0:00

Alexey

**This week, we'll talk about data science for social impact. We have a special guest today, Christine. Christine is a writer and researcher of tech and social issues. She is currently studying data science for public policy, and previously, she spent years managing social programs and exploring Data Science for Social Good. Welcome, Christine.**

1:30

Christine

Hi. So happy to be here. Thank you.

# Christine’s Background

1:33

Alexey

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

1:41

Christine

Sure, absolutely. For me, I started wanting to really figure out how to serve marginalized communities, so I studied political economy and sustainable development in my undergrad. Then for the next four years, I worked as a program manager for sustainability initiatives. And then after that, I sort of became fascinated with data, or I saw how powerful it was in those roles.

I made sort of a pivot into working as a data analytics person in the private sector for over four years. At that point, my goal was always to bring those things together and that's what I'm doing now in Berlin. I came here for a graduate program called Data Science for Public Policy.

2:30

Alexey

**What do program managers do?**

2:33

Christine

Yes, great question. In the social impact space, you're definitely managing different initiatives. For example, I was a program manager at a corporation doing their sustainability and corporate social responsibility. So I was the… I want to say “king” but that's not [laughs] – the “queen” of several initiatives that I took ownership of. So we had sustainability reporting, things like that, that I manage.

3:10

Alexey

**And by “managing” it means that you make sure that things get done? Or what does that mean?**

3:16

Christine

Yeah, it's a little bit more than project management. I think the main difference is that you not only become the face behind the cause – these are initiatives which don't have an end. There's no end. For example, sustainability reporting – the organization hopes to do that forever. So there are a lot of long-term stakeholder relationship developments and managing large data sets. But yes, ultimately, you need to make sure to drive the situation forward.

3:54

Alexey

**So a “program” is a general direction and then “projects” would be individual steps that you need to do in order to move in this direction?**

4:04

Christine

Yes, projects are something with a known end. Yeah.

# Private sector vs Public sector

4:07

Alexey

**Okay. You said you worked as a data analyst in the private sector. What is the “private sector”?**

4:17

Christine

“Private sector” is any organization which does not have social benefits. I think that's one way to say it. So it's just corporations – businesses. For example, I worked at Toyota, Wells Fargo, Blue Cross Blue Shield – which is the largest health insurer in the US.

4:41

Alexey

**So basically, just companies. Right?**

4:45

Christine

Yep, exactly. Private companies. Yeah.

4:48

Alexey

**So the opposite would be “public sector,” which are government institutions, right?**

4:53

Christine

Exactly. Yes.

4:54

Alexey

**Okay. So you were a program manager, you liked data and you decided to go more into data, so you became a data analyst, and then you moved to Berlin to study data science for public policy.**

5:09

Christine

Genau. [chuckles]

5:10

Alexey

**[chuckles] For those who don't speak German, you said “Exactly.” Right?**

5:15

Christine

Yes, exactly.

5:16

Alexey

**Okay. How's your German?**

5:18

Christine

It's coming along. I'm trying.

# Public policy

5:25

Alexey

**Okay. [chuckles] So you're studying data science for public policy and I was wondering – what actually is public policy? Can you tell us what that is?**

5:35

Christine

Sure, absolutely. Public policy is essentially – it's describing all of the laws which govern people and places. It's important in relation to social issues because it is the way that many social issues are addressed. Many social issues, obviously, have taken government intervention to address. For example, public libraries wouldn't have happened on their own – it requires funding, support, and space – things like that. So policies are the way that you get things done at that higher level.

6:21

Alexey

**So a policy would be something like “For every 1000 people, we need to have at least one library.”?**

6:30

Christine

Something like that, yes. I'm not actually aware of any library laws. [laughs] But that would be a good example. I can give other real examples. But yes – ensuring public welfare.

6:44

Alexey

**What are the examples that you have in mind?**

6:46

Christine

Sure, yes. Something that I worked on, specifically in the States, was environmental legislation regarding electronics recycling. This was something that companies did not want to do, companies like HP – people that basically create printers, or monitors, or things like that – essentially, there was no easy way to recycle these at the end of their life. People would throw them away and nine times out of ten, they would end up on a ship for a developing country where children would break them down and be exposed to toxic chemicals. There was just a lot that was problematic with that situation and the companies were not willing to create a new way for people to easily rid themselves of these products in a sustainable way.

I worked as a grassroots organizer with a nonprofit to canvass neighborhoods and gather community support to petition lawmakers to eventually pass a law which required companies to take back those products. They put stations in every electronics store where people could bring these products and there would actually be checks in place to make sure these products were recycled in a sustainable way.

