1:10

Alexey

**This week, we'll talk about improving our world with analytics. We have a special guest today, Parvathy. Parvathy is a CTO at Analytics for a Better World, a nonprofit organization aiming to use analytic techniques to contribute to sustainable development goals. She holds a professional doctorate degree in data science from Eindhoven University of Technology. Welcome.**

1:35

Parvathy

Thank you so much, Alexey.

1:37

Alexey

**The questions for today's interview were prepared by Johanna Bayer. Thanks, Johanna, for your help. Also I would like to thank Antonis Stellas for introducing me to Parvathy. So thanks a lot for that. Let's start.**

# Parvathy’s background

1:54

Alexey

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

2:02

Parvathy

Yeah. I started in 2007, in India. I studied for my Bachelor's in electrical and electronics engineering. After completing the Bachelor program, I really wanted to work in the renewable energy sector because with my background in electrical and electronics, it was the obvious choice for me to go into the renewable sector. I was also very passionate about sustainable development. At that point, even though I didn't know the depth and width of this field, being a very young graduate, I did a Master's in energy management and climate change technology.

I realized that collecting and using data is becoming more and more prominent in this field. I started learning a lot of modeling simulation, data analysis, specifically in the energy management sector. I started working in India itself. I worked in research and development in a few research institutes in India, predominantly working with solar photovoltaics and in the data side of that. I saw a really good job opportunity in Singapore at that point, at the National University of Singapore, where they were collecting a lot of data from cross-climatic photovoltaic systems and analyzing it. So I jumped at that opportunity to learn more about data science and I applied and got in. I worked there for more than two years.

Then I realized data is not just usable in the renewable energy sector – it can be used for other fields in sustainability as well. And I realized that I do not have an education in data science, so I decided to pursue an education in that field as well. I moved to the Netherlands and pursued a professional doctorate program in TU Eindhoven. I'm currently based in Eindhoven in the Netherlands. After graduation, I started working in the public sector, so I started freelancing for World Bank and United Nations, specifically UNDP in India, on a series of sustainability projects, leading the data part of it across healthcare, transportation, infrastructure, and others.

Last year, the University of Amsterdam, together with some private sector players, decided to set up this Analytics for a Better World Institute, and they were looking for a CTO. Thanks to my good luck, I got the offer and last year, I started working with them. I still continue to consult for both World Bank and UN on a series of projects, but the majority of my time, I work as a CTO there. That's basically my career journey.

4:38

Alexey

**Hmm, interesting. What do you do there as a CTO?**

4:42

Parvathy

Basically, what Analytics for a Better World does is connect nonprofits with research and private sector expertise – not just organizations, but with expertise. What we see is that many nonprofits have questions that sometimes require deep research, but in many cases, they require things to be set up – scaling, development tools, dashboards, visualization, etc. but they internally may not have the capability. So what we do at Analytics for a Better World is try to make this connection and start projects for nonprofits, either research projects or implementation and analytic projects. I lead some of them – I help in brainstorming ideas and coming up with digital solutions.

Also, we try to understand the maturity of the nonprofit. Based on that we create a tech and a tool set, because it’s not that with every nonprofit, you can directly go to machine learning and deep learning. You need to start smaller and iterate fast, because that's the only way we can drive digital transformation. We start from the beginning – we ideate, we brainstorm, we understand the level of maturity. We have a long term partnership kind of model of working with nonprofits.

I deal with most of the technical stuff, the data path. We have a managing director who looks at more of the business development – making partnerships work, the way they work, etc. I sometimes lead data projects and sometimes lead brainstorming workshops. I even create curriculum for training and upscaling nonprofits and others.

# Brainstorming sessions with nonprofits to establish data maturity

6:20

Alexey

**That's a lot of information – I'm trying to process it. [chuckles] So what you do is connect nonprofit organizations – you help nonprofit organizations with analytics. That means you set up all this technical infrastructure, and even before that, you do some brainstorming sessions with them to understand what they actually need – what kind of tools they need. Maybe we can start with these brainstorming sessions.**

**Let's say a nonprofit organization comes to you and says, “Help us with our digital transformation.” How do you start that? How do you work with them? What things do you talk about in these brainstorming sessions?**

7:05

Parvathy

Usually, we try to understand the different levels of the organization. We can summarize this into three categories. We have executives who make long term strategy plans for the organization. Then there are managers or project leads or program leads or department leads, who put together teams, as well as define, deliver and implement projects. Then there are hands-on practitioners who actually implement projects. And not all nonprofit projects are data projects. Their goal is, let's say, to deliver humanitarian assistance, or to make sure that schools in the neighborhood have access to food and electricity. Their goal is totally different, but they might have a data component in their project.

During the brainstorm, what we try to understand is, “Are these different levels in the organization? What level of data maturity already exists?” Then we try to understand, “What are the things that are required in their strategy for them to grow in data science? Do they really need to invest in resources? Do they need to invest in infrastructure? Or do they need to invest in a long-term strategy and plan?” So we try to understand this through a series of interviews with different groups within the organization, and then we try to come up with a roadmap.

