

# Podcast episode transcript: Juggy Jagannathan and Detlef Koll

**Juggy Jagannathan:** Welcome to the Inside Angle podcast from 3M Health Information Systems. This is your host, Juggy Jagannathan. My guest today is Detlef Koll, our global vice president of research and development. He's a brilliant visionary who has shepherded speed solutions from CMU buildings to a company founder, M\*Modal, to now leading the development and deployment of health care solution here at 3M at the cutting edge of technology. Thank you, Detlef, for finding the time to join this podcast.

**Detlef Koll:** Thank you, Juggy, for the nice words.

**Juggy Jagannathan:** Oh, you're welcome. I just wanted to kickstart the whole discussion going all the way back to your CMU days and share some of your experiences over there.

**Detlef Koll:** Yeah. So back in the mid-nineties, I worked at a research lab at University of Karlsruhe and Carnegie Mellon University. It was an interesting time period because it was right on the edge of when speech recognition systems started to become useful. So, if you think back mid-nineties, the commercial systems at the time would require you to pause between words to be understood. So it wasn't continuous speech recognition. It was isolated words, speech recognition. At the time in universities, triggered by investments from the government through DARPA and missed grants, conversational speech recognition but more natural speaking styles were becoming well while not yet mainstream, they were becoming reality in research settings.

**Juggy Jagannathan:** Yeah. I remember they used to have these discrete word, speaker dependent systems or something like that.

**Detlef Koll:** Exactly.

**Juggy Jagannathan:** Now you're talking about continuous words, speaker independent solutions.

**Detlef Koll:** Yeah. Actually, the government was interested in wiretapping application at the time. Usually, you can't ask terrorists to enroll and then enunciate well in order to be understood. So basically the tasks that they gave a couple of research labs were, again, no enrollment, through telephone lines, meaning somewhat reduced quality speech recognition and fully conversational. There were multiple tasks. We used to work on something that was called switchboard. At the time, all the big name research labs participated in the same research grants. Thus, they had to submit to the same kind of evaluations where the government, once a year, forced everybody to disclose the accuracy rate of recognition. So BBN, Dragon, SRI, University of Cambridge and the UK, all the big research labs were part of

it. Well, we happened to, well, have pretty good technology at the time. We won this evaluation twice in a row. BBN might dispute this, but I still think we won.

**Juggy Jagannathan:** Is that the one which led to the founding of M\*Modal or there was also Hauspie somewhere in between.

**Detlef Koll:** Yeah. So it led to us founding the first company. It was called Interactive Systems Lab, same founders for M\*Modal. Our advisor at Carnegie Mellon University, Alexander Waibel, we was involved and co-founder of this company. At this time, we had the brilliant idea to just commercialize speech technology, which means we built the technology and we will find somebody who uses it. But basically we just wanted to license technology and actually we got off to a good-

**Juggy Jagannathan:** No, no. How did you land up in the clinical space? Speech technology could be used for anything.

**Detlef Koll:** It was planned, actually not. We started the first company called Interactive Systems. Funding for this came from Japan at the time. In Japan, we started an interactive voice recognition interest group. So we signed up lots of companies like Toyota, Hitachi, Fujitsu, NTT, all the big Japanese companies. They would pay us \$10,000 entry fee and then we would engage them on building whatever applications they wanted to build with speech. It was reasonably successful. We signed up close to 200 people through our partners in Japan and yet we didn't get a single commercial application into the market through this. So what we learned in those early days is that if you have good speech technology, it doesn't immediately translate into a good product. Right. At the time to be honest, I mean, speech recognition it demoed well, but it still had its issues in real world use.

**Juggy Jagannathan:** Well, you're really the pioneers, right? We are talking about 20 years ago or even more, 30 years ago, 25 years ago. Yeah.