8:08

Alexey

**The companies did not have enough motivation to do this on their own, so they needed an extra nudge coming from the government saying, “Hey, you actually must give people a way to recycle products.”**

8:20

Christine

Exactly.

8:21

Alexey

**The government came up with a law and you were kind of the driving force behind it – you did this as a program manager, right? Or were you a data analyst?**

# The challenges of being a community organizer

8:27

Christine

[cross-talk] This was actually before that. I did this as a community organizer. There's lots of stuff in between and exploring social issues. But yes, I was a community organizer. I knocked on doors, actually, gathered signatures and donations, and then we brought everything to the Capitol. Yep.

8:51

Alexey

**You must be quite persuasive to accomplish this, right? Or is it easy to persuade people that “This thing is bad. This thing is good. Give me your signature.”?**

8:59

Christine

Yeah. You know, actually, I've worked for over 10 years and seriously – that is still the hardest job I have ever done. [chuckles] Yeah, it's not easy. You knock on their door, they don't know you, they’ve never seen you, they’ve never heard of this issue – and now you want money. [laughs] And you want them to write their signature or print out some letters. People are doing this – community organizers are doing this – everywhere. It's just the hardest work I've ever done. [sighs] [chuckles] Yeah.

9:31

Alexey

**But I guess they don't have to give you any money, right? All they have to give you is a signature.**

9:37

Christine

Both are definitely equally as important. The letter is what gets this issue noticed by our representatives, but the money definitely helps facilitate all of that activity – like going to the Capitol, all this time.

9:56

Alexey

**So you were a volunteer but somehow you still needed to get paid. Right?**

10:01

Christine

Oh, no – I worked there. It was my job to do that. But yes, the organization as a whole needed, obviously, funds to operate. So we would knock on doors for eight hours.

# How public policy relates to political science

10:14

Alexey

**Regarding public policy – I thought policy is maybe related to politics, or political science. But it's not, right? Or is it?**

10:25

Christine

It is, yes. I would say political science is just the study of the systems of governance and public policy is essentially putting it into action, I guess. I think one is an application, I guess, more so. Maybe that's a good way to say it.

10:51

Alexey

**I think you've partly answered that previously. I basically wanted to know how public policies are related to social impact and I think the example you gave is pretty good. I was able to understand it – companies need to recycle the products they do not want, so how can we encourage them or force them to do this?**

**The social impact is that the environment perhaps doesn't get harmed as a result – you don't have these chemicals that you mentioned and children do not play with printers. Is this how it's related? You come up with policies to have a social impact. Correct?**

11:41

Christine

Absolutely, yes. I think social issues tend to just be really, really complex. Usually, there's just multiple ways to address it. I'd say for example, in that situation, the bigger picture is like, “How do we increase sustainability? How do we increase well-being?” And this is just one little part of that. So yeah, exactly what you said. Yeah.

# Programs that teach data science for public policy

12:10

Alexey

**Coming back to “data science for public policy,” to prepare for this interview – I googled it. I found the program where you study and I found a couple of other programs. I think there was one in the States – in Washington University, I think, or it was from Washington, DC – and I saw a couple of other places. Then I also came across a book that has exactly this name “Data Science for Public Policy.” But, in general, is it mature as a discipline? Are there other textbooks apart from this book? Are there people who are seriously working on this? How can you (I want to save “marry”), but like how you can *do* data science for public policy?**

12:56

Christine

Sure. My sense is that it's very new. I started following this concept of ‘Data Science for Social Good’ maybe five or six years ago. At that time, (obviously, I was in the States) there was really only one program that I was aware of in the country – it was at the University of Chicago. They had a fellowship, actually – it wasn't an academic program – where they would help technical or policy people to bring these concepts together.

Actually, the reason I ended up in Germany is because, upon doing research – I had done research for years looking to answer “How can I combine what I want to work on?” And this was the first data science for public policy program I found. Now, yes, like you said, that there are so many that it feels like they're everywhere now. [chuckles]

13:47

Alexey

**Well, not so many. If you just look for university programs in data science, you will find millions of them. But specifically for public policy – maybe for social good and so on, there are more – but for this specific thing, public policy, there are not so many yet.**

14:06

Christine

Right, sure. Yeah. [cross-talk]

14:07

Alexey

**At least I wasn’t able to find many. [chuckles]**

14:11

Christine

No, you're exactly right. It's probably still like, less than 10. But that feels like a *huge* amount [laughs] when for years, there was nothing.