But also, not every nonprofit is the same – some nonprofits come to us with a really good, already defined problem statement. Then we help them implement the solution for it. It totally depends on what level of maturity the nonprofit is and whether they already have a problem statement. I would say it depends on the nonprofit, but this gives you a flavor of how those brainstorming sessions work.

# Example of an Analytics for a Better World project

9:00

Alexey

**Just curious – I don't know if you can tell us or not – but maybe you can just pick an example without naming the organizations, or if you can name them that would also be interesting and then a project that you did for them recently. It would be interesting to see how you took it from the start and then worked on that – did the brainstorm, defined the roadmap, and then worked on the solution.**

9:29

Parvathy

Yes, definitely. I can give you one example. This is not a direct example of a project because we cannot divulge all the information due to privacy and other reasons. Last year, in October, we started an academy where we trained nonprofits. It was an open, free program where we invited nonprofits from around the globe to apply – for practitioners. We received more than 340 applications from around the globe. From that, we selected 43 fellows, or nonprofit data practitioners to work with us, and go through an eight-week program.

In the first few weeks, they did some data science courses. Then in the second half, we brainstormed with them, “In your nonprofit, let's try to identify a problem that you would like to solve.” I will give one example from the United Nations environment program in Nairobi, Kenya. We worked with a fellow from there who had a problem – in Kenya, waste management is a huge issue. What they were seeing on the ground was that there were not enough waste collection points throughout the city. They wanted to use analytics and optimization models to understand how we can improve the situation. What we did is paired them up with mentors who are already experienced data scientists. We paired them up with some of our trainers from university, doing some optimization research, because it requires mathematical modeling as well.

They jointly did a very simple proof of concept pilot, where they used data that is openly available – population densities, onto which they overlaid the road network data and other data, to say “This is where you need to establish more waste collection points.” It was an AP program, of course, and the brainstorming came up with a really solid idea to take forward. Now, what we are doing is – we are having the student plus a small team supporting them to scale this up. Because that was a proof of concept and it worked fine as a test case, but now we are working with them to implement it as a proper solution that even uses, for example, satellite imagery to identify large scale dumping ground surveys to see if maybe in those areas, you need not just a data solution, but you might need more advocacy for waste segregation and plastic waste collection.

This is one example where we started with just a brainstorming and an idea, and we went all the way to actually developing a solution. I just want to make this very clear – the data part is just one part of the solution. You have other parts – you need to have public officials, you need to have advocacy for better data collection, you need to have advocacy for on-the-ground training for people to actually do this segregation. So the data part is helping the official reach there, but there is much more beyond that to get the solution to work and scale.

# The overall data maturity situation of nonprofits vs private sector

12:33

Alexey

**Yeah. I don't know if I'm right, maybe I'm wrong, but I expect that the level of maturity when it comes to data in these organizations (in the public sector) is probably lower compared to a typical internet startup – a typical internet company. Is it a correct observation? What kind of level of data maturity do you usually see in these organizations?**

13:03

Parvathy

It also depends on the nonprofits themselves. There are some nonprofits, or development sector organizations, that have really good maturity in terms of digital. They have a data science team, they might already have data scientists working with them from other organizations, they might have research teams, etc. But if you look at the overall landscape – yes, there is a resource gap. There is a lot of need for data scientists in the sector. The sector also has huge competition with the private sector, because the private sector has a lot of need for data scientists as well. Thus, the data scientists have an option to choose either to work in a public or a private sector company.

In many cases, what I see is – in a nonprofit, data is not their first order of business, like I said before. Their main goal is not to collect and analyze data, but their goal is something else and data is one part of it. Therefore, you also need data scientists who have that kind of affinity to work with them – who understand the big picture and not just start developing machine learning every time they see some dataset, but to understand where to contextualize it. There is definitely a resource gap and there is huge demand. I think I mentioned before that the academy that we set up – it was our first time setting up an academy and we are a new organization. We do not have that much visibility yet.

We do have a lot of partners and nonprofits working with us, but if you look at it, we are relatively new. Even then, we received so many applications from top organizations around the world and it really shows that there is a need. Another thing I have to mention is that it was a free program. Usually, in free educational programs, the dropout rate is high and the retention isn’t great. Because it's an eight week program, there can be much fewer people who actually graduate. But we did not have that problem at all because we really looked at the motivation and the fellows were motivated throughout the program. Therefore, we had a really high retention rate, attendance, and program completion rate. I have to say that there were extremely motivated resources that existed, but there is a huge skill gap – yes.

