

Podcast episode transcript: Juggy Jagannathan and Thomas Polzin

Juggy Jagannathan: Welcome to the Inside Angle podcast from 3M Health Information Systems. This is your host, Juggy Jagannathan. My guest today is Thomas Polzin, director of natural language processing. He's a brilliant researcher, shepherding an army of NLP practitioners, including me, to research, develop, and deploy analytical solutions for 3M. He got his PhD in linguistics from CMU in 1999, and joined as one of the first employees on M*Modal, which is now part of 3M Health Information Systems. Thank you for joining us, Thomas.

Thomas Polzin: Well, thanks, Juggy. And thanks for those nice words.

Juggy Jagannathan: I want to start it by asking you a simple question. How did you land up in linguistics?

Thomas Polzin: Good question. After going to high school, I was kind of a little bit lost and didn't really know what to do, so I started Spanish and German literature. And then, one of the requirements was actually to take a linguistics class, so I took a linguistics class, and I just left it. This was University of Bielefeld, a small but quite modern university in the northern part of Germany, then I switched to linguistics, which was really easy to do. It took me half an hour to switch these things around, and then I took classes in linguistics, and then one offering was computational linguistics, and it was completely new to me using computers at all.

And then applying it to natural language processing or natural language understanding actually was the coolest thing I thought. Doing this during the day and coming home and you achieved something, you wrote a little program, which did some kind of analysis was, I thought the coolest thing ever. It was very satisfying, it was kind of like Lego for grownups, so I loved it. So I got my master's in computational linguistics and in biology because I was even then, interested in neural networks. And then, I went to Carnegie Mellon University in Pittsburgh to kind of get my PhD and work more on machine learning.

Juggy Jagannathan: That's cool. I saw your PhD was on detecting verbal and non-verbal cues. That sounds like a fascinating subject.

Thomas Polzin: Yeah. So yeah, it was an interesting topic, but it was really stressful, to be honest. So I was trying to use prosody and verbal cues to detect emotions in speech. So what I did, I was transcribing Hollywood movies for emotions, so when there's someone was angry or sad or happy, and then I would use machine learning to try to detect the perspective emotion automatically. And it kind of worked. I mean, the data was of course, kind of funny, it was not enough data. Of course, I was running out of time, but it worked to some extent.

Juggy Jagannathan: That sounds like a lot of fun, playing all those movies.

Thomas Polzin: It was. Yeah, I still remember the movies, actually. One was with, and I shouldn't go there. But anyway, I still remember all the movies because I kind of knew them very well by then.

Juggy Jagannathan: So you were one of the first employees of M*Modal, right? Even before they were M*Modal, and here you are working on and NLU for the past 25 years, whatever.

Thomas Polzin: Yeah, so while I was at Carnegie Mellon University, I went into Detlef Koll, who's now the vice president of IT here at 3M, and Michael Finke, the founders of interactive system, that's how we call it, or ourselves initially. And I think that was one of the wisest decisions I ever made because I mean, working with these smart people was just great. And it was an exciting story. I mean, lots of changes. The suspect of the company, bankruptcies, buy-outs, IPOs, you are part of history and now we ended up at 3M, which I'm really grateful for. For many reasons, most of them, we are probably going to discuss later on. And now, I'm actually leading the Natural Language Understanding teams here at 3M. We have teams here in the States and in Europe around 100 engineers ranging from software engineers, software architects, machine learning, PhDs, so quite a lot of experience, I have to say. And it's fun working with these people and we have lots of interesting problems, you know that, yeah.

Juggy Jagannathan: That's fascinating. To actually manage this large group of people to act, solve. I mean, I know NLU and 3M, we have lots of different products and internationally, but one overriding team I keep thinking of is our customer focus. We are very focused on our customer and we want to satisfy them with all this AI that is coming out of our environment. We have things which will provide assistance to physicians, we have things which will automatically code. How do we satisfy customers with all these AI solutions that we are providing to them?

Thomas Polzin: Yeah, I mean, first of all, you're right. And in some ways that's how we grew up. If you are a small company, I mean, the customer comes first, because he's basically paying a salary. So that's kind of very ingrained in how we do things. If customer comes first, how can we serve the needs? I mean, lots of things you can do, I mean, one thing is actually working very closely with the customer because sometimes it's not quite clear what the customer wants. So it should be not a waterfall, it should be an interactive process to develop a product. And once the product is out there, this should continue.