# Data science for public policy vs regular data science

14:22

Alexey

**I do data science at work and I'm usually dealing with things like… I work at an internet company, so we look at clicks and transactions and things like this. So we, as data scientists, deal with things like linear regression, logistic regression, classification regression, and so on. So I was really curious, what are the use cases for public policy?**

**I checked out this book that I was talking about, Data Science for Public Policy, and in the table of contents I saw that there are chapters on transforming data, record linkage, exploratory data analysis, cluster analysis, NLP – to me it looks like “Okay, this is what I already do.” Does this mean I already know everything I need for public policy or are there some nuances?**

15:19

Christine

Yeah, there are definitely nuances. Yes, I investigated this book as well. I know it's fairly new. I don't know that I would consider it “the Bible” or the definer of the field at all. Just like you said, I do see that the majority of the book is just data science, so I think that that's something that's really targeted towards policymakers or political scientists that want to explore this field. But I actually think the goal of data science for public policy is more so data science for social impact.

In the field, I guess why you specify “public policy” is that I think there are not only a lot of tech issues – ethical and challenging issues with technology – that we are facing right now. We need people who are actually educated in what data science is, I think that's really critical. But also I think we're getting to a point where every type of organization, including social impact organizations, will need a data scientist to support them or help them make progress.

I think that's the future as well – having data scientists that have some knowledge of the landscape or the processes is very beneficial. From my experience being in both spaces, they are completely different. Working for a company, working in social impact, government, academia, everything – it's so different.

16:56

Alexey

**So it's aimed at people who are similar to how you were a couple of years ago before getting into the data world? It's aimed at educating them on what data science is, how it can be used, and how it can bring social impact, right?**

17:12

Christine

Actually, our program has quite a few folks that have a background in technology and there are lots of folks that are social scientists and myself – I'm like a mix of both. So I do think there's space for everyone that wants to have a very specific impact. I appreciate the specificity because social impact has so many layers. The school I go to, Hertie, is actually a school of governance, so all they teach is policy. They're experts at that part. I tend to think that that method is one of the most important methods to accelerating change. I do think people can come from different ends into this space and they just need a little bit of knowledge from both sides.

I can give a few examples of how this has worked. So for example, I am an organizer with Data Science for Social Good in Berlin, which is a great organization if folks want to learn more about this. We had a presenter come to our most recent meetup and they are supporting refugees that are making their way to the southern coast of Europe by flying drones, and then using that footage to identify boats in the water, which they can send aid to. This is actually a really challenging issue. They use computer vision to identify the boats. They shared some photos like, “Can you find the boats?” And I'm like, “No. No *human* could find these books.” [chuckles] So that's just a really critical way that this organization can support people that would otherwise receive no support.

I know in Scandinavia, they're using satellite imagery to identify all the flat roofs so that they can assess opportunities for rooftop gardens or things like that, to increase sustainability and livability in cities. I think if you have a specific social cause that really speaks to you or resonates with you, there is probably a way that you can apply technology to scale up the support for this issue. There's lots of examples but yeah – it can be policy-related.

19:51

Alexey

**So usually there is some social problem and then organizations like Data Science for Social Good and others have a list of problems that they think should be addressed, right? Then they get people who know data science, who know public policy, who know all these things you mentioned – they try to put them together into one place to solve the problem, right? This university that you mentioned, it kind of teaches how to oversee this process, or how to run this process, right?**

20:28

Christine

At this point, they are more focused on teaching the subjects separately, but there's so much to learn about the policy process – economics, policy analysis. Having just finished my first year, I feel like there's so much more to know. [chuckles] But yes, I think Data Science for Social Good Berlin does have nonprofit partners with data issues that volunteers then help them support, specifically.

# The importance of ethical data science in public policy

21:06

Alexey

**Coming back to the book – one of the chapters was not the usual regression or usual entity linking, but rather one of the chapters was about ethics. How important is the problem of ethical data science to public policy?**

21:27

Christine

Sure, I think it's absolutely critical. It's generally critical for all data science [laughs] you always hope you're doing ethical data science. But I think the way that it relates differently for data science for social impact or public policy is that, ultimately, we hope that laws reflect our values and they reflect the society that we want to have.

There are many things that are not illegal, but are unethical. So there's a gap where we, as policymakers, ideally step in and try to close. Yes, I'd say it's absolutely critical in helping these social impact causes and organizations to build a strong foundation in data science. But I also just think that we will have some challenges in the near future regarding new technologies and innovation. Folks need to understand ethics to deal with them, for sure.