# Solving for the skill gap

15:23

Alexey

**How do you solve this skill gap? I imagined that, if you compare the public sector with the private sector, I suspect that the typical internet companies that are in the private sector can probably pay more. Right? This makes it kind of hard to compete, because people who graduate from university, or people who are already experienced – they might want to go there because data maturity is maybe better there, there's better compensation, and other things. So how do you actually manage to close this gap and manage to find people who want to work in the public sector? People who want to contribute? How do you do this?**

16:10

Parvathy

What we see, at least, is that nowadays, young people are extremely driven by purpose, and not always financial benefits alone. I have worked both in the private and public sector, and I wouldn't speak for the rest of the world, but sometimes when you are working, especially in the public sector, you get that sense of purpose much more than when you're working in a private sector company. It's not always the same for all companies. Disclaimer here. But what we tried to help with… and I think one organization alone cannot solve this problem – it is a huge problem and we need partnerships and we need a combined effort from different organizations to actually achieve this.

But I can talk about what we try to do to help. We have Analytics for a Better World courses, both here at University of Amsterdam, and MIT. Actually, our Science to Impact co-director is currently in Boston giving that course. We do this course both at the Bachelor level and Master’s level. This is for, let's say, training new young people to know that there is this split. Everybody is not aware that there is this split in where they can go after they do a program in data science. It’s also to get them excited that the problems are actually quite mathematically challenging as well. It's not really straightforward problems that we have to solve. It's quite mentally challenging. [cross-talk]

17:41

Alexey

**Like one you mentioned about waste management, right? It's a very difficult problem – where exactly do you put these waste collectors? It's a difficult modeling problem.**

17:53

Parvathy

It is. It's very exciting to solve it. Not just from a social perspective, but as a technical person, I get challenging problems and I get really excited. We, as an organization, want to inspire young people to realize that this is definitely possible. There are these interesting problems to solve out there. Secondly, we also see that a lot of nonprofits already have people working with data, and we want to upskill them. Plus, we also want to get organizations to invest in improving their data maturity. Because like you said before, a lot of data scientists need to feel that they are working in a more mature organization, to understand how they can contribute and they need to have all the tools and techniques and support to develop these solutions.

That's exactly why I mentioned these three levels. We have executives, we have managers or analytics translators (a very commonly used term), and then we have data science practitioners. At Analytics for a Better World, we have an academy and through the Academy, we hosted a program for practitioners last year that involves hands-on data science skills. This year, we will repeat that cohort, plus we are adding a new course for analytics translators. And in the pipeline for next year, we have plans to post a program for executives as well. But we need some more preparation time for it because for practitioners, we have eight weeks, for analytics translators, it's going to be a shorter program, and for executives, it has to be two days or so, because you're getting them engaged for more. The curriculum requires more research and that's exactly what we are doing now.

We are planning to interview a lot of executives who are working in the nonprofits to develop their curriculum. It is also difficult to set up a curriculum for them. I mean, you probably know there is Coursera, there’s Udemy – there is so much data science content out there. But how do we tailor it for nonprofits? And how do we tailor a curriculum which speaks to them and their language? It's also not very easy, so we are also learning in the process. Last year was a success and we are very happy about that, but we can still improve and we are working on doing a lot of research into developing such a curriculum for these three levels within organizations.

# Publicly available content

20:14

Alexey

**As I understood, in summary, you solve the skill gap by education. You educate people at different levels. You start with the students from universities, like MIT and University of Amsterdam, but then you also work with the companies and upskill people there on all the three levels that you mentioned. I'm wondering – you create a lot of educational content. Is any of this content publicly available? Do you maybe have courses from MIT or University of Amsterdam somewhere in public?**

20:49

Parvathy

We are huge proponents of open source, as you can probably imagine. That's the whole purpose of this. Whatever we develop, for example, for the fellowship, most of the lectures you can find in our YouTube channel. Of course, the interactive sessions and the mentoring cannot be captured completely. But the courses are there. Many of the lectures from the Bachelor's and Master's program give workshops specifically about these for students. We also do a lot of showcases of our projects for people to see. It is all available in the open domain.

If I can add one more thing, not just about the Academy, but we also have a GitHub repository where whatever we are developing for nonprofits (not the data, of course, the data is proprietary and it stays with the nonprofit) but the algorithms and models, if we see it is applicable for other use cases, we make them also openly available. We try not to use proprietary tools in general, but in some cases, we cannot escape this. [chuckles] But in most of the cases, we try to use open source. Both knowledge and any deliverables or results, we openly make available.

# The Analytics for a Better World Academy

22:07

Alexey

**That's really cool. Please send us the links to these resources, so we will include them in the show notes. But I'm wondering, maybe you can tell us more about these courses that you have with the universities. What kind of modules do you have there? What is the focus? What do you teach there?**

22:26

Parvathy

I do not teach myself. We have a Science to Impact co-directors – two professors. One is Professor Dimitris Bertsimas, he is from MIT Sloan School of Management. And then of course, in Amsterdam, we have Professor Dick den Hertog. They are the ones who are leading, let's say, most of the educational programs in universities. And we have other professors like ‪Joaquim Gromicho, who also takes sessions. So it's not just us, we have a big extended team behind us who does these courses.

