**Industrial AI Podcast Transcript**

**Oct. 11, 2023**

**Guest: Kurt DeMaagd, Sight Machine Co-Founder and Chief AI Officer**

Peter Seeberg: My guest today is Kurt DeMaagd. I hope I pronounced that correctly.

Kurt DeMaagd: That's exactly right.

Peter Seeberg: Yeah, I did good. Kurt is co founder and chief AI officer at Sight Machine, and Kurt and I are going to talk today about democratizing industrial data with generative AI. Hello, Kurt.

Kurt DeMaagd: Good morning. How are you today?

Peter Seeberg: I'm fine. Thank you very much. Please introduce yourself, Kurt, to our Industrial AI Podcast listeners.

Kurt DeMaagd: Yeah, my name is Kurt DeMaagd. As we mentioned, I am the chief AI officer and one of the co founders at Sight Machine. And so we're a company that has been around for for a little over 10 years now, and we've been working with manufacturers around the world on how we can help them extract their data, transform it into useful formats and turn that into business value.

Peter Seeberg: Oh, more than 10 years. Before we get into the topic of democratization of industrial data with generative AI, maybe let's talk a little bit more about your company Sight Machine then. So what does Sight Machine offer since when we just heard 10 years, to whom do you provide and where are you based?

Kurt DeMaagd: Yeah, so we are a manufacturing data platform. Now, of course, there are a lot of folks out there working in manufacturing data these days, but a lot of what we focus on is actually getting down to the plant floor, getting data off of all your different plant floor systems, going sometimes straight to the PLC on machines, but also historians, quality management systems, ERP, individual sensors, whatever it takes to get the data that you need to prove out value. We take that blend it all together and put it into a format, which is easily understandable by not just your data experts, but something that even operators are able to use, understand, and start to integrate into their process.

And so we've been doing this for about. 10 years now, as I said, and we are based officially in San Francisco, California. I'm actually in our Ann Arbor, Michigan office, and we have folks scattered around the globe courtesy of the, the modern era where it's easy to work from anywhere.

Peter Seeberg: Okay. We'll come to that later as well. Okay, then why not tell us a little bit more about your manufacturing data platform that you just introduced to us?

Kurt DeMaagd: Yeah, so we try to get all of the data that is out there because what we've found is over the last 10, 15, 20 years, the history of things like Lean Six Sigma, which have had a good emphasis on data, have mined a lot of the easy value. And the next era of value is going to come from being able to look at the entire system of all of your data so that you can, you know, bring together information for both the quality management system and what's happening directly on the plant floor and from your ERP system, which you need to solve that next level of problems.

And so we want to have all that data blended together. We also want to make sure it's all in real time. You know, we work with a lot of customers where when they have data related questions, they may have to go and talk to the IT department or somebody like that. And then sure, you'll get the data in 3 to 5 weeks . We want to make it so you have effectively real time data as it's being generated, getting it streamed to the users who need it most.

And then I think the third thing that really is important about what we do is we take a perspective, which we call data first, which is, instead of just starting with a single use case or a single problem or a single application, and then you go find the data necessary, grab it, solve that one problem, then you repeat that process where you always "application to the data, back to another application, different data set." It's much more efficient today to get all of your data organized in one easy cleaned up centralized repository first, and then it makes it much faster to build lots and lots of applications on that.

So you're really accelerating your time to value there.

Peter Seeberg: Data first sounds great. That's something that one of the AI celebrities, Andrew Wang, you probably know, and our listeners know he's been driving this approach for not sure how long, but maybe you have been driving this approach for actually a lot longer than that.

Kurt DeMaagd: Yeah, we're big fans of Andrew Wang. So, yeah.

Peter Seeberg: Sounds good. So, why don't you share maybe one or two typical use cases around your manufacturing data platform?

Kurt DeMaagd: So, we go through actually a full spectrum of use cases. In fact, it's kind of funny, one of our most successful customers has been really focused on what I sometimes call arts and charts, which is you don't necessarily have to go all the way over to predictive and prescriptive analytics.

Sometimes it's just more valuable to get your data in the platform where somebody can see and respond to it quickly, where we'll see them, you know, just having the dashboards sitting right next to the end user you know, the operators on the plant floor, able to see when things are starting to go wrong immediately and to take action rather than two days later when it would have come up through the more traditional data gathering approach, and then you wait and talk about it in the morning meeting things like that.