22:35

Alexey

**In your printer example – throwing away a printer would not be a crime, right? But still, it's not ethical because it can contaminate the environment. Bad things can happen, but nobody will put you in jail for that.**

22:52

Christine

Right. Yes. If you know that a four year-old is going to be exposed to toxic chemicals when you throw it away, then you throw it away in a way that doesn’t result in that.

23:06

Alexey

**Do you maybe have some examples where data science was applied to do these things, but not ethically?**

23:20

Christine

That's a good question. Let me think on that. Maybe I will remember something later.

23:26

Alexey

**I guess people who work in these areas maybe think about these issues of ethics more than, let's say, the average data scientist. Maybe these topics of abusing something just to get the model right – these things do not come up.**

23:49

Christine

Oh, sure. Yeah. Actually, now that I think about it, I can think of a few examples. One thing that I think we'll probably come to later is new legislation in the EU about AI – the EU AI Act. One of the things that it deals with – its core method is essentially to assign risk levels to different technologies. The highest risk level explicitly references this social scoring system in China. This system essentially, through all of your goings and operations in society, each thing will give you a score, which will give you more access in society.

For example, if you get a car loan and you pay it on time or you pay it early, you might be able to get a visa faster. I think the consequences of that are actually big. Imagine you have one situation and your score is damaged and then your whole family is just… you're second class citizens now. I think that's kind of an abuse of big data and all of these technologies and things like that.

25:15

Alexey

**You've given an example from China, but I was thinking, “Wait a minute, isn't there a similar thing in Germany called Trufa?” You can easily damage this and then nobody will give you credit. Maybe it's not as abroad as the system you just described.**

**I can understand that the intention is maybe to make people behave well, right? Pay their debts, not litter, or whatever – but then the consequences could be pretty wide.**

25:59

Christine

Right. And I think the opportunity for abuse is obviously huge. Maybe you offend one of the maintainers of this database, then they just give you a zero score and then you can't access anything in the social sphere.

# How data science in social impact project differs from other projects

26:21

Alexey

**What is unique when it comes to applications of data science to social impact projects?**

26:29

Christine

I think the main difference that I see from being in the social impact space, and also working with organizations like Data Science for Social Good and similar organizations in the States, is that I think a technical person has to sort of expand their thinking. The social issue ultimately will not be solved by this one data science project. Often the technical solution has to consider the longevity or the bigger picture of this social issue to really make it effective.

Also talking to a lot of non technical stakeholders – people that, honestly, want nothing to do with tech and will not understand what you're doing – you need to really understand not only how these organizations work, you need to interpret all these different things about the specific social issue. What I observe is a lot of technologists saying, “Oh. Well, here's a project idea, but will that solve the whole thing?” And it's like, “*Nothing* on Earth will solve this one problem?” [chuckles]This is just one small part – we have one small goal and we need to make sure that it lends itself to further iteration.

28:06

Alexey

**Who are the usual people/stakeholders that data scientists working in this domain talk to? Are they policymakers? Who are they?**

28:20

Christine

I think nine times out of ten, they’re people very close to the social issue. For example, Data Science for Social Good (I'll just say DSSG) had a hackathon in December of last year – that's something that they do. They had a bunch of NGO stakeholders give datasets that have a very specific problem. For example, the German Red Cross was one of the stakeholders and they were addressing a hiring issue within their organization. Their stakeholder was a hiring manager (an HR manager). So that person is very close to all of the challenges, and you need to sort of “mine” them for information to create a dynamic solution.

29:17

Alexey

**Mine for information? [chuckles] How do you do that? I guess you ask them questions, fill out some questionnaires, right? Is this how you do this? Or do you just sit and talk with them?**

29:36

Christine

I think it takes a lot of proactive effort on your end and then, of course, you ask them questions. In this situation, for example, it would have been great if the data analysts did their exploratory data analysis and did research on the structure and history of the Red Cross so that they could be prepared for what the stakeholder will share.

A stakeholder generally has very specific challenges that they face, but they don't have a sense of how it can be solved with technology. So they really can't help you with the technical aspect. You really have to go to them with “What does the technical system look like in their organization? Who manages it?” And things like that.

