Transcript of Not Loud Enough Podcast

Can tech and culture work together to create meaning and question data? Feat. Aurore Paligot – Episode 9

Canan Marasligil

I'm Canan Marasligil in Amsterdam,

Laura M. Pana

and I'm Laura M. Pana in Vienna.

Canan Marasligil

And this is Not Loud Enough, a podcast that delivers authentic conversations about actions we take to build a more inclusive and empowering world, brought to you by two very good friends and their guests across various industries.

So today on the podcast, we are welcoming Aurore Paligot. She holds a PhD in linguistics from the University of Namur in Belgium and a specialization in advanced analytics and machine learning from the Ubiqum Code Academy in the Netherlands. She has more than eight years of experience working with qualitative, quantitative and experimental research methods and is currently working as a data analytics consultant at Positive Thinking Company and she will tell us more about all this. As a linguist turned data specialist she cares about delivering meaningful insights and human centric visualizations. Her research interests currently lie at the intersection between Tech and Social Sciences and part of her free time is spent exploring Digital Humanities. She's also an R Ladies enthusiast and a keen Tableau developer. Oh, you're gonna have to tell me more about that, because I have no idea what a Tableau developer is. Welcome to Not Loud Enough, dear.

Laura M. Pana

Thank you. Welcome, Aurore. Hi.

Aurore Paligot

Hi, Laura.

Canan Marasligil

We are so happy to have you because this is something Laura and I actually know very little about, all these topics. And we are very curious about your experience. But first, we would like our listeners to know more about you because you have a very interesting story. Because what brought you to data analysis from the world of linguistics, from the world of Academia, from Belgium to the Netherlands, and now back to Belgium? We want to know everything about you Aurore, please tell us.

Aurore Paligot

Well, thanks for having me with you, connecting Brussels, Amsterdam and Vienna together, it feels really good. And I'm happy to share a story that also Canan you've been involved in it for the last chapter. So, it's really nice to think about all the things we shared and how I arrived where I'm at now. A Tableau developer, what is it? Maybe let's start with that. That's my main activity, actually, at the moment. So as a consultant, I'm now

working in a telecommunication company. And my role is to develop metrics and to visualize that, to report them. And Tableau is the name of a business intelligence tool that delivers insights through data visualization. And so, I'm really busy with building graphs at the moment, but also thinking about my own practice, as you said, coming from this background in literature and how I can use data and data visualization to build meaning, share meaning and also question the data itself.

So how did I get there? Well, first of all, the first question is maybe how I got into Academia because it's a story by itself. It was really not part of the plan. When I started my studies, it was well, more than 10 years ago, I started a bachelor in [French and Spanish] literature. I was really hesitant when I started. I wanted to be an illustrator. And then I also wanted to be a biologist. So, you know, arts, literature, science in a way now, as I reflect upon it, are starting to mix again, which feels really nice. But then I started this bachelor in Namur, I was the average student: interested, I liked literature a lot. It's played a huge part in my life overall. But then I had a really impactful course about linguistics and I would say that this is where it all started for me. And when I mean, started, I mean, seeing myself as a researcher, which is something unknown. Well, from where I come, I think I was the first child of my family to go to the university. So, it's not like we have a representation of what being a doctor means. So, for me, being a doctor, you could either be a medical doctor or a doctor in law, that was my full representation of what it was. But then, luckily, I was in a small university, but that was really involved into research and also had a special program about linguistics, sign language linguistics, and also a specific, a very strong, let's say, theoretical background there about linguistics. It was really, I can say, an intellectual revelation in the sense that I can say that it changed the person that I am now. And then, in fact, what really got me into linguistics is already working with data, linguistic data, as a way to ... in a very clinical way. So that language was the medium between us and the word that we were studying by modelisation. So, all that we studied came around this idea that reality, or the object does not exist per se, but is built by the gaze of the observer. And so, for me, that was a really meaningful realization. And with that, I had a strong formation also in other social sciences, disciplines like philosophy and an epistemology. And it all came together for me realizing that this is what I wanted to do. And then I decided without hesitation to follow a Master in Brussels, about linguistics, that would prepare me to become a Researcher.