Juggy Jagannathan: I know, we have these very close relationships with big customers like Geisinger, Mayo. It is part of our customer obsession like you were just saying, that we work with them intimately to craft solutions to help them, right?

Thomas Polzin: Yeah, yeah. That's true. So there's this customer care aspect to taking care of the customer, but then there's also this technology aspect to it. I mean, in order to satisfy a customer, you have to be agile. You have to adjust this respect to the features you want to push out so it shouldn't take months. It should take a decently size feature to push to a customer should take maybe a month, and then the customer should see it. Other smaller changes should be more or less immediately, like a small change in a rule should be pushed to the production data centers more or less immediately, so that's what we are working on.

And there are lots of things you can do. I mean, one thing of course, which is really helpful is actually having the product in the cloud, because if it's in the cloud, then you can just push that change into the cloud and maybe even your build processes part of the cloud or even your source repository sit in the cloud, and then your build process in the cloud, your artifact is in the cloud and the product is in the cloud. Once you have this, you are basically half way there. It's essential, but again, it's always really important to have these communication channels open between engineering product team and the customer. I think that's really, really important.

Juggy Jagannathan: That's fascinating. The scale at which we do it, always amazes me. I mean, 3,000 hospitals, 4,000 hospitals, hundred thousands of physicians.

Thomas Polzin: I mean, yes, you're right. The scaling is in some ways daunting, but there's also a lot of synergy because the product is in the cloud. And we come to that maybe later on, we get feedback from the customers and we change basically one product so it's not that we change something for customer A. If we change something, typically it's for all customers because every customer would benefit from NLU change or some content change. So basically everyone wins by having the model or the product in the cloud. So it's a mutual benefit to have this actually in the cloud as opposed to have things individually on-premise. And this way, you can scale. If you are in the cloud, you can scale because in some ways, it doesn't really matter if you have 3,000 customers or 6,000 customers, but if you have 3,000 on-premise installation, there's the 6,000, that's something else.

Juggy Jagannathan: That's a prescription for disaster. I know that M*Modal always started out in the cloud. I think, they were one of the pioneers in this space. So it's been a long history for working with the cloud.

Thomas Polzin: Yeah, we started with speech recognition. Speech recognition, we put in the cloud. And I think they have, that's initially, I worked with that team, that's where I got some of the experience. And they have a lot of experience because they are doing it for quite some time. I mean, they have in speaker dependent model training in the cloud for, I don't know, hundreds of thousands of physicians, here in the States, but also in Europe and so that's really impressive. So the training actually is automatic. So the profile, the speaker dependent profile is updated automatically so they were really able to scale. So in some ways, we are doing a little bit of catching up here.

Juggy Jagannathan: We were doing cloud computing even before the term cloud computing existed. It was always in big data centers.

Thomas Polzin: I think you're right.

Juggy Jagannathan: I was just thinking that it's not just the complexity involved in scaling and cloud computing, we also have extra dimensional complexity in terms of the international markets we serve. And I just want to pick your brains on how do you deal with this? I know you pull your hair out on this front, but just to get your perspective.

Thomas Polzin: Yeah, I mean, so it's a good question, again. How do you deal with it? I mean, it's not the case that you can have what I just described, you have one cloud instance and all the customer from all the countries connect to this one instance. We have multiple cloud instances here in the States, in Canada, Australia, in the UAE, and in Europe.

Juggy Jagannathan: Doesn't Canada have a... I mean, a lot of these countries have regulations which says they cannot send PHI content outside their country or something, right?

Thomas Polzin: Right, I mean, exactly. That's why I was going there. So because of this regulations, because of these rules, the data cannot leave the country. So this is very pronounced in Europe. And within Europe, I think, Germany is probably the toughest one. So getting things into a European or into a German cloud, even that step is a challenge because people are really wary about their health information. They take privacy very, very serious. So it's a hurdle, but it's a hurdle we somehow have to solve. I mean, speaking for Germany right now and machine learning models, of course, you can train on-premise with the data from one customer. I mean, we can do that, we prove that it works, but there's a drawback. And the drawback is the amount of data. If you train a model just on the data from one hospital, you typically don't have enough data to train a robust, reliable model, in particular, when it comes to very infrequent diseases or procedures.