30:32

Alexey

**And how technically advanced is this part? Is it just an Excel spreadsheet? Or is there an actual database? Or is there a website? Or is it something like old-school books where you take notes? Or all of the above?**

30:46

Christine

Yeah, all of the above. I feel like I've seen everything. I guess that's the one of the big challenges that these organizations face – they're donor funded. They live on donations. A lot of times, you need to understand that some staff members are funded for a temporary amount of time, and that later, that person might not be there. So maybe you shouldn't build your technical plan around that one stakeholder. But yeah, I've seen it all. Unfortunately, I personally have never seen a technically savvy organization – it's always a mess. [laughs] That's my experience. Really bad

31:37

Alexey

**If you're not an IT company – you have an organization but you realize that you need some sort of IT system quite late in the life of the organization, right? You just try to put it somewhere, but it doesn't really fit.**

31:56

Christine

Exactly. Yeah.

31:57

Alexey

**That's the one of the challenges, I guess, that data scientists need to solve, “Okay, now you have a model that detects boats in the ocean for helping refugees. But so what? How do you use this now? How do you actually help these people who you can detect from the drones?” Right?**

32:21

Christine

Right, absolutely. I think that digital and data literacy is really a challenge everywhere. Every type of organization that is not a tech company struggles with some level of digital or data literacy. But I think the challenge of a lot of social impact organizations is that they don't have the resources to do this type of investment into increasing that digital or data literacy. Due to funding, they're so focused on the cause, even if that hinders their efficiency or personal progress.

# Other resources to learn about data science for public policy

32:56

Alexey

**There is a question, “Which book are you referring to?” We were talking about a book called Data Science for Public Policy. I think it is just one book of that name from Springer. I don't think there are multiple books with this title. I think it's just the one.**

**By the way, I think one of the comments you made when we were discussing a question was that there is another book – another source – where people can learn these things. Which one is that?**

33:24

Christine

Sure, sure. Yes. I have a bunch of resources. I think if people want to explore this space – that will be great. But essentially, the source that I recommend for people that want to explore specific applications is actually the Data Science for Social Good Fellowship, originally hosted by the University of Chicago. They have an entire page of specific projects. So it’s DSSGfellowship.org/projects. I will put the link in the comments.

33:52

Alexey

**I just shared the link right now.**

33:54

Christine

Perfect, thank you. Yes.

33:58

Alexey

**So it’s a list of projects and what we can do with this list of projects is go there, check out what these projects are about, and see what kind of data was used there – what kind of outcome was achieved. Right?**

34:12

Christine

Exactly. Yes, most of them have videos where people involved in the project talk about the challenges they face, what they did, and there's all types of tech used – computer vision, predicting text analysis, etc.. There's a great diversity.

# Challenges with getting data in data science for public policy

34:32

Alexey

**One of the things that I asked you about a couple of questions ago was what's unique about using data science for social impact. We talked a bit about this and I was wondering – how difficult is it to get data? For example, in this case where there are drones flying over the sea, I guess it's not that straightforward to get data and then actually label it. You said that for you, as a human, it was difficult to understand which thing is the boat and where it is, right?**

35:03

Christine

Right. Yes. This organization – I don't want to say their name wrong, so I'll share it in the chat – but they were creating their own drones, actually, because they had to be sort of special or fit their budget or things like that. They had to try multiple different things – they had to iterate their data pipeline and process to sort that out. I think the challenge in a lot of social impact spaces is that it's a challenge not only to gather data, but it's also really challenging to find clean data, especially coming from the outside. There are just a lot of gaps, I guess. So it is a challenge to get data, even if you're the one gathering it.

36:02

Alexey

**I guess in a typical IT company, you join a company as a data scientist and there is a database – you just do select and you join a couple of tables. Then maybe one of a few features is missing, maybe the data is not the cleanest data on Earth, but overall, you can do something. But here, in this case, you come and there is nothing. There is a problem that needs to be solved. There is no data. There is no IT infrastructure – just a bunch of people who want to solve this problem. Is that right?**

36:36

Christine

Right, exactly. Imagine a data scientist building their own drone. [laughs] That's amazing. You have to build your own drone to gather data. [chuckles] Yeah.

# The problems with accessing public datasets about recycling

36:49

Alexey

**Yeah. So a question from Matt, “Are there models – guidelines, websites, templates – for recycling projects that would facilitate their implementation by a small group in the regional community?”**

37:05

Christine

Hmm… Recycling in a regional community – am I hearing that right? Man, I think there are a few challenges there that have nothing to do with the data science. And I think there's probably a couple of ways that you can approach that. So I need a little bit more information – I'm sorry to give such an unsatisfying response. But Matt, if you do want to reach out to me, please do, and I’ll be happy to chat.