**Detlef Koll:** So we had the good luck that we actually could sell the first company because by all rights, we should have gone out of business. But before we got there, we managed to sell the company to a big Belgium company, Lernout, which conveniently went bankrupt shortly thereafter, so they had to sell the pieces in the bankruptcy auction a year later. So we had the opportunity to reboot Interactive Systems and that actually is in M\*Modal. So we founded a second company, bought out the assets out of bankruptcy from Lernout. We did actually learn some lessons in the first time around. The lessons was don't try to sell technology. Sell solutions to problems. We then took a year to identify a problem. We looked into three areas. One was voice portals. I mean, imagine that back in the early 2000s people believed that you would use your phone in order to access the internet. Well, at that time though, people believed that you would call up a number and say

what movie is playing in that cinema. So actually people were thinking that you would have voice portals where you access information from the internet to speech.

**Juggy Jagannathan:** It seems like Alexa before it's time.

**Detlef Koll:** Yeah, it was Alexa, but it was also a lot more primitive because it was not ambient. You would have to dial a number in order to then access the internet. Of course, the iPhone and before that NTT, what they called it N-mode, I-mode. There were different solutions, but again, early 2000s, this was a big boom. We worked on it. Luckily, we didn't pick this area because pretty much all companies who did went out of business. We looked into media mining, so indexing voice data, however it's collected. That I think would've been a good area to get into, but at the time we decided that the speech technology is just not good enough. I think there was a benefit of hindsight and probably we were wrong. But anyways, the third area that we looked into was clinical documentation. After a year, we picked this and since then focused exclusively on clinical documentation.

**Juggy Jagannathan:** That's unbelievable. That was quite fortuitous that you picked that area. I think it's definitely a winner.

**Detlef Koll:** Yeah. So to what made it a very interesting area is that dictation transcription used to be the primary mode of documentation for physicians. So after a patient encounter, physician would dictate the clinical note. This dictation would then go to a transcriptionist and would type it and send it back as a draft to the physician who would review it and sign it. Right. The beauty of this application is that there were transcription companies who were paid to perform this work, obviously. If you would use speech recognition in order to create draft documents, making the transcriptionist more productive, the transcription companies would have the benefit of reducing their production cost. We would get the benefit of, a), getting paid for this and, b), that we are getting paid for collecting data because the transcriptionist still would've to fix the speech recognition errors in the draft, right. We call it closing the loop, where we provide output and then collect user feedback on the output we produced.

So starting in 2004, I believe, we offered this as a service. Once we had the service deployed, we would collect data. As we collected the data, the service became better and better over time. So it's a beautiful model.

**Juggy Jagannathan:** And you targeted all the transcription service providers, right?

**Detlef Koll:** Yeah, indeed. Well, there were two companies in the space trying to do this. One provided this as an integrated platform, and our approach was as an OEM technology provider to transcription companies. I think pretty much every sizeable transcription organization in the U.S., except for our competition, signed up to our offerings. So it was a good run.

**Juggy Jagannathan:** Yeah. I noticed you didn't mention MedQuist. MedQuist was your big competition at that time.

**Detlef Koll:** You're right. At the end, I view MedQuist as a good friend right now because we ended up merging, but indeed in the early days we were competing as.

**Juggy Jagannathan:** It was an interesting time. I think everybody was... You were targeting all the service providers and MedQuist was just gobbling up transcription service providers. It was kind of an interesting period, I think. It's all through early 2010 or '11 when we merged with MedQuist I guess. Can you talk a little bit about what happens after the merger with MedQuist? I mean, we're pretty much the largest technology and the service provider standing.

**Detlef Koll:** Yeah, maybe the second interesting lesson... I mean, the first lesson we took from running a company was technology doesn't sell. The solution sell. The second lesson was don't disrupt the business that you depend on. That we learned over the next 10 years. Again, we introduced technology successful in the transcription market. This kicked off a large wave of consultation and the companies who were quicker and better in adopting technology had a significant advantage. They ended up buying the smaller transcription companies leading to a wave of consolidation. Then this collided at the late 2000, like the first decade of the 2000s, speech recognition became good enough to be viable as a documentation modality for physicians. It still was a little bit painful, but now more and more physicians started to adopt front and speech recognition.

