experience and knowledge that nobody else does. Again, in my team, I have this guy who studied aerospace – God knows what. But he knows everything about Boeings and it's really interesting during lunch to hear everything about airplanes. But as soon as we're talking – I don't know logistics optimization – and it involves airplanes, *he's* the guy we go to. Because he knows airplanes.

# Open source intelligence (OSINT)

45:59

Alexey

**Another thing you and I talked about just a couple of days ago – you mentioned already that there might be a problem with famine and hunger in the next months again, right? I think it's related to war in Ukraine. This is something we've talked about recently and one of the things you mentioned is that, as a data scientist, you can do what you called “open source intelligence.” Can you talk about that a bit and also about hunger?**

46:30

Marijn

Sure. Let me start out about open source intelligence, also known as OSINT, is the field of grabbing social media data – stuff that's open, that's out there – and using it as valuable data to do stuff. I was already doing that 10 years ago when I was studying to grab Twitter data in order to analyze which parts of the Netherlands would and wouldn't be open to receiving refugees. And the war in Ukraine is the first modern information war on earth. It is incredibly interesting what we're seeing over there. History is being written. I'll skip all the doom scrolling because I just spent three months just doom scrolling. But just yesterday, did you see the deep fake of President Zelenskyy being posted?

47:24

Alexey

**You shared it. I try to spend too much time on…**

47:29

Marijn

Smart move. That's a very smart move. But this was the first time *in history* that people used deep fakes to impersonate a president to spread fake news. And it's not just Putin dancing and singing Abba. No, it's actually a serious misinformation campaign based on machine learning techniques. Now, back to OSINT – first few days, Ukraine was using GPS Intel based on Russian soldiers who took their phones along to figure out where the soldiers are and to send drones and bombs after them, which is terrifying and horrifying. Because this is war and I am doing – I don’t condone the violence and that kind of stuff. But simultaneously, it is so interesting to see two modern states, with modern technology, at war. We've never seen that before.

Syria, Iraq, Afghanistan, Palestine – it's all horrifying what's happening over there, don't get me wrong. I hope we get more sympathy for those people and their plight thanks to Ukraine. But the images we got from there was usually night vision goggles of people really zoomed out with turbans, so to say. Now we have high quality imagery hitting the internet 24/7 and we have a huge social media response to that. We also have people going on Tinder, like I was last week. I was Ivanka. I was using a GPS spoof to pretend to be on the Ukraine border and I was matching with people in Russian uniforms and we were all just reporting our location and the amount of miles – the distance to that person we matched with – because somebody else also matched with Ivan, but was over there. That way, you can triangulate that the Russian soldier, who you are Tindering with, on which part of the Russian/Ukraine border (because that's where he’s stationed) on which part they are. Thus we can, on scale, analyze where the reinforcements are on the border and the Ukraine Government actually uses this information.

It's absolutely crazy and I haven't even started about Anonymous using DDoS attacks to take down Russian news, for one, which was also really interesting. Because then Russia as a country started blocking foreign IP addresses, so you couldn't do DDoS attacks like that anymore. And yet, a few hours later, the DDoS attacks continued, because they were coming from within Russia itself. And this is all open information that we're using for very violent purposes. But it's never happened before. And it makes our field as data scientists relevant in ways that it has never before been relevant. I think that will only increase in the future.

50:46

Alexey

**If somebody wants to help with these efforts – with OSINT and other things – how can they do this?**

50:53

Marijn

Google is your best friend, because there's so many task forces – you have Anonymous, obviously, people on Reddit (unfortunately, especially Reddit) compiling lists. The whole Tinder thing is a Reddit thing, I'll admit this. Just gathering data. There's tons of open maps where we aggregate Twitter reports and sightings of “Oh, hey. Here, your bombs are falling. Oh, we saw track tanks driving over there.” I help out sometimes with that as well in the evening – of just pinpointing those tweets. There's no GPS, because the Ukraine government asked the population, “Please turn off your GPS,” because that's a way to target you. But these reports are coming in and we're mapping those – pinpointing them on maps – to see movement of troops.