And so we do some sort of arts and charts types of basic visualization all the way over to some of our most advanced analytics. One of my favorite ones was one where we actually were blending together more traditional supply chain management data with plant floor data to solve scheduling problems.

And this was actually for a large dairy processor. And we were able to take the official plans that, you know, get updated in a traditional ERP timeframe of once per night. But often we're not adapting to the reality of raw materials arriving in a much slower fashion or assets on the production line going down.

And so being able to quickly and immediately reschedule in an optimized format so you can either meet or at least get as optimally close as possible to the original plan. And so we can sort of blend the, the old school ERP scheduling with on-the-fly adaptation to create the sort of consistent output.

So we, we sort of do everything from A. I. down to basic visualization.

Peter Seeberg: Okay, we come to the details for that. I understand that you work in different stages in your manufacturing data platform. Can you share those with us?

Kurt DeMaagd: So just to clarify here, when you say different stages, I could interpret that from a technical perspective of the data we process, but also a more use case oriented, which stages of the manufacturing process. Can you clarify?

Peter Seeberg: Yeah, I guess it's like from the connecting to the building, analyzing and just share a little bit of an idea of maybe of how the potential use. I mean, in the end, it's like potential users of your platform listening. How are they, would they be using your platform?

Kurt DeMaagd: So we originally started off with our focus being sort of the core data transformation, saying that, you know, recognizing the value of taking many different data sources in real time, blending those together. So kind of this, I'm gonna call it middle stage of any processing pipeline of what was traditionally called, you know, ETL. Of course, these days, ETL has gotten a whole lot more complicated, but that initial sort of that, that data transformation stage.

But of course, we don't just do data transformation because it doesn't do you any good to just have a wonderful transformation layer if you don't have all the raw data. And so that actually drew us further towards the connect side, where we work through all the connectivity going down to the PLC, talking various protocols there, being able to interface with a large number of quality management systems, MES, ERP, etc.

So building all those connectors to get data from traditional plant floor devices. And of course, you know, we work in an ecosystem today where many customers have already done that basic connectivity and have data up to the cloud. So that means that we work with and partner with many cloud providers so we can extract the data there.

So you know, we we've had to go back to the data sources and make sure that we can collect as many different data sources as possible. possible. And then that gets us to what I mentioned earlier, the the ETL pipeline that we have to get everything blended and transformed and cleaned up in real time.

And then, of course, we have to solve the next problem after that, which is after you have this wonderful data set, what do you do with it? Data is no good unless you create business value with it. And so that then led us into all of the AI machine learning, visualization tools, other workflows just to help manufacturers. And so that does give us the full, end-to-end connectivity data wrangling all the way through to the final end- user applications technology stack.

Peter Seeberg: Sounds great. So when, then, or where does then, you're in charge of AI, where does AI, where does machine learning come in, and maybe you can give us an example of an application use case or machine learning technology example?

Kurt DeMaagd: So we break that up into a couple of different scenarios. So one is actually on the initial ingestion side, where this is, less visible often to many of our customers, but when you have very large volumes of a lot of different types of data, getting that efficiently blended together, either is going to take a lot of manual work, or based on our decade of experience, we have started to develop our own algorithms to help automatically detect and arrange and organize what types of data do you have so you can get into that data pipeline, monitor it and ensure that you have a high quality output. So one of the most important parts that we actually have is not visible to folks where we're just using AI to make sure that you have good, clean data.

Now, what more people will see is, when we're using AI at the application layer in terms of how we use the data then to deliver value to the operators and engineers on the plant floor. And so we have out of the box applications. These could be things like our cookbooks application is a recommendation tool that is looking at real time data on the production plant floor and trying to understand where you have your current machine settings today, how that compares to ideal settings in the past, and then makes recommendations to the operators on how they could be optimally running their machines. So we have things like that which are fairly out of the box in our application, but we also do other sort of customized analytics for very specific use cases.

So for example, we have a customer who is a magnet wire manufacturer. So magnet wires exploding in demand right now with electrical electric vehicles, and we're helping them with quality and they have a lot of these cases where. It's really funny to sort of talk with them because the engineering teams often have sort of an idea, everybody has their own sort of pet interest of what they want to work on next. But you always get these battles between the engineers of where they should be focusing their attention. And we've been able to actually come in with this new analytic for them, which looks at their production processes and actually makes recommendations to them to say, "Hey you know, among the many opportunities you have today, this is what the algorithm says would be the most productive area would give you the most benefit, least cost in order to do your process improvements."