I received a grant to pursue a research with a project that was starting at the time about documenting French Belgian Sign Language, which is the sign language spoken by the Deaf community of Brussels and Wallonia. And I was working there under the direction of Laurence Meurant who is doing an incredible work. And basically, we were a group of four, five researchers and Deaf teachers, and then Deaf annotators and translators, and little by little the team was growing around this idea that was gathering us : how can we represent, how can we document in a fair way, sign language, to study it, to archive it, to use it as a linguistic tool to teach, etc, etc. But we did... I knew nothing about data. And neither did my colleagues. So, it all started by us recording Deaf people in the studio, which was, let's say, not the easy part, but a part of the process that is quite easy. But what is more complicated are questions like: who do we invite in our studio to create this corpus? And then how do we transform all this videos into a machine-readable material that we can then study? And so, it was all a fantastic adventure, and really, we had the hands on the data all the time.

And I had the chance back then, it was in 2015, to go for a five months internship at the Radboud University of Nijmegen where I collaborated with the team there. And they already had a more advanced corpus than ours. And then it was like a gift, like, all the data that we were building in Belgium, I could play with it, start to play with it in a more advanced state, and then for me, I was feeling like a child. "Okay, so this is it when we can go until then and when we can ask those questions. So, we can ask questions in a way that was not possible until now in sign language research." So that was fascinating because without the data well, linguists of course, were limited.

[In Belgium] some connections were starting to be created between us and the IT faculty. I was feeling really at my place in between those worlds, between languages and tech and having this curiosity about the data and then understanding both languages. And then when I finished my PhD in 2018, I just realized that for the new questions, the new research questions that I had in mind, I wanted to, yeah, to go… to have new tools in my hands, that would really allow me to answer questions that I was thinking, imagining, but without having the technical possibilities. I had the ideas, but let's say not the tools. And also, it was a time, I would say after my PhD when I was not sure if I wanted to go further for a specialization, follow a postdoc or do something else. And then I really decided to take the time to explore something new, maybe reopen the perspective because doing a PhD and a research is really an act of focus, and then you specialize and then ... and I wanted to reopen maybe a few perspectives, but also, yeah, I wanted to be surprised, maybe. So that's how, Canan, you entered my story.

Canan Marasligil

Very happy I entered your story, because you moved to Amsterdam to study this at Ubiqum. How was that experience? So, because studying in Academia is quite a very different way of just studying and learning and getting knowledge then what you did here for six, it was a six months program, right?

Aurore Paligot

Yeah, yeah. Yeah. Well, I really enjoyed it. I really enjoyed it, because of the philosophy of the program, and this is what I was looking for. Yeah, that's what I wanted to say. Back then, when I finished my PhD, I was at first looking inside Academia of ways where I could get those skills. And I searched for a long time, but nothing was really quite what I was expecting. So, this is when I went for Ubiqum because it was different. It was practical. And it would give me, it would let me be in a context with people coming from different backgrounds and really have this tech feel, atmosphere, that I wanted to discover. I was feeling that there was something there. But yeah, I loved the bootcamp. The program was really much... well it was super intensive, but really enjoyable. It's really precious to have for once a parenthesis in your life that you can just dedicate to learning a new skill. And of course, you have to organize yourself because when you 30 you don't decide to take six months leave just for learning like that. But it was really worth it. And what I loved is that it was really a learning-by-doing experience and it resonates with my way of learning. When you're in Academia, you don't have a path. So, you're always confronted to errors, bugs and questions. And the questions are the drive of you wanting to go further. And what I enjoyed in this data program is that I realized that yes, data is really a way to open new questions.

So, for me, it sounds like a perfect fit: always having something to feed my curiosity.

But also, much, much more I was surprised way beyond what I think I would be at the beginning. So that's the fantastic part of it, I would say. First of all, meeting the people, having the time to go round in Amsterdam. The thesis was done so I didn't have to feel guilty about you know, going out and not writing the last chapter. But the thing is that I realized ... my ambition in the beginning was to learn how to be autonomous with me doing statistics to be a good researcher. And then I realized the data was way broader than that, that I could with it collaborate with doers, makers, create applications, create visualizations, develop new technologies. So that the possibilities and their combinations were in fact endless.