37:41

Alexey

**But in general, is there any public information on recycling projects – like the one you mentioned about printers – for people to read and get inspiration about them? Is there public data about this?**

37:53

Christine

I don't know that I'm aware of a public dataset about any type of recycling program. I don't know that they keep track of how much they gather, how many people are involved – these numbers might be buried somewhere in a public budget, but they will not… you only receive like one number, “We deployed 5000 cans,” and you don't know to where necessarily.

38:29

Alexey

**I guess the companies, the industry, who produce this electronic equipment are not super incentivized to make all this data publicly available. Because, first, it’s bad for their image and they will have to do something about this.**

38:54

Christine

Right. Yes, absolutely. Any corporate data would not be public.

# Christine’s potential projects after Master’s degree

39:03

Alexey

**Okay. So we have a comment – “This podcast is great.” Thank you. But the question for you is, “Do you have an idea on what project you would work on when you finish your Master’s?”**

39:14

Christine

Oh, man, great question. Is that my professor? Is that my dad? [laughs] [cross-talk] Perfect. Great question.

39:26

Alexey

**There is no name. It could be either.**

39:27

Christine

Could be my dad. [laughs] Yeah, this is a great question. I think this is one of my main goals in being in this process. I'm currently in a fellowship working on a few different areas to see where I think I can be the most effective. The issues that, not only do I care about, but I think are the most critical issues of our time, are: climate justice, women's issues (like gender equality) or just women's issues as a whole, and then ethical or responsible technology. I think there's a lot of ways to address all those things. Obviously, they're huge issues. But it would be my hope to find some data science applications to one of those areas, for sure.

# Gender inequality in STEM fields

40:19

Alexey

**Speaking of gender inequality, just yesterday, we had another conversation where one of the points was that, just in general, females are not encouraged to go into STEM fields – to go into science. Some don't like these areas, but in general, society doesn't encourage them to follow this direction. Is there something we can do as data scientists, let's say, who work for public policy? Is there any policy that can help with this issue?**

40:53

Christine

Sure, yes. I think there's a lot. I actually saw you yesterday with Olga – she was really grilling you too. I was like, “What's happening? Tables are turned.” [laughs]

41:03

Alexey

**It was fun.**

41:04

Christine

Yeah. [chuckles] Good. That’s good. What I have observed and my understanding, is that a lot of times women are not only not encouraged, but they are actively prevented from joining this field. At one of my jobs, I have a manager who is a woman and she wanted to go into STEM, but was actually not allowed to.

I think on many levels this needs to change, but I do think there are ways that we can encourage more transparency. Because I do think a lot of this discrimination or subtle discouragement is really nuanced. It's hard to say, “Make that illegal. Don't discourage.” You can't throw someone in prison for that. [chuckles]

41:58

Alexey

**That's what I thought as well. It probably happens from the familial level, right? Maybe the dad says, “Hey, don't do this. Go study law,” or something like that.”**

42:09

Christine

Right. “Marry a lawyer. Don't study law – marry a lawyer.” [chuckles] Absolutely. I think these social issues are so ingrained. I mean, we really have to all, individually, do the work to study the lessons our society has taught us. But I do think increasing transparency in this space is one way to sort of change the conversation.

For example, in the EU we have this corporate social responsibility legislation, and one of the things that's not required but is a value in that space, is salary transparency. I think by requiring organizations to record this information – report the makeup of their companies – this is one way for people to see really clearly what's happening and keep companies accountable. I think there could be more education about creating safer organizations.

I know y'all talked about writing job descriptions that are sort of gender-neutral and there's some material out there for that. But yeah, these are some ways I think policy could come in or community efforts could come in, and try to address that issue.

43:30

Alexey

**Is it something that you can potentially work on as a part of your graduation project? Or as a part of what you do after the Master’s?**

# Corporate responsibility and why organizations need social impact data scientists

43:38

Christine

The rest of my life after Master’s? [laughs] Yeah, we will have a thesis, sure – like any other graduate Master's program. But I definitely hope to. I think that corporate social responsibility really will be critical to addressing all those issues. Corporations have such a huge impact now – they have such a huge reach – we really need to require them to be responsible citizens of our world. Legislation will help with that, for sure.

44:13

Alexey

**I imagine that in a project like that, the modeling part will not be the most difficult one. Maybe it will be just fitting a linear regression, but the most difficult part would be actually getting the data, analyzing the data, preparing the data, and then conveying the results in a way where the problem is clear, right?**

44:32

Christine

Sure, yes. Because of this corporate social responsibility (CSR) initiative in the EU – the amount of data in that space is actually growing really rapidly, so I think there will be potential for more sophisticated applications in the future. But you're right. I think right now, it's somewhat more of a policy issue for sure – a social issue. Yeah.