And so that's been very tailored to their machines, their production lines, their processes, but still based on that same underlying data platform that, that we provide. So we can do customized work for them too.

Peter Seeberg: The first phase you mentioned the detecting sounds a bit like only now, what, a couple of weeks old ChatGPT, forgot the actual word of it recognizing whatever file typically CSV, I will give it. And will tell me, okay, this is sound, seems to be like a manufacturing file with rows and columns X, Y, Z. And again, here, we'll come to that later. That's the actual theme for today, but you, you have been doing that again, also probably for maybe since the beginning, or at least since since a longer time detecting the quality and the quantity of the data that is available.

Kurt DeMaagd: Yeah, exactly. It's started off with us doing it all manually by hand. The customer, you know, gives us a CSV file or we connect to their database or, you know, go to their OPC server or something like that and pull the data. And then over time you start to write a little bit of automation for it. And then you start to figure out the rough parameters to build up some algorithms, train them on your historical data. And, you know, just sort of evolves the point where you can start to create a lot of tooling and let the machine learning handle it.

Now, of course, I don't want to overhype AI. I think we're perhaps even entering an era where AI is getting so overhyped and expectations are getting so big that it could end up hurting the market.

But, so I don't want to overhype it, but it is pretty cool what we are now able to do with detecting and responding with machine learning.

Peter Seeberg: Very good. I'll come to the OPC a bit later on as well. For specific reasons, I want to stay a little bit longer here in the area. We don't want to go in too much detail of when talking about machine learning, but I saw that you are using several kinds of approaches. I saw gradient boosting, I saw support factor machines, which I have not worked with myself, but I do recognize colleagues having worked with them years ago. I haven't used the genetic algorithms myself, Monte Carlo simulation. I know also from chess engines, etc. Again, without needing to go into the details again, who will, or where is it again, we're going to come to the user of your solution later on, but. Is the user selecting or is that again an algorithm algorithm? Are you running different approaches in parallel and whatever is the best for this specific use case will will be chosen. How does that work?

Kurt DeMaagd: A lot of it is what is the best solution, and I do think this is an important flexibility that we have where since we aren't coming out as an AI algorithm company, really Sight Machine is about having AI-ready data. We're focused on the input into those, and that gives us the flexibility to choose whichever algorithm is going to be the best fit.

Now, of course, some of that, you know, when we're creating products that are gonna be out of the box for everybody, yes, Sight Machine is going to choose that and you know, customers don't get to just say, "Oh, well, we know your optimizer is an svm. I would much rather use a random forest. So please change it."

We don't do that level. But when we're doing more the customized applications, you know, we do actually work and collaborate with the customer. Now, a lot of those customers are just much more on the end user focus and they're sort of, "I don't care. Just give me the most accurate, the best fit." But we do also work with customer data science teams and some of them will have a more stronger opinion about how they would like us to tackle the problem.

Peter Seeberg: You already mentioned availability. for the cloud. Is this exclusive partnership and or how about edge if people want if potential customers want so?

Kurt DeMaagd: So with the way Sight Machine is architected, we do have edge support. In fact, we have an edge device at every customer. And the reason for that is, of course, we're pulling data from the manufacturing plant floor, so we need something down on the plant floor to push the data up to the cloud.

You know, obviously there's massive security implications if we were to try to go from the cloud and pull. So we do have an edge device. There is, I'm going to say, a modest amount of compute there, and so we can do lightweight work at the edge. Now it's not designed at the same level of here's a gigantic compute cluster for training a machine learning model, but we can do some lightweight work there.

We're also working with other partners though, where we may need a little bit more of a heavier edge compute availability. And so working with companies like Siemens has been the most recent one taking advantage of some of their technology at the edge. But we still push a lot of stuff up to the cloud. Very strong partnership with Microsoft there. And that's where we try to do the heavy lifting for anything that's going to require a lot of compute. We also really try to put as much stuff up in the cloud, too. So then it becomes available to everyone within the organization, not just somebody in the sort of fall four walls of the manufacturing plant.

Peter Seeberg: Okay, so what industries do you sell your solutions? Then what are your typical customers? Is it more in discreet? Is it in process as well or both of them?