Also all the war crimes being reported, that's going straight to Hague. They're gonna use this data that we collect on Tinder and Twitter these days – they're gonna use that in court in X months time. That is big. That's all data. And that's all data that's out there that we can use. It's free for the picking, really. To make matters worse, we basically knew this was coming for half a year, because for half a year people have been tracking the fact that there were tons of military trains everywhere in Russia going towards the border. That was mapped. We knew all the tanks – that data was also present.

52:28

Alexey

**Nobody wanted to believe it would happen.**

52:31

Marijn

I didn't want to believe it would happen either. I have friends on both sides, for the record. And fuck – I pray for their safety. But historically, these are… this has never – this is unique. Never before. That brings me to the last part and that's the famine. I'm also working on a few projects for that. Did you know that the Ukrainian flag – the blue and the yellow – is actually grain? The blue is the blue sky and the yellow that's this golden grain – because it's the breadbasket of the world. So we're gonna see huge decreases in grain production over the next year, which will lead to higher grain prices, but also higher meat prices worldwide, because a lot of that grain is actually used to feed animals.

53:23

Alexey

**Not only that. That's only one of the parts, right? Then there’s also energy and many other things.**

53:31

Marijn

Oh, yeah. *A lot* of things.

# The importance of soft skills for data experts

53:34

Alexey

**Yeah. Okay. I don't know if I should change the topic or not because we have some questions. One question from Mike (you probably don't see this question because it’s from Slido). Mike is asking “Should data experts learn soft skills, like people skills, communication, networking, drinking, as well as hard skills to go further and faster?”**

54:00

Marijn

It goes both ways. Whatever makes you unique – what makes you better (I don't mean you should paint your hair purple, by the way) but whatever makes you stand out and gives you an edge works. Even if it's being the one who can also talk to the people, the one who is not afraid to stand on stage – I did opera as a kid. That is why I can talk about it like this, I'm pretty sure. Because when you sing Italian for 500 people in pants that are way too tight, you can tell 500 people about data. But the point here is – it doesn't mean it all has to be social skills.

You can also excel because, I don't know, it's all an Azure and then this client suddenly wants AWS – you're the only one who also worked with AWS. That works too. Just don't be the same as everyone else, which you want to be. I know a guy who did eSports for a few years, and thanks to that, he excelled at his project because they were mining (I don’t know, League of Legends or some shit data) and he happened to have played it in the past. So he was like “the expert”.

55:18

Alexey

**[laughs] Usually, when you look at job postings that require a certain profile – this is often a description of an average data scientist with three to five years of experience – and this is who the company usually wants.**

55:34

Marijn

Those are means and averages and stuff – and *we*, as data people should know how horrible averages and means are. The *average* person has one boob and half a wiener. Think about that.

55:54

Alexey

**[laughs] Okay. Well, maybe to summarize, a way to stand out is just taking what you know and doubling down on that. Right? I think this is something you said at the beginning. Let me check if we have other questions. In live chat, do you see anything? There was a comment about…**

56:15

Marijn

Oh, yeah. The Tinder sample I got? Yeah, I'm sorry. Ivanka was a really bad name. I had to come up with something.

56:26

Alexey

**[laughs] Yeah, it's not a typical name, I think. At least not in Russia. In Ukraine, maybe.**

56:32

Marijn

No, I was using auto-translate and pretending to be an exchange student. I didn't really work that well. I think, even though my dad never gave me much advice in the world of data because he really doesn't get it, he did give me dating advice, which is “Don't pretend to be someone you're not.” And career wise, it's exactly the same. Don't pretend to be someone you're not. At least not in the long term. That doesn't help you and that doesn't help him or her, or your clients, or your boss.

57:08

Alexey

**But also your father said (correct me if I’m wrong) your father told you that social science is not really a good way to go, right?**

57:18

Marijn

“You don't get a job with that.” Turns out, you do. But that's just because his data – his experiences of the past – told him that.