44:57

Alexey

**What kinds of organizations need data scientists that specialize in things like social impact and public policy? Is it governments? Is it universities? Is it somebody else?**

45:09

Christine

Well, I will be the annoying person that says “every organization” [laugh] in the future. And I say this because not only does every nonprofit hopefully have a data person, I honestly think that the whole data space is kind of evolving towards more of a data persona person. So I think, obviously, all of those organizations will hopefully have support in analyzing what they have so that they can more efficiently make decisions.

But every corporation has a public policy department as well. Those people also need to understand the technology and the way that their business relates to technology. All the time, I see things like “public policy at Amazon” “public policy at Airbnb”. And these are tech companies, they need to understand what’s happening – the gaps in the AI Act and things like that. So I think in the future, that type of person will be very valuable.

46:25

Alexey

**But right now, I know that in the States, there is a position called US Chief Data Scientist. What's the name of that person? DJ Patil. I think DJ was the first Chief Data Scientist. I'm wondering, what kind of problems does this role actually solve? All these problems that we talked about here in this podcast?**

46:52

Christine

Sure. You know, I'm actually not too familiar with his tenure in the Obama administration. I did follow him quite a bit afterwards. He's very dynamic and is a really huge advocate of the humanities, education, things like that – diverse education for backgrounds of data scientists.

I’ll need to investigate, because I think at that point, the US does lack a lot of complete open data as well. So I think it's like data science but maybe just data at first – I don't know.

47:29

Alexey

**I guess the idea is like, “We already have so many data sets, now let's have somebody who can make use of this data.”**

47:37

Christine

Right. I think it was a savvy choice for Obama to at least send the message that “We want to be a country that values innovation and technology, and is proactive about accommodating this,” because I think at that time data science was like *the* buzzword. It was very foggy as to what data science actually was. So I think at that point, it was more like… I don't know what.

48:09

Alexey

**Marketing?**

48:10

Christine

Yeah, yeah. [laughs] I don't know. Politicking? I don't know. I will investigate, though. This is a great question.

# What you need to start making a social impact with data science

48:19

Alexey

**Let's say somebody is interested in learning about this and working with data science, making a social impact, and making a difference with their data skills. What do they need to get started?**

48:37

Christine

Well, I would just congratulate that person and really just encourage them to keep pursuing that. But I think I would recommend two things. Not only is Data Science for Social Good an excellent organization, but this is a global organization as well. So if you're not able to join – if you don't have a local Data Science for Social Good group, there are ones where you can volunteer remotely. If you feel ready to dive right in, this is a great way for you to learn from folks that have been doing this, get exposure to social issues and things like that.

I think the second thing is I would really encourage you to see if you have a particular issue that resonates with you, and really start your research and understanding the current landscape of that issue. For example, I think we have the Sustainable Development Goals, created by the UN and that lists out very clear, specific social issues and goals. I think if you read through those, find something that resonates with you, and do a little research, for a lot of those (because they were presented a while ago) there's beginning to be more data for each of those specific social issues. So you could get far on your own if you do have a cause in mind.

50:04

Alexey

**Is there a list somewhere? Because I know that there is Data Science for Social Good Berlin. I guess there is Data Science for Social Good Germany, then there is one in Portugal, there is one in Poland. We have quite a few organizations. Then there is Omdena, I think they're also doing things like that. Do you know of Omdena?**

50:24

Christine

I actually don't. There are others, but I have not heard of this one.

50:27

Alexey

**I think they are also doing some projects and then they get data scientists who want to get experience so they can coach them and let them work on their data skills. So there are quite a few of them. It's quite… I want to say decentralized, for lack of a better word.**

**What I'm trying to ask is – is there maybe a place where all these problems are listed and then instead of your geographical area – where you go to Data Science for Social Good Berlin, for example – instead of that, you can see a list and think, “Okay, this problem really resonates with me, I want to work on that one.”?**

51:06

Christine

Would you say a list of local issues, or…?

# 80,000 hours

51:10

Alexey

**Yeah, for example, or a list of issues that might need some help – sort of like a job posting or job board with these kinds of problems where help is needed.**

51:23

Christine

You know, there actually is. This was another resource I wanted to share with the folks watching. There's an excellent organization – I think they're called 80,000 hours –essentially, their mission is to help people effectively create impact. They do have a job board, and they do rank jobs based on how critical that issue is to saving the world, basically. [chuckles] Like, “What's the greatest risk to humanity?” They will rank all the issues and organizations affiliated with those issues, in that order. So I will also put a link to that organization in the comments.