Just to give you, in a nutshell, what the courses talk about, if you look at the data analytics ascendancy metrics, as we call it, you usually start with descriptive analytics, which is just knowing about what is happening and you visualize, you get the data clean and all those things. Then you do diagnostic analysis, where you see why things happen and what the root causes are. Then you can do prediction – you can forecast, you can create early warning systems, you can make machine learning models, you can do some image detection, geospatial analysis. But the last step, which is quite difficult, and it is more impactful, and that’s optimization. This talks about, “How can we make things better?”

I will give you a very quick example of what we did with the World Bank, or what we are actually continuing to do with the World Bank. We went to Timor-Leste, together with the World Bank to study how the current hospital network in Timor-Leste is and whether people actually have access to primary health care centers. We made a map of all the population points, we mapped the existing hospitals, we mapped roads, then we calculated the travel distances and we see that at five kilometers, 10 kilometers, 15 kilometers – how many people or what percentage of people have access? Now, this is descriptive. Basically, you just know that asset scenario, but then we try to do some more analysis to see why this is so. Is it just the road network? Is it that new hospitals are required?

But then, the last step is the prescriptive part, where we are actually using a mathematical model to say, “Where do you place new hospitals now?” That is a decision-making point. Because knowing all the previous steps is necessary to make that decision, but that model – the optimization model – is what will give you the decision-making capability to say, “Put a hospital in this county, or subcounty, to give access to X people more in the population. And this is going to cost you X amount. If you have 5 more, you go from 50 to 100% (just an example, of course.)” But giving that power is what we try to incorporate into all our training programs. How you can actually go from descriptive and diagnostic to this optimization model and how you can use mathematics or other things to actually come up with that decision-making capability. I think that summarizes the core of all our training programs.

# The Academy’s target audience

25:36

Alexey

**From what I understood, of course, you have training product programs and each of them has a different target audience. But the ones in these universities – MIT and University of Amsterdam – what is the main target audience? Is it people that study management or is it people who study data science and analytics? Are they more technical, or are they less technical?**

26:10

Parvathy

That's a good question. They are technical. We have, for example, an MBA in Big Data program. That's an MBA program, but they have a Big Data component. They do learn some of this. We also have a Business Analytics course, both here and at MIT – the Sloan School of Management. These are the kinds of students who take up these courses.

But we also have been brainstorming (it's not done yet) on seeing if healthcare or education or these kinds of programs can also have such components. But that's just an idea at this point. Like you asked, in universities, most of our courses are targeting students who are learning business analytics or MBA in Big Data – so both management as well as technical.

27:02

Alexey

**So they already know analytics, and you just show them the application of analytics for nonprofits. Right?**

27:12

Parvathy

Nonprofits, and also maybe a more of optimization kind of problems. So we teach them more prescriptive analytics.

27:20

Alexey

**Yes, but the point is that they already know analytics. You just show them how they can apply this to solve real-world problems. Okay. That must be quite inspiring, I imagine.**

27:33

Parvathy

Yeah, the feedback we get from students – they are all very excited. We have a lot of them working with nonprofits during their thesis as well. That shows that they do get inspired and they want to work with nonprofits for their thesis. Interestingly, we also get requests from other students who are not really from these programs, writing to us asking if they can do their thesis with us. We do encourage that as well. These are not just open to the University of Amsterdam, or MIT, or any of their affiliate universities. We do encourage other students to also come and work with us and do their thesis.

# How researchers can work with Analytics for a Better World

28:19

Alexey

**I see this question coming up in the community in Slack quite often. People come and ask, “Hey, I need to write a thesis, but I have no idea what to write about. Can you suggest good topics?” So maybe for these people, for these students, who come and ask these questions – how can they contact you and find out how they can help with their skills and how can they take part in this? What's the best way of doing this? Maybe you have a public list somewhere with possible topics? How do they go about this?**

28:56

Parvathy

Actually, this is not that straightforward, because we usually define it together with nonprofits and the problem can evolve over time. So it's not like there is a list somewhere and people just choose. It is a bit more personal, let's say. Just to give one example, we have a student who says they want to do a thesis with us. We know that some of our nonprofits have some pressing issues. We send the student and the nonprofit to a brainstorming session together with us to define a problem statement that works for both of them. Because it's not just that it should match what the nonprofit wants, it should also appeal to the students.

They are investing a lot of time, so we want them to get excited about it as well and be passionate about it. I wouldn't say it is that industrialized yet where they can choose, but they can always reach out to us. In our website, we have a page specifically for researchers and students, where they can just write a message. We always get back to them. I think you also have our LinkedIn page, you have my LinkedIn, Robert’s LinkedIn – you can just reach out to us. We are always open to all these. We are not so busy that we don't have time to answer these questions. So just reach out to us personally.