Then when the meaningful use in this initiative, the big stimulus package ARA well hit the market after the 2008, 2009, it was, yes. So this incented hospitals to deploy electronic medical record systems and part of the cost savings of an electronic and medical record system at the time was touted as the ability for physician to immediately elected to self-enter clinical documents into the computer, thus saving transcription cost. So there was a lot of pressure on the transcription space in this time. Again, so we relied on the transcription industry for most of our revenue. It was a rapidly consolidating market. The market itself became under pressure from hospitals increasingly trying to force their physicians to self-enter documentation. So the perfectly disrupted market well encouraged us to look for partners to de-risk this.

So we saw the writing on the wall, that clinical documentation is going to be more and more going to fall on the physicians and dictation transcription is going to reduce over time. So if this transition happens, we had two options, either get out of the market and figure out another business to sell speech technology in, or to double down in the market and be a disruptor, right. So we are being disrupted, but I mean, we could also, if we align in the right way with the big players in the market, be the disruptor and be the entity that helps provider organizations to make this transition from transcription to self-documentation.

So we chose the second, and we had the opportunity to pitch the idea of using technology and in an increased way to the biggest transcription vendor at the time, this was MedQuist and MedQuist ended up buying a model, then rebranding as a model and trying to be the entity who helps provider organizations find the ideal mix between dictation or documentation modalities. Whoever wants to dictate and transcribe, we will support this. Whoever wants to use front-end speech recognition to information directly in the EMR systems, we would support this as well.

**Juggy Jagannathan:** I mean, I think this whole notion of supporting the enterprise needs and whatever modality that they wanted to create clinical documentation, I think this was first

proposed and executed by us, right, in this model. That was brilliant because I think it was a cannibalizing model. We were cannibalizing the transcription service, but we got all of them in speech. 10 years, hence, I think that model has broken particularly no physician wants to use anything other than speech.

**Detlef Koll:** Yeah. If you look in speech recognition today, and we use a speech recognizer, even for somebody like me who is not speaking proper English can use it, but you didn't use to be like this. So even in the 2011, '12, '13 time range, physicians needed to be motivated to make it work. Even then, for some people, it would just not work as well as it needed to in order to be a real productive tool. So if you would sell a solution to a hospital with 100 physicians, probably 10, 20, 30 would absolutely love it, maybe 60 would tolerate it and 10 would absolutely hate it. It's a difficult to deploy product if you go live and you have 10 people for which it doesn't work well enough for them to enjoy the experience. If you just sell speech recognition as an isolated product, this is difficult. If you sell it as part of a package and could say, well, for those who doesn't work well enough and who prefer transcription, you continues this as part of the same package. So it makes it way easier for hospitals to deploy something like this.

**Juggy Jagannathan:** Right. Of course, we all know where the EMRs solutions are saying, let the physician be secretary, how it turned out for physicians. I mean, created a wave of burnout syndromes and they hate it. Physicians hate using the EHRs.

**Detlef Koll:** Yeah, and it's an interesting story because how EMRs used to be positioned in the market initially. I mean, the early EMRs systems, they fully expected physicians to self-enter information and they expected information to be nicely structured by physicians so that it can be stored in a database. I think we were always proponents of the narrative because a narrative documentation is, from the human user experience perspective, close to optimal. Physicians can say whatever they want and whatever words they want. Whereas if you have a structured input modality where you only have a check box, allowing you to say patient has X or doesn't have X, then patient might have X or I cannot rule out X or patient might be at risk of X, cannot be documented.