52:09

Alexey

**Yeah, I'm looking this up right now. It's actually a London-based non-profit organization and they have their own Wikipedia page, which says that they’re previous serious.**

52:18

Christine

They’re legit. Yes.

52:22

Alexey

**Yeah, and the link is 80000hours.org – do you know why they chose this name?**

52:26

Christine

80,000 hours? Yes. I think this is how long the average person spends in their career. That's quite a long time if you think about it, and you can make a meaningful difference with that time, instead of just… yeah.

52:42

Alexey

**So this is a way of encouraging people to spend all these hours on making an impact instead of bringing revenue to yet another… corporation.**

52:57

Christine

Soulless… [laugh] No, I'm kidding. [chuckles] Yeah.

53:00

Alexey

**[chuckle] I had this word in mind.**

53:03

Christine

[laughs] Sure, yeah. We won't say that, or we won't go there. But, actually, they’re quite a savvy organization – they're quite educational. I think there are many ways for people to make their career have a huge impact. An example I know they give is like this Google engineer who loves his job at Google – and he makes a pledge to donate like 20% of his income for the rest of his life, for the rest of his 80,000 hours. And there's other things – they provide many different options and they give quite a lot of education as to why this is impactful, or what's meaningful, and how to make a choice. Yeah, they're quite dynamic.

# Other use cases for public policy data science

53:56

Alexey

**I have quite an interest in learning a bit about other use cases. Maybe from your experience, or maybe now you're learning this with other students – with your classmates. What kind of problems do you see that they’re solving, or you see in general that are being solved now with data science?**

54:16

Christine

Yeah, so I think in my program right now, we're diving more into the deep end in both of these topics, and have not yet had much of an opportunity to bring them together. I know we had some group projects, for example, that addressed a lack of census data. I worked with a group, where we used satellite imagery to try to predict poverty levels.

It’s an incomplete project, but essentially, there are some countries that don't have this census – Afghanistan is one of them – so it would be really helpful to have estimations of what poverty looks like in these countries so that aid or support can be accurately given. Of course, we use computer vision for analyzing the satellite images. It’s an incomplete project but I guess it’s one way that you can sort of combine those issues.

55:34

Alexey

**Interesting. I guess as a part of a census – I don't remember the last time somebody actually asked me this question. I think it was like 15 years ago. I don't remember what was on there. They asked me my nationality, what I do, and I think that was pretty much it. So I don't know. Even though I was a part of the census, I don't know how much useful information they actually learned about me, apart from my nationality.**

56:06

Christine

Hmm. Interesting. Okay. Yeah.

56:11

Alexey

**Anything else that comes to mind? Maybe some other problem?**

# Coffee, Ethics & AI

56:22

Christine

Um. Nothing comes to mind right now, but I do actually want to pitch one thing if you want to explore the space. I moderate an open coffee club that focuses on ethics in AI. The goal is to democratize this conversation. Every meeting, every other week, we have different ethical challenges in the space. It's just a one hour casual coffee chat – start your day with an existential crisis [laughs] (I’m kidding) a challenge for the human race.

57:03

Alexey

**So what was the last challenge at your last meeting?**

57:09

Christine

Yeah, we actually met yesterday morning and we talked about LaMDA, the Google chat bot, which people are debating whether it's sentient or not and maybe the potential consequences of the response, or just the situation. It was an interesting conversation.

57:32

Alexey

**I’ve never heard about LaMDA. I heard about a bot from Microsoft that became racist pretty quickly. Like because it was trained on data from Twitter, so people on Twitter would teach the bot and then, of course, this would happen – if you just let people do this without controlling it. I should look LaMda up. Okay. If somebody wants to reach out to you, what's the best way to do this?**

# Finding Christine online

58:06

Christine

Sure. I'm on Twitter and LinkedIn. I do have a website with an email newsletter that you can try to contact or sign up. But yes, I think these are the main ways – Twitter, LinkedIn, and ChristineCepelak.com

58:23

Alexey

**So that's first name, last name.com and then you have a newsletter there. Okay, we will make sure to include all these links in the description under the video. I guess that's it for today. Thanks a lot for joining us today. Thanks a lot for sharing your experience, the stories with us – all these problems.**

58:44

Christine

[laughs] No one have a crisis. [laugh] No, I’m kidding. [speaks different language] Thank you so much.