Kurt DeMaagd: So we're not necessarily fit to just one or the other. We've sold to, you know, everyone from, you know, chemical manufacturing to auto companies. And so we run the full gamut.

What I would say is kind of interesting. We've had the most traction recently, kind of in these companies that blur the lines. So a lot of consumer packaged goods, food processing, that sort of thing, where technically there is a discreet object moving down the production line, but often at such a high volume of production that starts to look much more continuous, behaves kind of like a, a continuous flow process. And so the fact that we can sort of float between the two is powerful for us.

One other thing that I would also mention is we often work on production lines that kind of go from one type to the next, where we might have a continuous flow going to a batch process, going back to a continuous flow, going to discreet for final packaging. And so we get a lot of variety.

Peter Seeberg: Now listeners of the industrial podcast may have heard this before, but I actually was in charge, that's probably now 10 years ago of, in that case, it was discreet. And the person opposite of me was in charge of process. And we had to decide, you know, who's going to take , where's the line.

It's exactly that line that you talk about. And I think we said that as soon as fluid out of process, it doesn't matter if it's chemical or Coca Cola, as soon as it gets bottled being packaged as a, I believe that's where we sat. That's, that's where the line is. Now I was, it's not about sharing this information with our listeners, it's more about does that mean that the approach in one or the other, or exactly where the process is changing is actually different, I mean, from the perspective of the, the capabilities of the platform.

Kurt DeMaagd: So that's a really good question. For capabilities of the platform, we've really tried hard to engineer and ensure that you can get the same types of insights wherever you are.

But as you point out, this is a huge challenge where, for example, a recent customer that we we've had is exactly the same thing where you're going through one type of process. There is some sort of distinction kind of from the processing to the packaging line or something like that. You have to draw the line.

But when you're somebody like us, where we say we are trying to give you the system- wide view, you're not just looking at packaging. You're not just looking at processing. You want to understand how packaging affects processing or processing affects packaging. You need to make sure you have those capabilities.

And so we've invested a lot of time and effort into making sure that we can off the same features and offer that same traceability across the somewhat arbitrary boundaries of the manufacturing line or even not arbitrary, very clear boundaries. We want to make sure that you can monitor the entire system no matter what's happened.

Peter Seeberg: Okay, what are the typical users then of your solution? Let's say pre factory copilot, which we'll talk about later.

Kurt DeMaagd: Our typical user was either a process engineer or a quality engineer. So, somebody who is, you know, on the plant floor, clearly, but often somebody with a little bit of data savvy mindset so that you know, they were somebody who in the past was at least connecting to, say, the historian and looking at some basic charts or maybe even did a little bit of introductory statistics types of things. But you know, we're now starting to stretch beyond the ability to just stare at charts in their historian and wanted to better understand and solve problems by getting that system wide multiple different data source type of view.

Peter Seeberg: Before we're going to get into the Factory Co pilot theme for today, democratizing industrial data. What makes you unique? What is your USP? Why should I, why should the listener consider maybe to, to visit your website and get a little bit more information?

Kurt DeMaagd: So we do have this approach to doing our data modeling, which is that data-first approach, where we want to ensure that we're not just giving you some data for an application, we want this to be a data foundation, this tool, which then enables many, many, many use cases. And I would say that is the key differentiator in terms of overall approach.

Now we can also go down into some of the underlying technology because we have been trying to connect to so many different types of data sources we have a pretty good, a fairly robust set of different connectors to be able to extract data. We have lots of experience just working through, we'll call it the IT security process of accessing data. So as an organization, we're pretty good. Our technology stack has been designed to address common concerns from IT security groups.

Our ETL pipeline is really focused on the types of data problems you would have as a manufacturer if you want to have real time data, which can also handle a lot of out of order data, for example, quality samples coming in late, things like that. So a lot of out of order reordering of your data, stateful transformations, which is maybe just a big mouthful for things like being able to do anomaly detection or other sort of data wrangling to clean up your data as the data is streaming through and of course, real time blending of multiple data sources. So the ETL engine we have is, is pretty cool too.

But as I said, at the start here, I think the fundamental difference for say an end user is really our philosophy, our approach, the way that your data is going to come out of this with this data first foundation for other applications.