30:16

Alexey

**Yeah. We'll make sure to include all the links you send us. What I understood is – if there is somebody, a student in India or in Germany or any other country, which is from a university that is not MIT or University of Amsterdam – if there is a student and if they want to write a thesis with you, what they do is go to your website, where there is a section for students and researchers. They simply contact you, and then you take it from there?**

30:46

Parvathy

Sounds good.

# Improving data maturity in nonprofit organizations

30:47

Alexey

**Okay. I want to go back a bit, to when we discussed your different educational initiatives. One of the things you mentioned was that you want to improve data maturity of existing organizations, and you do this by upskilling people, and perhaps something more.**

**I'm really curious, how exactly do you go about this? You work with an organization and you see that the data maturity could be improved, so what are your next steps? What do you do for that?**

31:20

Parvathy

First, we need to measure the level to actually improve it. We do a series of interviews, like I said. We also have standard questionnaires, standard formats for measuring maturity, let's say. Then, using the template, we say, “Okay, currently you are here in terms of people or processes in technology.” It is not just about technology, it's about the process and people as well. Then, together with the management, we look at their strategy for the next year, three years, five years, etc. Then we say that, “Based on this strategy, this is the roadmap.”

We help them create a roadmap for data and digital transformation, “You need to upskill people around these sectors. These are the options for fundraising. These are the training programs. Also, these are the tools that you can internally use to maximize operational efficiency.” Because sometimes we see that a process is not correct. They have set up something, but it needs some optimizations there, and we help them set it up. So it's not just about training, but it goes much further beyond that. I can give one example without naming the nonprofit. We saw in the nonprofit that there is already an Azure Cloud environment setup, but there were configurations which were actually not optimized. They were spending more money than necessary.

A lot of the unused resources or clusters were not turned off, or they were still running, which ran up costs. There we had to, first of all, shut them down to reduce cost, but also we had to come up with a standard operating procedure so that the next time, anything in the organization set it up, they followed a set of criteria. “If you're no longer using it, this is what you have to do.” So it's not just about shutting it down and optimizing, you have to document it and make them aware that these are the standard operating procedures.

Again, I have to go back to what I was saying – it's totally dependent on the nonprofit. Some nonprofits require more investments in people, some more in technology, but we kind of try to tailor the whole process to the nonprofit. In any case, usually it goes: maturity scan, discussions, and then we create a roadmap. We help them implement the roadmap in stages. It’s not like you come and change the whole organization all at once – it's a long process. Usually, we have short-term goals and long-term goals in the roadmap. And we implement it in pieces.

# People, processes, and technology

34:06

Alexey

**You mentioned three aspects, or three dimensions, that you measure during the interview: people, processes, and technology. Can you maybe give us a little bit more insight into what exactly you measure there in terms of people, processes and tech? What exactly do you look at?**

34:27

Parvathy

Like I said, if the strategy of an organization, let's say, is to collect and analyze more health-related indicators. This is just one example. Say they are working on the ground with healthcare professionals and hospitals. We look at, “Are they collecting the right amount of data? Are they having the capacity to analyze this data?” That is where resources are measured. [cross-talk]

34:52

Alexey

**Is this the tech part or the people part or the process part?**

34:54

Parvathy

Actually, let's start with the people part. If you look at people – say they need people who know how to collect the right amount of data, people who know how to structure the questions, let's say, or even set up that collection system or a mobile app to collect this data.

35:15

Alexey

**Researchers, social scientists, or people who have this background of asking questions and making sure they're not biased and all that stuff, right?**

35:27

Parvathy

Yeah. Plus, you need some technical experts or data science people who can help them create a web app and connect it via an API to an analytics part. So it has to be a team including, let's say, people who can set up the questionnaire, but also digitally collect the information, put it into a server somewhere, host it, etc. You need a team who can do all this, right? That is the people part.

35:58

Alexey

**So if they have a team, then they are already more or less mature. If there is no team, then you understand, “Okay, there is a gap. If you want to collect data, how are you going to do this if there are no people who know how to do this?” Right?**

36:11

Parvathy

Yeah. Let's say that there are people who can collect data, but there is no analytics capability – no people who can analyze this data to make decisions – then they are at a different stage in their maturity. But if they don't have any data collection at all, then they are at a different stage. If they do everything, then they are totally at a different stage. So there are groups of categories.

36:34

Alexey

**I understand. Okay. And then the process part – what is there?**

36:39

Parvathy

In the process part, we look at, “What is the process that the team has to go through to set up this system? Organizationally, do they have all the support required? Are they actually looking at privacy security? Are they following all the best practices out there?” You can create a form and anybody can create a form, but “Are they actually following the best practices? Are they following the right order of things? For example, before collecting a form, is there clear guidance from the management or from the team? Who needs this data? Why do they want to use this data?” So you need to have that process clearly defined and in stages, so they know that they are doing the right thing in the right order. That will be the process part.