So they are actually, you would profit if you were able to aggregate the data in the cloud somehow, using maybe de-identification or some other privacy tools so that the GDPR requirements in Europe are satisfied. But that we still have this mutual benefit of aggregating data in the cloud and pushing the benefits back to the customer. So I think, there's some work we have to do with our customers together, convincing them that our technologies are safe. I mean, we have, there's this paradigm of federated learning, and then we use de-ID tools. So we have technology which really makes it safe, but there is some convincing to do, but there's a really strong argument or there are really strong arguments to do it. I mentioned the mutual benefits by aggregating the data and training a common model, which is then much more robust than the little models would be on-premise. And also the agility with which we could update compared to on-premise installation.

Juggy Jagannathan: The federated learning is indeed a fascinating concept. And I know we go the extra mile to make sure that the PHI is protected, not just in Germany, but here as well. I mean, we take extraordinary measures making sure that PHI is preserved and it's not accessible to others and...

Thomas Polzin: Yeah, yeah. I agree, I mean, it's important and we take this extremely serious. Even if we de-ID, de-identify patient information, we still treat it as a PHI. So even de-IDed documents will always be in our safe zones where they're really safe.

Juggy Jagannathan: We only allow access to researchers on a need to know or need to use basis. And they all have to go through all kinds of HIPAA training and other trainings. If you are working with European data, I guess, we have the GDPR training.

Thomas Polzin: Every year.

Juggy Jagannathan: We take it seriously. I can attest to all the different trainings I've gone through. No, PHI is important so we pay attention and we... What can you say about all the deep learning research? I mean, we've been doing a lot of deep learning stuff, particularly the last half a dozen years.

Thomas Polzin: It's an astounding technology and I don't think we have seen the effects right now. I mean, this technology will impact everything we do. And by us or by us, I mean, everyone on this earth, not just people working with computers or information, science or something, it will come to everyone. And you're right, I mean, we played with this technology early on when it was still kind of in its infant state. We have now products actually, which use it in production and the speed with which you can develop models and produce products, it's astounding. And in particular, for other languages in English.

For English, you have a lot of data. But if you go to Europe, you have many, many languages. Some of them are not spoken by many, many people, but you would still want to support the doctors in the hospital workflows with improving the documentation, saving their time. And to build a rule-based system, for example, Dutch, it's extremely expensive. I mean, you have to find the subject-matter experts, it takes years. If you just have the training data and you are able to train some transformer model or convolution networks, you can do that in a couple of months. I mean, you're not done, but I mean, you have something which proves that this is a valid way to pursue and this is technology which will make your product work.

And then lots of things you have to do in order to operationalize this and quality control and all these things, explainability is another thing, which is not always coming naturally to this transformer models or convolution networks or whatever. And typically, our customers, they want to know why we did something. Why would we alert this doctor? So they want to know why we did it and so that's something the models have to be able to do. So again, I mean, we just scratched the surface and we are pushing this deep neural networks into our products and it's quite amazing.

Juggy Jagannathan: The one fascinating thing that comes to my mind is we are always in the mode of assisting somebody. So we are assisting a code, we are assisting a clinical documentation specialist, or we are assisting a physician. So our models are engineered to help different people do their task better. So in that sense, explainability is, at some level, is incorporated in the way we actually work with the actual solution. So no, explainability is quite important, but when you're engineering human in the loop, it helps us get these products sooner into the market because it's helping them do their job. And so coming back to the customer, which is the human in the loop, so I was just trying to get us going to say...

Thomas Polzin: I was wondering where you were going. I think, I know what you were trying, about where you were going. So the human in the loop, training data, gold standards, quality, I mean, quality agility. I think, we talked about it or we touched it briefly at the beginning when we talked about our customers, how important it is to work with them. So obviously, you have to make sure that things work. And in order to these things become complex. I mean, clinical documentation, improvements, those are quite complex algorithms we use in order to issue an alert to the doctor.