Peter Seeberg: Okay, and that's where I had, I already announced kind of at this this follow up question, you mentioned connectors, but then it's more about the OPC way. When you talk about contextualized data foundation and now I'm involved in OPC Foundation podcast and OPC UA as a, as an architecture, you mentioned that you are capable of reading from OPC UA data sources in addition to other connectors, whatever they are, I'm not sure, PROFIBUS, PROFINET, I think there's typical European, American market approach differences there, but how about, does, does OPC UA play a role or, and in the markets in which you're working today and, or do you use your own kind of contextualized data foundation approach?

Kurt DeMaagd: For OPC, especially OPC UA, it's music to our ears when we hear that the customer has data sources that are OPC UA. It's just so much easier to connect you know, OPC DA you know, that's a little bit more problematic, of course, given the, the original, yeah Windows security issues and stuff like that that you run into commonly, but I won't rant about OPC DA here. OPC UA though, easy to get access to. It's got some nice security protocols in there too, for those who care about security, easy to explore.

Like I mentioned before, a lot of what we're doing is the sort of AI work where we need to sort of discover which data fields are available. And so that ability to easily walk the sort of overall hierarchy. So if they don't have a predefined data model or data schema, we can more easily discover what is there and then help them formulate those schemas, et cetera.

And so, you know, OPC is not required for us, but certainly when we see that it makes everybody's life a whole lot easier and really accelerates projects.

Peter Seeberg: Perfect. We completely agree on that. I mean, independent of, but you know, I've said in the past, you know, OPC UA makes your data fly, you know, typically when you, when you, when you start, I mean, and that is again, different if maybe you're using your solution because you do exactly the same, you know, the data first and maybe it's not going to take necessarily whatever of nine weeks that you have for a small project.

Maybe you're not necessarily going to need the typical 80 percent of, you know, do I have data, what is the quality. And I've always said, if I've, I heard you exactly say the same, if you have OPC UA, you can just start immediately by looking at the information model and like, you know, 10 minutes later, if it's an hour later, it doesn't matter, you're starting.

So we, we strongly agree there. Very good.

Now let's come to you introduce. I think July, it was a solution called Sight Machine Factory CoPilot of which you say democratizes industrial data with generative AI. Can you give us a quick first introduction, please?

Kurt DeMaagd: So we are bringing this out as a way to start bridging the gap between who is using the data today and who could really get some of the most value out of that data. And as I had mentioned before, a lot of our end users today are the process engineers or the quality engineers or people who have been the, we'll call it the traditional data users. But we're trying to get that into the hands of people who are not typically thinking at that level.

And one of the obstacles that we sort of saw is folks who just kind of look at the data system and say, I don't really know how to use that. I don't understand that. That's not really my thing. And so how do we actually change this data into a, or change the interface into a way that everybody can feel very interested and able to ask questions. And so, so to turn it into this, here's an expert here that you can just type and ask a question and get an answer.

And one of the really cool things that we have found beyond doing things like more traditional UI user experience research is when you can work with something like a ChatGPT type of interface, suddenly people are realizing, "Hey, I can take this question and do something really silly with it." Like, you know, have it reframe the answer like a country music song or talking like a pirate or something like that. And suddenly people are thinking, Oh, Hey, you wrote it as a country music song. Can you also make it, write it as a rap song?

And now, obviously, if you're doing that with your manufacturing data, you know, it's not directly going to create value, but it suddenly flipped the script where instead of somebody saying, Oh data, it's not for me you know, I'm intimidated, I don't want to touch that, too. They're actually wanting to get in and starting to play with their data, and it just breaks down those barriers that were in place in the past.

Peter Seeberg: Very good. Yeah, I completely hear what you say. So it's almost implicitly means, well, but the question still is, who are the users now? I mean, what is, is it that you're targeting? It's a new solution. Understand the same as before or typically more utter. I mean, you mentioned before the process engineer, quality engineer, they are probably still you know, your target audience, but will other kinds of users be using this solution as well?

Kurt DeMaagd: So we currently have two different users. In the early testing and use of this, one are the folks who are commonly not using the data in the first place is often operators, maybe on the second shift who aren't necessarily interacting with the engineering team. Folks like that who don't have that expert to go talk to and weren't using data through the traditional interfaces, but now they could just go and ask a question by typing a question into this interface, as opposed to having to understand how to use all the different Sight Machine tools. And so we're opening things up to people who weren't using it, such as operators on the plant floor.