And I felt really inspired also by the tech community that I started to explore. What played a huge, huge role for me at this point was going to programmers' group and I joined the R Ladies group, which is a group of ladies, well not only ladies, but it's organized by ladies and with a strong emphasis on diversity, inclusiveness and empowering women in tech careers' that are still underrepresented there. And R is the programming language that I was learning, it's really well known for data science. So, there was this group and then also the Women in AI community that pushed me even further. Women in AI is focused… has… let's see, R Ladies is great for the workshops and everything that you will learn and how you can gather and then be geeking around and then you know, I like it. And Women in AI is also about having a vision and dare seeing yourself as a founder, and a creator. And it really, I think, it gave me a shift of perspective. Because in research in a way, you always have to wait one year or two years to dare starting your program or your research. And here the focus was on "but how? Okay, so you have an idea and your ideas can have an impact in the world." And then how can you find investors, for example. So, it's a really different mindset. Also, very new to me, that challenged me quite a lot. I do not feel like this is the path for me right now. But I have this in my mind, you know, so just for this, it was a great, great gain.

Laura M. Pana

You talked earlier about working with the sign language or in sign language. And hearing you talking, I have this first question coming. And that is I wanted to ask you, how inclusive, how included are these communities in the creation of the technologies that are prepared for them?

Aurore Paligot

Yes, that's a really good question, a really, really good topic and also a difficult one, because it's linked to the history of Deaf people. So usually Deaf people have been and are still excluded from the conversations or when they are included, their voice is still unheard. So, it is a really important topic that we also have in the sign language linguistics community. How do we make sure that becoming a linguist is accessible to Deaf people, that education is accessible to Deaf people, that the way we organize our conferences are also accessible? So, these are really important questions. But concerning the technology, well, of course, AI is fascinating. And tech is fascinating, because it gives us the dream that everything will become possible. And we've seen all the progress of automatic translation, for example, and I think that it's really amazing.

But about the sign language history, you know, everybody that, like me, has been working with sign languages, receives an email of a friend or a

colleague thinking about us, like "hey, look at what I've seen in the news. And there's this great new invention and it's a magic glove and it's gonna translate sign languages automatically". And you know, I'm still on this very polite way of saying "well, thanks for thinking about me". But then now the linguist in me wants to say "hey you know how much I like AI but let's reflect on the concept of the magic glove".

Well many of these inventions gained, along the years, public attention, because regularly a tech student, specialist in AI would think about that, as "Okay, so we have the technologies, we can put sensors and capture information about the movement of the hands, and then we can... "translate" is not the proper word, but we can match those gestures, those shapes of the hands, of the fingers with the equivalent of a sign or of a letter that is then considered as a translation of a word [or letter]. And it's really fancy, because it works, right? You have the glove and then you can finger spell. So, for the people who do not, who cannot see me: what is fingerspelling? It's a way to use your hand to visualize letters. It's like a visual alphabet. So it's the A, the B, the C, so you have different way of fingerspelling in, for example, in American Sign Language you use one hand, and you see that the C shape looks like the letter C, but then in British Sign Language, it's a bit more complex, and then you use pointing of one finger on the other finger palm to do the same thing. But so, this magic love, the idea behind it is that you would put it in and then you would recognize the shape of the C and match it to the letter C. Right? But how is that improving communication? And for who? Which is also an important question.

First of all, the idea is based on the misconception that sign languages are only using the hands. But sign languages use the whole body that has a grammatical value, and the space has a grammatical value, the eye gaze, the movement of your eyebrows, and the movement of your torso can have a real grammatical value. When I say that, it's not just for making it sound like complicated, but because it's really… like the direction of your gaze can define different grammatical values. If you raise your eyebrows, you can have a marker for asking questions, for example. So, it's not anecdotal at all. And so even though if we could, which we cannot at the moment, capture perfectly the movement of the hand then we would totally miss the other important aspect of the embodied language. So, first of all, there's this.