37:33

Alexey

**So having zero processes means low maturity. If there are some processes, then it's mid level. And if there is a process that explains everything, like you mentioned, these standard operating procedures – if they have these things in place and if everyone knows how exactly to collect data, or process data, or delete things that are not being used, what the sequences of steps that you need to do for that, then it's good. Right?**

38:05

Parvathy

Yeah. You are saying three categories, but there are many more. It's not that clear-cut. It's fuzzy. It is not just three levels – there is like a whole “level of maturity” scan. But you are right in the overall picture, yes.

38:22

Alexey

**And then the tech part, I guess, is also quite interesting. You mentioned the cloud at some point, right? So what are the stages there, in the tech dimension?**

38:32

Parvathy

So in the tech part, I would say, “Do they have a collaborative working style?” Because in many case, you will be working [cross-talk]

38:42

Alexey

**Like Git?**

38:43

Parvathy

Yeah, like Git, of course. “Are they all working on their local machines, having local copies of the data? Or do they have a centralized one source of truth? Is everything accessible? Is the tech configured properly? Also, from the organizational side, are they aware of the technologies that they require from the management?” We look at that as well, in terms of technology – not just whether they are using it, but are they aware of what the alternatives are and whether they have chosen something that is best. We look at their tech stack, we look at their analytics stack, and see if something needs to be optimized. So that will be the technology aspect.

# Typical tools that Analytics for a Better World recommends to nonprofits

39:28

Alexey

**Yeah, I kind of wanted to use this as a segue to talk about different technologies and analytical tools. What are the tools that you typically see or typically recommend that your clients use? How do you go about suggesting these tools to the customers?**

39:50

Parvathy

That also depends on the maturity of the organization, but we have a few sets of categories. We do not usually make decisions for the nonprofits, but we give them the options and the pros and cons of each, if they don't have anything, let's say. Let’s just take the example of data visualization. Within the organization, if they want to make dashboards, many organizations prefer one platform, which they are already using, and they are aware of, like Power BI or Tableau. So they already have capabilities that they use, then we try to upskilling them in that, because they usually have a shared platform and the organization has invested in it. So we try to upskill them in that.

As an organization, we do not have priorities for any technology, to be very fair. We look at the nonprofit agency and see that they already have capabilities in something, then we try to identify from our side, if there are people who can help the nonprofit in those particular technologies. And when we see the technologies, we also try to tell them if there are open source alternatives or tools, for example, Python, or others, that are very easy to configure for a data scientist, but then connect them to good visualization and dashboarding so that the managers and executives don't get too overwhelmed with the results so that they aren’t afraid to make decisions. From the technology point of view, we see a lot of dashboarding because many organizations are at the beginning their transformation journey. So they are still exploring visualizing data.

When we need to do machine learning or predictions, we prefer Python or R, depending on the organization. Then, when we have to develop a product and deploy it, we suggest different cloud options, like Azure, AWS, or GCP. Then we see which one fits the organization best. Many nonprofits already have such environments set up, and in this case we try to stick with that and go with that instead of changing the entire thing. Because, like I said, we do not really have any priorities in tools. Just to add, from the technology point of view, we see that a lot of nonprofits have different types of data.

It's not always just an Excel file with so many rows or data frames – they have a lot of geospatial data, they have a lot of text data. To handle these, they might need different technologies. For example, for geospatial data, they might need more knowledge in platforms like QGIS. And for text mining in Python itself, they need to know more about packages that specifically deal with text. So we try to identify that as well and then suggest a tech stack that will fit their requirements.

42:52

Alexey

**Just to summarize, you mentioned multiple categories – you mentioned database visualization, which could be tools like Power BI, Tableau, or some open source alternatives. Then you also mentioned machine learning, where you suggest going with things like Python, or Python packages, and R. Then you also mentioned cloud a way to deploy these machine learning solutions, for which you mentioned Azure and others. What are the other categories? I think you mentioned a few specific examples like geospatial data, text data, and for them, there are specific tools. Do you have any other categories when it comes to maybe collecting data or storing data or something like that?**

43:41

Parvathy

We currently see these categories in nonprofits, like dashboarding, visualization – we have Python, R and of course for deployment. By other technologies, do you mean defining technologies or defining categories? What exactly?

44:01

Alexey

**What kind of categories are there? I imagine that you need to keep the data somewhere, so maybe you have some data warehousing solutions. Or maybe you need to collect data, so you need some data tracking solutions. What are the categories you look at?**

44:18

Parvathy

For data warehousing and storing data, a lot of nonprofits prefer open platforms. There are solutions specifically made for nonprofits, so they tend to use that. We see a lot of PostgreSQL databases because it's open. We also see KoboToolbox being used, actually, which is part of the Digital Public Goods Alliance. It is meant for nonprofits to collect data in a structured format, even including images and storing them securely. KoboToolbox has a humanitarian section which is built specifically for nonprofits, and it's registered in the Digital Public Goods registry.