The second use case that we're seeing is actually the same traditional engineering users where maybe there was somebody who had to come in an hour early to pull down this data, aggregate, synthesize information to be ready for the morning meeting. Well, now, instead of having to do this all as a manual process, it can much more easily go through and just aggregate the information, sort of write the report for you so that you're ready right away for that morning meeting so you can kind of more, more efficiently synthesize information.

Peter Seeberg: Right, so it's still the traditional plus, or you're extending your target users. Actually, you mentioned. Typing, yes, I can imagine. How about talking? How about talking to whatever, talking to the plant, so to say. And if, if we talk about talking about languages, do I today, have I in the past, have you been providing English only? I mean, how about other languages, Spanish, German, French, et cetera.

Kurt DeMaagd: Yep, excellent question because we have been working on both the talking side and the multilingual side. So today you can actually do text to speech if you have it built into, well, of course, you know, most computers these days, you can just hit a button and do a text to speech and then do that with our interface and it will reply.

So that is an option, or it is kind of neat, you can do that right on your phone. If you go to a copilot application, just kind of go there, hit the button, ask the question. So you can imagine being say in the car, you want to be properly hands free, but you want to ask a quick question about, you know, is this happening right now? Technically you could do that. Hopefully you're keeping your eyes on the road too, but that is now an option.

That said, we have not gone super deep. There have been a couple side conversations with other vendors who do things like the walkie talkies for the plant floor sort of thing. Can we integrate with that? So there's some interesting opportunity there that I see in the future, but we don't have anything deployed today.

What, though, also is really cool is the multi language, multilingual parts here. For example, the magnet wire plant that we've been working with, I mentioned before, is in Mexico. We have therefore a lot of folks who are going to speak Spanish and with our interface, one of the first steps that does is checks which language it is and handles that appropriately.

So if you wanted to ask a question in Spanish and also make sure it's replying in Spanish that that is fully supported today. And so it's still something that we are testing to make sure it works since we are relying on open AI on ChatGPT to do that translation. And sometimes, you know, these language models are not necessarily trained on manufacturing jargon, and so you run into some issues even like with an English. If you ask about performance, performance could be a generic statement about. Hey, how is stuff working? Or it can mean like capital P performance in an OEE type of sense. And so making sure that things like that translate properly. And so we're still working through a little bit of that to make sure that the edge cases of manufacturing jargon translate correctly.

But today you can, you know, do something like ask a question in Spanish and German and French and whatever language you may want to try.

Peter Seeberg: Muy bien, sehr gut. Yeah. And of course, I mean, the topic of the languages is in relation to the large language models is a topic for itself, where we as, let's say, a global community, independent of manufacturing, probably want to make sure that we're going to continue to have, not sure how many we have, 150, 200. probably even more dialects, maybe a thousand so that they're not gonna die out because we've been doing everything. And that's the next question. Everything has been trained on English mainly, not exclusively. So but your solution then typically is what, is then also, as you mentioned your partner, Microsoft OpenAI has been trained, has a large language model, I assume as a basis, but has been trained in addition on specifically customer data as well.

Kurt DeMaagd: Yeah. And so it's trained on a combination of customer data. And then of course the core large language models. So specifically we use GPT 4 underneath the hood. Now I should of course mention here that although we feed it a lot of data from customers, that is all done sort of locally within a deployment and through a combination of, you know, sort of model tuning and prompt engineering so that none of that data then goes into anything that can be retrained on a ChatGPT someday. So obviously you're not going to just type in something a year from now and suddenly discover that your confidential information is somehow leaked out. So we do take that part very, very, very seriously.

But yeah, a lot of our effort then goes into how do we take that general purpose, large language model and fit it with the necessary data and know how of manufacturing to solve these problems.

Peter Seeberg: Very good. Yeah, we've all heard about the colleagues from Samsung, who probably were the first, at least we heard about it. Something else happened within there. We don't want that to take place. Now, your colleague, co founder Jon Sobel had a quote, AI can't work with a raw factory data. AI needs to fully contextualized data. You want to comment on that statement from Jon?

Kurt DeMaagd: Yeah, there's a lot going on in that statement. Biggest thing that we are really trying to address when we are making our data AI ready is you can't just take a whole bunch of CSV files, parquet files, whatever you have, shove it into an algorithm and just hope that it magically figure things out.

So a lot of what we're doing when we're getting our data AI ready is trying to blend things together into a coherent format so that it has the necessary context to minimize the spurious correlations, to clean up the types of data anomalies that are going to trip up an algorithm et cetera.