So basically the documentation modality impacts what information is collected. We always use this as a big deficiency of the structured reporting modes. If you collect data a certain way, then you distort the data that you collect. Narrative was free of this bias. Narrative is perfectly interoperable. You can send a clinical note to another physician and he would understand it. So all our tasks were just to address the issues that we have as narrative. The issues with narrative was if you use dictation transcription, it's expensive to produce. If you use physicians entering it in a computer, it's even more expensive to use because they have to type it. If you use front-end speech recognition, it reduces the cost somewhat. But in terms of physician times, it's still much slower than dictation transcription. So you have to address the cost of creating narrative as one item.

You have to create address the immediacy of how quickly information is available. So if you use dictation transcription, it uses very little physician time, but the result is only available the next day. So if you need a ER note or anything where you have a transition of care with a patient, well, for provision of care, you need the result immediately. Dictation transcription is

just not ideal. Finally, if only you have narrative, it's really difficult to deal with the information embedded in narrative. So if you want to drive care automation or care pathways from discrete information, it's very easy. I mean, if you check certain check box in the EMR, you can use this information to then influence care. If you dictate the same information in a free form narrative, it's much harder to act to. That includes-

**Juggy Jagannathan:** That's when all our efforts to build and use this actually understanding systems come into play, I guess, right?

**Detlef Koll:** Exactly. When we talked about narrative is the best modality to document care, and we don't mean this dogmatically, I mean, there are, of course, settings where you want structured information. But for the majority of use cases, we believe narrative is ideal, so we have to address the shortcomings of narrative with this cost, immediate of availability and accessibility of the information in the narrative. That's why we said from the very early days, we have to pair understanding this documentation or tagline used to be always understanding.

**Juggy Jagannathan:** Well now it's create time to care, which is just as good or better actually. Let's do a slight pivot. I mean, the past decade has been very interesting in terms of the evolution of technology. You fairly early on took us in the direction of scribing, almost like half a dozen years ago. What motivated you to go in the direction? What do you think is the technology evolution, I guess, in the past five, six years?

**Detlef Koll:** Yeah. So first scribing business is about paying intelligent human being without clinical knowledge to listen into the physician patient dialogue in order to have this scribe prepare a draft note for the physician. Now, the beauty of this is that the physician doesn't have to document anymore as a separate work step. Even if you have to dictate for a transcriptionist, it is a separate work step after the patient leaves the room. I mean, if you could just have an intelligent human being listen into the physician-patient dialogue and create a draft document from this, you can even save more physician time. This is a model that was proven, I want to say, starting in the 2010 and following, and it proved valuable.

It is a very expensive model because you pay a human being to follow a physician around and basically be present whenever the physician sees a patient. So those services can easily cost between 30 and \$50,000 per physician per year, which means it's somewhat limited in how broadly you can deploy it and be very intrigued by this model very early on. But we just didn't see a way to introduce technology to solve this problem or to help this problem. Now, also in the early 2010s, actually 2012, Jeff Hinton published his seminal papers on deep learning for speech recognition and image recognition. So that's when the current AI boom started, in 2012. He applied deep neural networks to speech recognition and proved superior accuracy. At the time, we didn't quite realize how big of an impact it would have on the industry. We knew that we had to adopt it for our own front-end speech recognition products. Within two years, we had a commercial product on the market using deep learning acoustic models.

But then what happened in the industry is it happened when Alexa was released. Slowly, the market realized that this was a technology revolution. It was a breakthrough that enabled whole classes of new applications. So things that were completely unthinkable, for example, recording human to human conversation in a fast speaking mode. So where we have adjusted

microphone somewhere in the room and being able to transcribe it, that's an accuracy level that allows a system to then automatically summarize it in the clinical report. This was science fiction even 2012, '13, '14.

But as we continued playing with the technology, we came to the conviction that eventually this technology is good enough to actually automate what a scribe can do. Actually in 2016, we pitched this as a vision to our board of directors and convinced them that this is the future. We didn't tell them a timeline. I mean, I still would be hesitant to commit a fixed timeline. So it was a little bit science fiction in 2016, but it was clear that eventually the market would move in this direction. So that's when we started. It is reality now.