And then, well, of course, you have the question of how much data do we need to train the algorithm. But let's imagine that's not an issue. The other issue is an understanding of the needs of the Deaf community. And so, when we focus on the glove, literally, what we say is you as a deaf person, you will have to adapt your signing pace so that my machine can recognize you so that I can have a translation. So yeah, what about doing it the other way around? And there are some technologies like working the other way around. And so, I'm thinking about a guy who started a tech [start up]. He has a deaf sister, so he knows about the context very well. And he started to do a live captioning system. So, you are deaf, and then you can go to the hearing person who has to adapt, and then articulate a bit more clearly. And then you can have some kind of live subtitle. So, you know, it exists, but you have to know the concepts.

And more importantly than that, what is important now for the deaf community is really having the resources to have access, fair access to every areas of society. And for that you need interpreters, sign language interpreters, and for having good quality interpreting, you have to have material or to train them. And I believe that having a good data like the

corpus that is being built at the University of Namur is a good example of how we can use tech and data and linguistics to provide something that is relevant for the community and for the community itself.

In the first stages of my research, we were building this database, that is a library, but the story doesn't stop there. The idea now is to make of this library or a full bilingual, contextual dictionary, a bit like the one that you have on Linguee, and to make it accessible through text, but also true sign recognition based from a webcam. But the idea here is so that: Imagine that you're a kid at a deaf school, and then you want to know how to learn a bit more about French. And you have a sign in mind, you don't know how to translate it and you want to access the full nuances and diversity of the language. Then you could use this interface to enter the dictionary and see different videos of real contexts of signs, of other signers, and how they have been translated in French and all the nuances of the language. So, it's not like we have a letter and we match it. It's like we enter the linguistic process through this tool, having the possibility to say for example … that's the one that's on the introduction page of their website, but for example the French word "personne", you can have different meanings rights? But also in signs, you can have different signs. So how can you magically match one another? You cannot. But what you can do is learn the language through putting side by side the different nuances of the word "personne" or the sign "personne" and all their different meanings. And that's the perfect combination of what can be done with the inclusion from the start of the community that is involved, of its primary beneficiaries, let's say and, yes, the specialists that are not only… of course there is Anthony Cleve from the university, the IT faculty who is responsible for the AI and the tech parts, but then in collaboration with Laurence Meurant, who is the head of the [sign language] linguistics department, and I think that's the way to go. So, I'm really looking for projects that are bringing people from all the different worlds together, I think that's the key for more inclusion.

Laura M. Pana

I totally agree. And then going back a bit to what you just said that maybe we also, everyone else can make an effort to have more accessibility. I'm very interested, for example, to create accessibility through the content. There are now tools also on social media, Twitter Able, for example, and there are little things that everybody can do that can help to include these communities also. For example, I don't know that when you write the hashtag, you just write camel cases, for example, and that makes it easier for these people to read, or be more conscious about how to use emojis. Because the same, they are, the description of the emojis is very different than when you write, when you include like thousands of emojis in a text, it becomes really hard to be read, or the usage of alt text for images, so description of the images. So I started to be more and more conscious about that, and starting to get a bit more into that and make my content more accessible and the way we communicate the more inclusive and more accessible as well. So this, Yeah, it's very important, I think.

Aurore Paligot

Yeah, that's really, really interesting, really important. That's something I try to be mindful of, as well. And I see that on the internet, more and more creators are proposing subtitles, for example, for the videos or transcripts of their recordings, and then you can rely on tools like voice recognition to make it quicker because of course, when you are an independent content creator, you have always a question of the resources

and the time that it takes. But I think it's really nice to Yeah, to have this reflection.

Laura M. Pana

True, because these are small things that everyone can do. And everyone can do it in through the social media. So, it's not much effort from our side, I think it's just a matter of being more conscious about it and just apply it. And I think a lot of people are still not aware of this, or companies. I saw now there's also a discussion, more companies are starting to get interested into that and making their websites more accessible, their content, and so on. So, this is great to see that as well.

Canan Marasligil

If we come back to the work you do with data and what is actually data, what are we talking about? Whose data are we talking about? who collects that data? What do we do with that data? Who creates that data? Do I create data? What happens to the data I create? What do you do with it as an expert? All these questions for me, it's so big. Sometimes we get lost, we say data, data, data, and then data controls everything, AI is controlling our lives, or algorithms are deciding him. You see, all these words we use daily, and I'm not sure everybody actually knows what we're talking about. So, could you maybe guide us a bit into this and explain us what you mean when you talk about data? What do you mean about data?