So there are specific tools that nonprofits do use and prefer, and they are specifically targeted and made for them, which is amazing. They all follow what we call the nine Principles of Digital Development. That's one of the resources that maybe the viewers will also be interested in seeing. It's about developing solutions that can scale, specifically in the public and nonprofit sector. There are tools like KoboToolbox that are very frequently used. Also, like I said, for data collection and databases, they use PostgreSQL quite frequently as well.

# Profiles in nonprofits

45:51

Alexey

**I’m wondering what kind of profiles there are for these tools. I imagine that you need different profiles – data analysts would be usually more involved in data visualization and using these data storage databases like Postgres to analyze the data, maybe also collect data. Then, perhaps, for the data scientist profile, the data scientists are more involved in machine learning, training models, and also getting data, deploying them to the cloud.**

**Do you have other profiles like data engineers, or deployment engineers, or some other people who are also in the data world?**

46:36

Parvathy

This is also totally dependent on the nonprofits themselves, but many of them that we see, if they are at lower levels of maturity – you see a lot of unicorns, let's say. [chuckles] They try to do everything because there is a resource gap. These roles: data scientists, data analysts, data engineers, architects, cloud engineers, MLOps – these are all somehow merged together. There is not a clear definition – I think even in private sector companies, there’s not always a clear line. So it is not always very clear what rules there are. But if you look at a particular project, I can tell you that in a team, there are always people who are, let's say, collecting and analyzing data.

There are always people who try to make business decisions out of the data – who are trying to understand what these models are saying and putting it into the context of operations. And usually, in nonprofits, we see a lot of research teams who are looking at it from a totally different, research perspective, because they need to know what more data has to be collected and what more use cases can be developed. We see a lot of data science profiles, but to be very fair, most of the nonprofits are not at the stage where they are deploying large scale machine learning models on cloud. It's just not there yet, at least with the nonprofits we work with, or the development sector organizations we work with. This is not what we see. Most of them are still defining key performance indicators, so it's still descriptive, and in some cases predictive.

Just one example is an early warning system. You might have heard that the Great Horn of Africa has had the biggest drought in many, many years. How can you predict or give early warning for these kinds of situations so governments can act well in advance? So early warning systems – a lot of prediction on weather and climate using satellite-based observations and historic data – a lot of these things are done. But when it comes to deploying to large-scale models, I do not see a lot of such applications. Right now I don't see a lot of MLOps or cloud architects in the sector yet.

# Does Analytics for a Better World have a need for data engineers?

49:15

Alexey

**I understand. Well, the reason I'm asking about this is that right now, we have a course. It's happening right now – we're on week four. The course is about data engineering. In this course, we show how to use different tools for collecting, storing, and processing data. The focus is more on, “How do you move data from one place to another place in a reliable and reproducible way?” Also, as part of the project, they will need to create a dashboard. So what I want to ask is – if people with this kind of profile want to take part in some of your initiatives, how can they do this? Or do you even need people with these kinds of skills?**

50:06

Parvathy

We do need them, definitely. But it is not like nonprofits are going to directly hire them. Not all nonprofits – there are some who need this, but because there is no constant source of these kinds of projects. But we do see some of these applications and I can give you one very simple example. We work with a researcher who has developed a very amazing solution for smallholder farmers in Brazil. It's a very impactful model that they have demonstrated on the ground. But how do you build that gap between research and then scaling and deploying it? We want to create a mobile application that will fetch all the required data and process it – immediately give results so that they can immediately act on it. This has to happen, because it's not just visualization – we are running machine learning models at the backend to give those results, right? We were looking for the right set of people to support such a project.

Another example, as I told you about, in Timor-Leste – we went and we were optimizing for the location of healthcare facilities. We saw that a lot of these use cases recurrently appear. For example, we worked together with the World Health Organization during the COVID pandemic in Nepal and they wanted to know, “Where should we place COVID test labs?” So it's a similar problem to what Timor-Leste did, but the data stack is different. We are actually building a web application where in the backend, these optimization models will run and give you results depending on the data and the country that you feed it. For bigger countries, it is actually a very resource-intensive problem.

If you look at Vietnam, we tried to do this in Vietnam for healthcare facility access – the population is huge. It's a large area to run this optimization model. It requires a lot of resources. So yeah, we do need machine learning engineers for specific projects and we need these deployment capabilities. We are always looking for such people who want to join this mission. But also be aware that it is not like we are constantly looking for these resources. It comes when such interesting projects come to us.

52:31

Alexey

**Do you have any open positions right now? Not necessarily in data engineering, but data analytics too. There is a question from the audience whether there will be any open positions in the future. But I guess you can also talk about them if you have them now.**

52:50

Parvathy

Right now, no. Not as far as this morning, let’s say. [chuckles] Things change very fast. But what we also try to do is – we have a list of people and resources that we can tap into, so when a project comes, we can reach out to them. Also, like I said, we have a lot of private sector organizations who are our partners. One of these examples is Ortec. Ortec is one of the founding partners of Analytics for a Better World, together with the University of Amsterdam. They are a mathematical optimization company.