So it was back in 2016 when M\*Modal decided to become a scribing company. So we started this human service business in order to then have an in-house lab to develop the technology and deploy it through our services in the same way that we did it for the transcription, what worked great in transcription where we have technology deployed through a human service and humans fixing the mistakes of the technology. We replicated this in the scribing world.

**Juggy Jagannathan:** This is the same closed loop system that you mentioned earlier.

**Detlef Koll:** Exactly. Yeah. It's fascinating for me to see how technology breaks through and probably not even Jeff Hinton would've realized how whole new application area has become feasible now due to a breakthrough in technology. Alexa is only here because, well, it was a brilliant idea from Amazon, but without the Hinton papers, it wouldn't have been possible. It's not that the idea itself needed to wait for its time to become feasible in order to become to the market. Interestingly, it happened very closely after the Hinton papers as well.

**Juggy Jagannathan:** I remember having a discussion with you in 2016 where you talked about the translation systems that Google had. They had like 50 or 60 people linguists and they fired them all and replaced them with deep learning solutions. There was a lot going on at that time, which I mean was tipping point, I would think, with a lot of these technologies.

**Detlef Koll:** It's still going on. It's still accelerating. I mean, we just released our latest speech recognition models. In the past, 10 percent speech recognition error reduction per year was a very good year. If our research team came out with algorithmic improvements or better ways to train the models to reduce errors by 10 percent, our customers would be happy and me as a development manager would've been happy. Now we just released a new model that is better than the previous iteration. So you're still seeing this year after year, and this revolution is continuing.

**Juggy Jagannathan:** It's totally incredible. Let me touch base with the last chapter, not last chapter, at least the next chapter about 3M acquiring M\*Modal. So how has the transition been? Now that you are the head of all HIS development and research, what is your new perspective on where this technology is going?

**Detlef Koll:** Yeah. So I sold M\*Modal to 3M. So I have to conject a little bit how the motivation of 3M at the time to acquire M\*Modal, but there's the logic of the deals that M\*Modal was a

clinical documentation vendor helping physicians to document. We had a technology we called CAPD, computer assisted physician documentation, that allows us to interact with a author of a note while a note is being created in intelligent way. Which means we would try to understand the content of the note that is being authored, put it in the context of what we know about the patient and interact with the physician in order to make sure that the note is compliant and complete. 3M's biggest deploy base or biggest product was the next step of the revenue cycle, where then the note is used in order to, well, be coded for reimbursement by an insurance company. So the primary use case or one of the most important use cases of a clinical note is to get paid from the insurance company.

So it was a natural combination of combining the documentation step with the billing coding step, because then again, you can now close a different loop. You can close the loop from the coder and to give feedback to the physician how you document and encounter in a way that enables appropriate reimbursement from the insurance company. Then there's a different product between the two, which is called clinical documentation improvement, which is a process of making sure that the note is complete and compliant for billing coding. Again, so there's the benefit. The immediate benefit for 3M was to tie those two applications together.

**Juggy Jagannathan:** Now they also have a population health component, right?

**Detlef Koll:** Yes, exactly. Then there's the same way that you can use CAPD in order to make sure that the billing coding process is supported, you can also use it in order to make the physician aware of population health needs. For example, in order to manage reimbursement in shared risk settings and capitated payment models like HCC, hierarchical care conditions, which require a certain clear practice in order to properly treat patients. In there too, we have intelligently engage with the physician in order to make sure that the note is appropriate for the purpose.

But now there's a big promise or what is exciting right now is we can increasingly automate documentation through those ambient documentation technology by automating the scribes, virtual automated scribes. We can automate the documentation process and you can now take this concept of CAPD and say, we cannot only interact with the physician intelligently while he's documenting a note in the EMR system, but also during the care process, while the patient is still in the room, have the ability to potentially engage with the physician in a respectful passive way. We would never interrupt anybody, but in an appropriate passive way, we can help improve best practices in care. Yes.