Aurore Paligot

Yes, well, data, that's true, it's a buzzword. Even when I described myself as a data analyst, it's a bit of a rebranding, like I'm a researcher. I know about data and the research process because of my work, and now I know some programming so I called myself a data analyst, you know. [Laughter] But yeah, what is data? I think that's not a trivial question. Because I think the data is behind everything that we do. Well, I think it's the only way in a sense, only way to approach reality. Because there, as I was saying, before, I don't think that there is reality, per se. So we need to, to build something. I see in data a real… a gesture, a gesture of cutting something, deciding to put something in and something out. And based on that you have decisions. Or not, sometimes you have biases, that are unconscious, but you can have conscious decisions to select the data.

And then I would say the data is what will feed your artificial intelligence. There is an example that I think is really funny about AI. While I don't believe that AI is going to rule the world, I believe that we do have the choice as AI users, as AI providers, inventors, to embed them in an ethical conversation. And this is my perspective. And this is why I, in Amsterdam, I chose to collaborate with actors like Tapp, or the City Innovation Exchange Lab, that are all about using data to build smart cities, but then integrate them in an inclusive and ethical view about the society.

But so yeah, going back to my example. All right, so there is a lot of artificial intelligence that is being built in the realm of visual recognition. And there was this algorithm that was designed to recognize fishes. Alright, so a quite simple task. But then we gave the AI certain sort of data. And in fact, the database was a database of fishermen that were proudly holding a fish in their hands. So it worked very well. Because then what the machine learned was to, you know, recognize the fingers holding the fishes. But how do you recognize a fish in the sea? If it's not that fish in the hands of a fisherman? So, this is what it is, when we say

it is intelligent, you know, this is the intelligence we create, and it's fantastic, but is as limited as it is. So, all the power we can give him is through the data that we feed him. Thinking about this, this data ... I don't have all the answers. But I believe that it's really important for people from the humanities and philosophers to take, to be active or, for the citizens to be active in this conversation and discuss what we put in the data and how we make intelligences that are not maybe that limited in a way.

Canan Marasligil

I had these discussions a lot in the literary translation field, you know, because translation, AI translation is actually doing very, I think it's working really well, in many, many ways. When you use the tools like DeepL, or Google Translate, it's extraordinary what they, what the machine is capable of translating, you know, really complex sentences. And so it is getting scary. As a translator you think maybe in 50 years' time, they won't need me anymore. But these discussions we have actually, it's very interesting what you just said about the need to discuss between different disciplines. So, we cannot leave the discussion only to data scientists or to programmers or to IT specialists, we really need to include the humanities, we really need to include people working in literature, we need to include people working with languages and with you know, this is really important. And I think this is something you advocate for throughout your workout.

Aurore Paligot

Yeah, this is really the key message. Going back to where I come from it's Yeah … You know, at the beginning when I was starting to look for a job as a data analyst slash scientist, I was a bit ashamed of saying "Yay. I have a bachelor in literature, you know." And then well, thanks for me, I got this great experience working here next to them with Tapp, which is a smart city initiative that builds prototypes for the city of Amsterdam, and other cities as well. And then I realized that, in fact, it is what brings me value when working with other data specialists. It is the very thing I was ashamed of that really is bringing a different perspective, I think, on the table. But it took me time to recognize it, but now I cannot go back and unsee that.

There is a great book I just received this morning. It's called "Data Feminism" by Catherine d'Ignazio, and Lauren Klein, and I haven't read it all, but I just found a table that really sums up how I see it. And it's a table between what's the difference between data ethics and data justice, with the idea that we can do more. And so, the authors are comparing every time two sides of the story. I really like this one that says, okay, so if you want to have a concept that secures the power, you can just understand the algorithm. But then if you want to integrate concepts that challenge power, then you also need to understand history, culture and context.

Canan Marasligil

What can you do with that data? One of the things we can do is data visualization.