We also are currently working with a lot of other organizations that support us. When such requirements come, we first reach out to them and see, “Do you have resources who can work on this project for four months, (or five months, or sometimes one week or two weeks, depending on the project, of course).” But we would really love to welcome others who would like to contribute as well and see if we can work together. I'm always looking for opportunities to connect with people who are inspired to join us.

# The Analytics for a Better World team

53:58

Alexey

**And the first part of the same question is, “How many data analysts currently work at Analytics for a Better World?”**

54:07

Parvathy

Like I said, we are very new. Last year, we set up and we currently have a core team of, let's say, four people. That’s the core team. We have our managing director, Robert. I am the CTO. We have Prof. Dick den Hertog and Prof. Dimitris Bertsimas who are Science to Impact co-directors, and they lead Research Academy and other ventures. We also have other data scientists who are working kind of full-time with us. One is Claudia, who’s based in the University of Amsterdam. We have Brett, who's a PhD student. We also have other researchers supporting us. But we are a very small core team.

However, we have a huge extended team of researchers from universities. We have supporters from our organizations that are supporting us. Just to give an example, during the Academy, we had more than 20 professionals from the private sector giving courses. All of them came from our network – from the people who are in organizations – and if I can look at the number of projects we currently have around four or five running in parallel. All of them are staffed with like three or four people, but it's not always data scientists. We sometimes have data scientists, data engineers, and others. So we do have an extended network and a big team, but our core team is very small.

# Factors that help organizations become more data-driven

55:38

Alexey

**Thanks. We have a question from Antonis, but I should have asked this question earlier. I don't know if you'll be able to answer it in two minutes, but we can try. The question is, “What are the most important factors that help organizations become more data-driven? Organizations that you worked with and organizations that, perhaps, start from zero data maturity and then eventually become data driven – what are the most important factors that help them?”**

56:10

Parvathy

I think a long-term strategy to include data in their portfolio is one. You asked “What are the factors?” If I have to explain that, two minutes is not enough. [chuckles] So I'll tell you what the main factor that we see is. This includes organizations that have a strategy and an understanding that data is required, organizations that are motivated to come up with projects, and want to upskill their people to become more data-intensive in their work. They need to make the right investments. And for investments you need drive from the organization perspective as well. Like I said before, the investments have to be in three sectors: people, processes, and technology. Understanding this as an organization is the most important thing.

We see a lot of value in data and analytics being added into the public sector organizations. We believe, truly, from what we have seen in the last year, nonprofits can benefit equally from data. But as an organization, most of them know that data is important, but have not made it a strategic part of their goal yet, in many cases. I think that a clear vision and strategy to include data is the first step. After that come all the investments on people, processes, and technology, as well as a clear understanding of why and how data can be used. And that's what we try to do with our Academy – making them realize “How can data help?” That's not always very obvious, I think.

You don't immediately see results, right? Especially in nonprofits, you cannot see the healthcare system in a country improving in one day – it takes years to see real impact. Of course, there are outputs and outcomes from programs and upskilling initiatives, but to see the change, it takes time. Having that patience and long-term vision is equally important. Thanks for the question. It's a really good one. I wish I had more time to answer it. [chuckles]

# Parvathy’s resource recommendations

58:22

Alexey

**Yeah, I should have asked that a bit earlier. Maybe one last question. This is usually the question I ask everyone – do you have any good book or resource that you can recommend to the listeners? I think one thing you mentioned is all these educational initiatives that you have. That would be a great resource that you probably will recommend to the listeners. Is there anything else that you would add to that?**

58:50

Parvathy

Resources? For, I always read a lot of technical stuff. For example, I spend every day reading at least 30 minutes in Towards Data Science. [chuckles] It's very techy, but I really love it. On books, I can recommend a book that I just finished reading. It's called The Culture Map: Breaking Through the Invisible Boundaries of Global Business. It's about how different people from across the world – because we usually live in our bubble and we know what is happening in our neighborhood, we know what is happening in TV media, but world is so much bigger and in different cultures, you have to understand the context of how people think, lead, and get things done. It's actually a really good book called the Culture Map. I am also currently reading a book called The 7 Habits of Highly Effective People. I also find it very interesting. I do have a lot of technical resources, not just “our” resources, let's say from Analytics for Better World, but others as well. They could be interesting. I will share this across with you and maybe you can send it to your subscribers. I think it will be useful.

# Conclusion

60:00

Alexey

**We will definitely include all the links for all the resources that we talked about in this interview (and I think there are quite a few of them) in the show notes, in the description. With that, I guess, we are a bit over time. Thanks a lot for joining us today, for sharing your experience and expertise with us. Thanks, everyone, for joining us today, for asking questions. It was fun. Thanks.**

60:27

Parvathy

Thank you so much for the opportunity. It was really good chatting with you as well

60:31

Alexey

**Yeah. Have a great weekend, everyone.**