So we develop gold standards to make sure that by some free thing, we don't affect the quality of our products. Some little cold chains should not affect the quality of the product. But sometimes not obvious, but we have to take care of it to make sure this does never happen. So we do develop gold standards while we actually develop, for example, a CDI, clinical documentation improvement rule, and then we go back to the customer and we create gold standards with their data. But it's an expensive thing to do, it's kind of slow. And discussing that, sometimes becomes stale because I mean, the customer was on new diseases, COVID come. So you don't test for COVID in your unit test on your gold standards so you have to make sure they are up-to-date. But even more importantly is, I mean, the scale, but again, more important or a much better way to do it is actually to look at the customer.

How is the customer using the product? Is there something which allows us to use the customer, while he is using our product to implement a feedback loop and observe, I mean, using telemetry and observe whether the things we do are correct. I mean, for computer assisted coding, it's straightforward. I mean, you have a model, it predicts ICD-10 codes, CPT® codes, you propose it to the coder, the coder accepts it or not, and maybe introduces some other codes, but you get basically feedback with respect to what you proposed, so in some ways that's the exception in clinical workflows.

In particular, if you want to deal with the doctor, the doctor doesn't have time to give you feedback. I mean, and they shouldn't, it should not be the doctor's task. "Hey, you did a good job like this, ICD-10 code or this ICD, CDI alert was perfect, thanks so much." He shouldn't do that, he shouldn't even be aware that he should do it. So we have to find feedback loops, which are indirect. For example, "Does this CDI alert go away after a second or two?" Which would indicate that maybe the doctor said something in order to make it away, things like that, or "How much time did he spend looking at his window?" Or something.

It doesn't have to be direct, but it has to be massive because it'll be noisy. But you need, and therefore you need a lot of data. And if you're able to get these feedback loops from hundreds of thousands of doctors using your product, then noisy data is really, really good data because you have so much of it and you can use these feedback loops to make sure that your product is working and you can even use these feedback loops to make sure it's actually improving. We might use machine learning to improve certain things. So data, again, is important, is key.

Juggy Jagannathan: I know we spend a lot of energy looking at these workflows. Our user experience designers spend a lot of energy trying to work with all of our customers to not only design these feedback loops, but also to actually help us design it in the first place, design the solutions.

Thomas Polzin: Yeah.

Juggy Jagannathan: Do you have any closing thoughts on the massive effort that we have underway in terms of all the work we are doing, any thoughts?

Thomas Polzin: Yeah, I mean, again, I think, the impact of these deep neural networks, it's hard to predict, but I think it will be immense in the health care settings, in the health care documentation processes and everywhere else. So it's hard to, and I don't want to predict

anything here, but the impact will be substantial. In general, I'm not a big fan of throwing technology at problems, I'm I guess, old fashioned. But I think, what we did in the last 20 years, our products, I mean, developing technology to help the doctor with her documentation. Actually, I think, it does work, it makes a difference. I mean, we hear feedback that it actually helps.

I mean, the typical example would be the incidental findings. I mean, the patient comes in because he has trouble breathing, the primary care physician thinks it might be pneumonia, he sends him to the radiologist, the radiologist rules out pneumonia, but she finds some nodules in the chest X-ray, nothing to be concerned about, but there should be a follow-up in five years. So please follow up in five years, she writes this in the document. It's a string of characters in the document. If nobody acts on it, there's unfortunately a chance that the patient might die from lung cancer, so using NLU to find these...

Juggy Jagannathan: I know we had a number of success stories in flagging these incidental findings across different health systems.

Thomas Polzin: Yeah, at Geisinger, we were actually able to prove, I mean proof, we got feedback from, or emails from Geisinger telling us, "This year, you saved X numbers of patients bleeding to death because we found this aneurysm and we acted upon it." So I mean, you can be cynical about technology, but I think, you've proved in the last 20 years that actually it does help. I mean, these are dramatic examples we are talking about. Even saving that doctor 30 minutes every day is something worth working for. So I'm not cynical about my work, I take it very serious. I like what I do, it's very rewarding in this respect so that's good closing, that's good closing.

Juggy Jagannathan: I wanted to say that, that's perfect closing. And I just wanted to make one comment which says that, "Physicians are very happy to recover time, they increase their pajama times." So that's a quote from actual physicians due to our technology. Thank you, Thomas, it was great chatting with you today and-

Thomas Polzin: Same here, as always.

Juggy Jagannathan: Thank you.