Aurore Paligot

I love thinking about data visualization, because it's really powerful. And you have this say: "An image is worth a thousand words", which I partly agree with and I partly disagree with. Yeah, I agree with it because yes,

you can represent things. Well, of course, through sign languages, one of the things that I really like is visual communication. And I think that visual meaning is something that is really strong. But what we forget is that first of all, numbers are not just numbers, they are not reality, it's a construct. There are decisions that are made on those numbers. And the visualization is, well, a point of view, a choice. Well, there is really the idea of we add a layer on top of that.

I have the idea that there is a lack of education about data literacy. It's not because that it's visual, that it's transparent, that the meaning is… that you can read it as is. So, you have to have the keys to be able to read it. So that's one of the aspects towards inclusion. But also, from the point of view of the data designer, let's say, there is the idea that you are the one making the choices of what is the story that you want to tell and two maps can be a valid representation of the reality, you know. But what do you want to know? What are the fine grains? What are the nuances that you want to add? Or? And so these are ... but also sometimes even more basic things like, is it readable? Or is your public, going to more aesthetical things, but is your public color blind, for example? How do you read it? How do you read? Did you include it in, in a way? Or did you put a legend? Do you explain what is the number behind it? And then can people themselves play with the data? Now we have so many tools, and that's why I like Tableau, but others as well, where people are given the possibility to play with the data and create the story.

So, for example, we can imagine having this graph and going for it and looking for the story behind it ourselves. And data visualization, I think is really important at two very important points of the data generation process such as, first of all the exploration part of it. And then after that the communication part of it. And my idea is that we should encourage people, well I mean I'd like to generate curiosity in people to share, to look themselves for those visualizations. I'm thinking for example about Google Trends. It's a really easy way that you can play with very visual data and you can look for numbers, trends across different countries, different periods of times. But then, when playing with that I was thinking a lot about data visualization, how to make it meaningful. And it's not obvious, like the visual can give you answers, like part of the answer, not everything, but part of the answer. It can tell you, "hey, there's something that's happening here". And this is really important in the knowledge process to have always something to ... that grows, where your mind can focus. But then to make it relevant, I think that it's important whenever we have a visualization, to make comparisons. It's really difficult, I realize, to talk about visuals without having the visuals with me to show the examples. But the idea is that, yeah, you have to make it meaningful, like giving to the people the possibility to know what is the data for and what is it that you are representing. Are you talking about humans? Are you talking about rates? We don't know. So it's important to communicate about this.

Canan Marasligil

And you've been doing a lot of interesting work. One of these includes social media and collecting data through Instagram, on the Marineterrein here in Amsterdam, so in the center of the city. Can you tell us more about that specific project and what was the role of using social media and for what purpose and for who?

Aurore Paligot

So yes, this project is portraits of, how I call it, a portrait of Amsterdam Urban Beaches, and I came up with the idea because at Tapp we're working with building Smart City prototypes. And so addressing architecture and urbanism and mobility in a modern way that includes both the citizens and the technology. And I was interested to see also the perspective, to bring the human perspective on the table. So this is really me with my Digital Humanities question, let's say, I wanted to understand how using social media data can inform us about how citizens living in those smart cities inhabit those places and how we live in there, how many communities are sharing the spaces? And what do you do in those spaces. And so this is a bit provocative thoughts that we don't want to anymore have an architect that brings a project that he has carefully thought and then that brings it on the place, but more we want to understand the ecosystem of have a place to create something with it. And our question, Tom and I was, could we use this as a tool to understand and measure and inform our Smart City projects. So it was an open question, but one that I really liked, because I could dive into my favorite topics. So going on to social media, using text mining, so this is a technique to gather textual data, and start to observe what is really happening on Amsterdam. So we gathered data from the Marineterrein but also the Pllek, Entropotdok and other hotspots. And it's been a really nice journey and results I'm happy with.