Then if you think further, the billing coding process is increasingly automated as well. The same technology that helped us achieve those breakthroughs of the speech recognition, both allows us now to automate more and more of the billing coding process. Now you would have the hope of automating to the degree that if a patient leaves the room, in principle, the note could be done in a draft state and the billing coding could be done in a draft state. So in principle, when the patient leaves the room, you could give them all the reimbursement information or copay everything in the hands or you can spin it forward and say what if the insurance company would engage at the same time? You can automate including the claims' adjudication through the insurance process.

So this is now still science fiction, but all the inefficiencies of the revenue cycle process from documentations through coding, through the reimbursements and the claims through rejection process could in principle be automated. That is something that would be of huge benefit to the U.S. health care system, would take massive amount of inefficiencies out and allows you to think about accountable care plans, or you talked about population health management. The goal of population health management is to treat each patient appropriately according to the best care practices.

Now you have a way to actually implement this, not relying on distributed health insurance payment models where there's not a single actor having a full picture of the patient. So just in today's world, you just have to hope that somehow the care continuum and all your care providers, all your physicians somehow interact in spite of the fact that their care systems are not integrated. With this innovation, you have your hope of actually creating entities who have a full view on a patient's record and really changing the way population health is provided in the U.S.

**Juggy Jagannathan:** Right. You can basically individualize care, a complete view into that care continuum.

**Detlef Koll:** By being able to interact with the physician while care is provided. In principle, the dentist could tell you about your diabetes follow up. I'm not sure that's a great use case, but this is a connection. When we interact with a patient, those are very valuable. Those are golden settings where you can influence a patient in a way that maximizes the health outcome to keep patient healthy before they get sick and seek treatment for a condition. I mean, you want to engage with them before they get sick, right.

**Juggy Jagannathan:** That's a good segue to what I was thinking of next, which is 3M has announced that they're going to spin off the Health care Business Group. You mentioned dentists. We do have a dental group in the health care business group as medical solutions. So what is your view thinking forward now that we know that the entire Health care Business Group is being spun off as his own entity? Of course, it's still a year down the road.

**Detlef Koll:** So I think the Health care Business Group as in 3M is a phenomenal organization. We have great technologies across medical devices, wound care, oral care, food safety. So it's a very broad and powerful collection of technology assets and customer touch points. Now the hope and in spinning this out or my hope in spending this out is that we have a singular focus on health care in the future. It is not that 3M would be defocused, but if you look into a diversified organization like 3M, and we do have, I think, over 50 divisions, building tape for home users and automated systems. It's a very, very broad variety. There's a very broad variety of concerns. If you have the ability to focus on just the drivers in health care, it's still a huge industry. It's still a very diverse industry worldwide, but it's less diverse than what 3M has to manage today. So now I have high hopes for what we can do if you focus further than just solving the inefficiencies of the health care system and providing better care.

**Juggy Jagannathan:** I completely agree. You mentioned customer touch points. It's mind boggling how many customers we have and literally the thousands of hospitals and clinics and

across all the different health care business groups. I think we have a big role to play in health care. Any closing thoughts, Detlef? Thank you for taking the time.

**Detlef Koll:** Closing thoughts, if I would look back to the beginning of my career, I would've never thought that health care administration is a cool area. But now understanding the societal impact that health care administration has, and it's two trillion plus industry in the US alone, it impacts all of our lives. Quality of care, cost of care are real societal issues, well there are problems and there are areas of improvement. Our ability to impact this through technology is phenomenal. If you talk about right now automating documentation creates time to care for physicians. According to some estimate, it saves between two to three hours of physician times a day. If you think about it, and we have 800,000 practicing physicians in the U.S., that would be the equivalent of adding another 150,000 physicians to the U.S. in terms of work time. Accessibility of care, I mean, the impact is huge. As a geek like myself, through technology, we have the ability to impact something of this size and importance. That's a dream job I think. It's a great area to be in.

**Juggy Jagannathan:** I completely concur. Well, thank you that Detlef.