And then also make sure that we are doing the analysis in a way that takes advantage of that context and can be the, the best fit to the final output. So certainly we can dig into some of those topics a little bit more, but fundamentally, if you go back to the early big data era, there was sort of this idea of, Hey, just get lots and lots of data, shove it into a neural net and magic is going to occur. And I think, within a year or two, people started to figure out, Oh, wait, that's nothing more than a spurious correlation engine.

Peter Seeberg: Yeah. Yeah. Right. Didn't take me 10, 10 10 minutes. Yeah. Right. It was sounded very much like at that time, marketing, a speak of, of specific companies. Kurt, thank you very much. Tell us a little bit more about your company, company size, maybe a global representation. You, you already mentioned. One, two pieces. Not sure. Maybe you're looking even for new colleagues if so, where and what capabilities should they bring.

Kurt DeMaagd: Yeah, so we are a global organization. It's kind of hard to be in manufacturing without being global these days. So I believe we have representation around the world, around the time zones. In total, we are about 160 people and that is heavily based San Francisco for engineering team. A lot of our we call them the outcomes team, but sort of data engineers, data architects, continuous improvement, Six Sigma, all types of folks are based in Michigan.

We have a nice little cluster of folks in the UK, but then as I mentioned before, we're also just scattered around the world wherever we find the talent.

Peter Seeberg: Sounds great. Coming to a close. So what then is, as far as you see, the status of this democratizing industrial data with generative AI? in different parts of the world. I mean, I believe you're one of the first, if not the first, feel happy to confirm that. Is it the U S only today? I mean, exception of the fact that you are global is what do you see happening in China, Asia? Have you seen things in Europe, UK, you just mentioned, and from your perspective, how is then Gen AI going to change the industrial AI world over the next five, 10 years.

Kurt DeMaagd: Yeah, it's going to be very interesting cause we have everything from variable regulatory rules, which are going to be changing a whole lot as people are really becoming, I'm gonna say it's aware of how much data AI consumes and the potential risks. I think AI has been, you know, these, these risks are nothing new, but suddenly ChatGPT brought it into a format where people individuals can start to see that.

So, obviously, anything that, you know, I speculate about is subject to regulatory issues that may be coming up in the future. But let's just assume that we continue to have a pro- AI regulatory environment. I think a lot of what we're going to start to see is a little bit more appreciation for what it can and cannot do.

It has been really interesting to see, even with things like ChatGPT, how people describe it has been so different. Where in the past, you would have, for example, seen somebody write a blog post about some AI for helping them make investments in the stock market. And they would talk about their algorithm and how it's helping make decisions.

But it's fascinating, now they talk about I gave ChatGPT $20,000 to make an investment as though ChatGPT is, I'm going to walk down the street to the bank and, you know, make an investment. And so this personification of AI is really interesting to see how that's going to change here in the next, you know, a few years, are people actually going to look at even, you know, for Sight Machine, are they going to sort of look at our CoPilot as an expert that they are talking to, or are they just going to treat it more like a traditional chat bot that you may have had for customer service or something like that for the last 10 years? And so we want it to be as democratized as feeling sort of human as possible to break down those barriers, but at the same time recognizing that there are limitations to this.

You can't just ask it anything and it will magically figure it out. We have to kind of put those skills in underneath the hood to make sure it has the data to make sure that we're controlling the way it knows how to answer it to ensure the accuracy. Things like that. And so a lot of what we're trying to figure out right now is how to balance that, you know, ensuring people feel comfortable that this is your expert partner, but not missing out on the fact that it is not actually a human in the background that somehow magically is omniscient and can answer everything. And so we're doing a lot of work there trying to understand understand that.

Thank you very much. Closing off on this notification of personification, which I interpret as, you know, like a personal assistant who will be helping not only the people you said before, you know these specialists, but also the people who have been a lot further away from the data. Thank you very much for that. So listeners who want to get in touch with you can best do so. You suggested on LinkedIn Kurt with a K - K U R T. And then there's a space D E M A A G D. Although you pronounce it as de Marge, you suggested. Otherwise, if you dear listeners, if you have any questions, comments, as always, please send a short email to Peter at AIPod.de. A great you stay with us so far, looking forward to have you with us again. And Kurt, thank you very much. Have a nice day.

Thank you. My pleasure, Peter.