So we made three visualizations that you can play with, to see how people relate to the places thanks to the hashtags, the evolution of them. And personally, I was really intrigued because I discovered things I didn't know about Amsterdam, thanks to this project. For example, I didn't know that the Marineterrein was also the Gay Beach of Amsterdam, and that you had a really active community of queer people gathering around there. But then you also have frictions between different communities, for example, the queer community, and then you have this strong, this big ship of the Scheepvaartmuseum that can be seen as an outdated symbol of Dutch colonies [by some people]. There was a Police community that was really active on another place. That was totally unexpected, and they were engaging with the population. So this is for the people, inhabiting the places but then you also have the question of what we do in those places, what we use them for, and what's our representation we have about them. And this is really what inspires me.

I've been replicating a similar idea for getting to understand how we live in the cities in times of Corona, in fact, looking into the data. And I think it's really interesting, I could do so using this technique going on Instagram. I'm gathering a lot of comments of photographs that I anonymized. And then I look at what people say and how they have lived the lockdown, for example. I observed different attitudes. You had like a sideration in front of emptiness. You had many posts, just focusing on how empty are the streets going outside and then creating almost a genre of the void, emptiness of the city. And then you had all the other representations of the city from the inside, from your house. We are inside now, we are locked down and we stay at home. And next to that, I also hum… Yeah, in those two positions, like I found a sort of romanticization, almost, of how the new city looked like. But then you also had another trend that was quite contrary of people feeling really bad about it, not posting anything related to the present, but then going to those famous hashtags like "Throwback Thursday", and just putting in photos of a really crowded Amsterdam and how we miss these days. And so this is something that I like to observe.

Canan Marasligil

I do have a question about this because it's super interesting, but when you talk about looking at Instagram only and at representation, I want to ask you how representative is it actually, because not everyone is using Instagram,

Aurore Paligot

We can look at different data. For example, on another project on the Marineterrein, I've seen that there were many dog owners on the Marineterrein. I don't know if they are on Instagram, maybe I should check. But of course, it's always just a partial representation of the reality. Of course, I'm from the younger generation, and I'm an instagramer. So for me, it's a playground. And what I like to include in my project, is this dimensional… how we live in a place, both online and offline. So and how they mix up and what the layers that are composed of. But of course, I'm yeah, I'm aware, it's a really partial version of the reality. There is a big part of the population that does not have an online presence, and that's totally fine. And that explains, for example, why an app like the corona app is not working because it would mean… well, it's only one of the aspects but yeah, do older people have smartphones? And are using apps? And are they not the ones to be protected from Corona?

Canan Marasligil

Another point I had about all this, this data collection part, we didn't talk at all about privacy, and surveillance, you know, what are the risks, you know, of having our data used like this? And is it anonymous? Do we have a say in how that data is used? Could you tell us a bit more about that aspect of that data usage?

Aurore Paligot

[In the context of Smart Cities projects] Well, I would say that to have a say in that it can be done at the level of the city maybe. So you're really, anytime you enter a station or a public place, and there are cameras, then you really don't have many liberty to move out from that camera. But these are the questions that Tada are asking, I really recommend you to go and see their Manifesto. They wrote a manifesto about, yes, how… I think, six important questions that we can ask to be sure that when you're building a project with data, you leave the possibility for the citizens to, for example, opt out of the technology. And that's why we are experimenting. There are places like the Marineterrein to experiment on. We can test the data, we can test the technology, and we can also not agree with it. Yeah, there's no easy option. I think that it's really important to act at the level of the municipalities. I think that you can do many things at that level, because then you can have consultancy with the people. And then you can decide whether or not you want this technology for your city, for example.And also you can ask, between different technologies, which is the one that you prefer, like, for example, are your data going to the cloud somewhere? Or are they stored locally and protected? And what guarantees do you have? So the technical part of it, the hardware part of it is also really important, even though I'm not an expert.

But regarding Social Media, I would say that everything that you put there is publicly… is more or less, it is public. So we all know about the Cambridge Analytica affair. For example, I'm on Instagram, and recently I've been receiving targeted ads to freeze my ovocytes. So you know, this kind of thing happens and how... Yeah, I've read an article recently from

the Algorithm Watch about how Facebook is using discrimination based on the pictures on how to target the people in the adds without the consent of the advertisers themselves. So of course, we have to be careful. Well, from my perspective, I agree to be on this platform. You can always opt to not be on those platforms. And whenever I'm doing a research myself, that is using this data, I make sure that I do not include anything that could, that is unnecessary for my project. For example, I want to know about the group or I want to know about a hashtag that is frequently used or about the habits or about a feeling, a sensation. But I don't know, I don't want to know that you, Canan, you were that day on the Marineterrein, even though I can do it. I can store your pictures very easily. But it's part of my, you know, my procedure. It is data that I want to fully anonymize. I think that they're fun to play with. But it would not be fair for me as a researcher to link them back to their creators.