# Marijn’s LinkedIn growth strategy and tips

57:30

Alexey

**Yeah. Right. Well, we still have time for one more question maybe. This is something I wanted to ask you – so one question about LinkedIn. I see that you're pretty active there and you have quite a few followers. What is your growth strategy there? Is it about posting cats and memes, or?**

57:50

Marijn

LinkedIn is an algorithm, same as any other. It's the nearest-neighbor-style matching algorithm, I believe. You can optimize that. Open secret – I think you should post on Thursdays and Tuesdays between eight and nine in morning, your time. There's an optimal time to post. There's an optimal amount of hashtags and people to tag in your posts. Type of material you post – you trend a lot less if it's an animated video. If it's without a picture, it's circulated even more. There is a ton of stuff you can do to optimize your posts to get a bigger reach. I may or may not have done the same for dating apps in the past. But now this scratches my itch, you know.

If you optimize (I could do a whole lecture on how to optimize LinkedIn) you will get a bigger reach and that is fun in and of itself. I am basically trying different things and for the first year or so it was terrible. I was just trying to reverse-engineer what works and what doesn't work. And for me, that is – try to post on weekdays between eight and nine in the morning. Because your posts will basically circulate for one hour and after that, if it doesn't reach a required amount of likes or comments, it will stop trending. Comments are much stronger than likes. If you want to support a post, comment. LinkedIn is a comment-driven algorithm, not a like-driven algorithm. If it reaches the threshold of attention, then it will circulate for like 12 hours. This is why I need to post early in the workdays, because then you will have the morning, everyone checks their phone while they're on the loo or having a coffee. Meaning, you optimize your odds of getting liked in that first crucial hour. Then you'll also be trending over lunch and dinner. If you post three at night, very few people we're going to like it, unless you have a lot of Australian friends.

Then there's the content thing, which is – there's three types of posts on LinkedIn: bragging, information, and humor. I remember a big marketing manager from a company I shall not mention going like, “Hey man! Posted any memes recently?” Even though I trend harder than I bet his entire marketing department. Most people and organizations just post bragging stuff, like “Oh yeah! I have the certification!” Or “I got this promotion!” Posting about your promotions is great. That's good. But people like that because they know you. They don't care about the information or the humor in there. But if you want to post every week, every month, or every day, you can't just brag about yourself, unless you're like Apple and you *only* post brags. But nobody likes that. Then only people from the same organization like your shit and nobody else does. But if you post information that's actually informative to people, *with humor,* you can post about that every day.

That's basically what I do. Okay, sometimes I brag. But I try not to post bragging posts, but just information about my field, about my experiences, about things I learned, about the things I did wrong – especially very funny, usually humor about me screwing up – like with the managers and burnouts, like I mentioned just now. That makes for great posting material and I have that every week. So if I inform people, people like that. Bragging they will only like if they know you or your organization. But if it's informative and relevant to your field, people outside your organization and outside of your field of connections will also like it.

I have more followers and people (you can analyze this). I have more people liking my stuff from Accenture and KPMG than from my own damn company. Are you listening to me Capgemini? Like my shit more! But that does mean it is a value to them. Don't try to brag about yourself or be funny about yourself. Give them something that informs – that they can use. If you just keep sending information, giving information, you will start getting information back – same as talking with the seniors.

62:24

Alexey

**I remember one of your posts. It was a swamp and there was a data engineer and data scientist. I think it was wombats or what were the animals?**

62:33

Marijn

Oh no, that was capybaras. I have to be very specific on that. Those were capybaras – with the tiny one as the data scientist and the engineer below it. I have to point out that at least half of the memes I post are my own. The other half might be from 4chan.

62:50

Alexey

**[laughs] So which one was that one?**

62:53

Marijn

The capybaras, those were mine. I'm a big fan.

62:57

Alexey

**Okay, I think we should be wrapping up. So the best way to find you – LinkedIn. Am I right?**

63:05

Marijn

Oh, yeah.

63:07

Alexey

**Okay, then I guess that's it for today. Thanks a lot for joining us today, for sharing, for telling us all these awesome stories, for telling us how to hack a data career. Thanks also to everyone for being active, for asking questions, for watching us. Everyone, enjoy your weekend.**

63:26

Marijn

Take care everyone. Let's stay in touch and let's stay safe.