Laura M. Pana

I just wanted to make a reference to a very interesting article which backs to 2016. And that we are going to link in the show in the show notes. It's from the Wired Magazine, where Barack Obama was invited to be guest editor. And there was a very interesting discussion between him and the director of MIT Media Lab, on artificial intelligence. And even on the pandemics it's very interesting, because you read it in 2020. And you read that Obama in 2016 was very worried about, was much more worried about the pandemic than being invaded with tanks and army. But one interesting thing that, for example, that the director of MIT Media Lab said, regarding who builds artificial intelligence, he said that predominantly is the white male who are doing that, who build the core computer science around artificial intelligence and the perception. So he was saying this in 2016, but I'm not sure much has changed since then. And there is this belief that the artificial intelligence will help us solve all the problems, including social and political problems. And this is very problematic, because they will not know you, if you leave the machine to solve the societal and political problems. It's not. It's not okay. But it's a very interesting because he also raises the question of how do we build societal values into artificial intelligence. And back then he was opening up these questions. They didn't have answers to that either back then. And it's something that needs to be more and more in conversations, what kind of values do we also create when we create these technologies, because we will need to balance these moral decisions. Here, they were giving the example of the electric car where for example, you are in the situation where you need to make a move and avoid killing a pedestrian. But by doing that, you hit a wall and you die. And the moral question here for the machine that will have to make what decision it will take. So it's in this kind of way that he speaks about "What kind of values do we bring?" It's very interesting, I will link it in the show notes for everyone to read.

Aurore Paligot

I think it's really, really interesting, especially in a society where we track data of workers. We were talking about Academia, of course, you know that academic success is measured through publications, number of publications. But if you're a worker in a warehouse, then maybe your productivity is measured to how fast you are to pick up some products on the food aisles. And in both cases, in any of those job, is it how we want to measure what is your worth in your job, and if you're doing a qualitative job, and so I think it's important as data analysts to remember

that we are the one generating the data that are measured, and how do we want to measure it and why?

Canan Marasligil

We didn't answer all the questions, of course, because as we can see, this topic is a process again, we're going back to process everything is process. We are all in the learning process, of course. And we are all doing, trying to work with data in different ways in all different sectors as well. I'm very happy to see all the work you do Aurore and all these examples you gave, I think were very inspiring for all of us. Thank you very much for joining us today. It's been a blast really to listen to you. Aurore before you leave us. Can you tell our listeners where they can find you? Where can they follow your work? You are on social media, please tell everyone where to follow you.

Aurore Paligot

I'm super open to collaborations at the moment about those projects gathering, looking at social media data. There is this corona project and I'm also interested in looking at multilingualism on Instagram and other social platforms, so something that is more linguistic. But of course, if you're curious, if you're motivated, if you have ideas, let's share. And well thanks for having me, it was a pleasure for me too. You can find me on my LinkedIn, because I'm checking it quite often. I'm on Twitter, sometimes. But I'm also now building my own website, so you can find it. It's called aurore.rbind.io. Rbind is the dedicated domain for the R users. So this is where I'm hosting it. It's under process. But I would like to share more about my ongoing projects and ideas. So, let's get in touch.

Canan Marasligil

It will all be in the show notes. We will put all the links, the articles, we talked about the social media profiles, everything will be in our in our show notes.

Laura M. Pana

We are happy to have your questions, your feedback. And yeah, do follow us and if you like what you hear our discussions with our guests, please do write us a feedback or a review. It helps us make the podcast more and more visible and more reachable and accessible.

Canan Marasligil

It gives us the feeling that we're not speaking to the void.

Laura M. Pana

We don't speak into the void Canan.

Transcribed by https://otter.ai and minimally edited by Aurore Paligot