**Dan O'Connell:** 0:00

The piece when you mentioned that the data set of what you trained with, which is unique for us as being a communication provider, we power all of these conversations and our founder, five years ago, you know, had the foresight to say look, if I can go again, power communications on any device and transcribe it. Those trains, the script, it's the training data that we can then go and leverage and build to power this large language model. A lot of our features are designing for a career, teams for sales organizations and customer success and service teams. Those costs that we absorb ultimately are sold or passed to the buyer right. There's very few AI features that we do. Those are just three features.

**Craig Smith:** 0:40

Hi, I'm Craig Smith and this is my AI. We're excited to have Dan O'Connell, chief AI Strategy Officer at Dialpad, joining us today to discuss how artificial intelligence is transforming business communications. Dialpad offers a cloud based communications platform aimed at understanding conversations and driving automation, and Dan provides an inside look at how they're using AI like natural language processing, speech recognition and large language models to deliver features such as real time transcription, call summaries, sentiment analysis and virtual agents. I hope you enjoy the episode as much as I did. Before we begin, I want to give a shout out to our sponsor, YouAI. They're a fascinating company that does a number of things, but I'm most interested in their Mind Studio, which is a platform for building AIs on top of large language models that you can deploy for free or for profit. Think back to the days when smartphone apps were first getting started, and if you were able to have been among the first couple of thousand people to build a smartphone app, you would have learned a lot and may have earned a lot of money. So give Mind Studio a try. Visit [youai.ai/mindstudio](https://youai.ai/mindstudio) to build your own AI today.

**Dan O'Connell:** 2:24

Yeah, I'm Dan O'Connell, chief AI Strategy Officer here at DialPod. I've been here for five years. They acquired my startup five years ago, which was a real time speech recognition we start up called OQAYIQ. Largest investor was Salesforce Ventures. It was our belief at that time too. If we could capture conversations and transcribe them and analyze them, then it opens up a lot of opportunities around driving, automation, insights and assistance. And DialPads it's a cloud communication and collaboration software. We focus on customer intelligence, which is how do we power communications on any device anywhere in the world and then how do we help businesses understand what's happening within those conversations? And my role today is I oversee our strategy, our partnerships and our product teams.

**Craig Smith:** 3:16

OK, how does AI play into this? And I'll use the bit earlier where you described what DialPad does, but how are you using AI?

**Dan O'Connell:** 3:34

Yeah, and what's unique for DialPad is we were a spin out from Google. For a little bit more history, our founding team, myself included, came from Google. Our founders had built what was called Google Voice. At that time, Google didn't want to commercialize that product. This is roughly 12 years ago. They left. Google Ventures wrote them the first check, which minors the founders, the Android Operating System, freddie Bates that's a lot on the iPhone, that's on our board and that really sparked the genesis for this communications company called DialPad. And then over the past five years, as I said, it became regularly apparent to say look, our focus is not just about powering communications on any device. When I approached Craig or I and Craig as our CEO around partnership, it was hey, if you can go. Power communication, that's one thing, but every business wants to understand what's happening within those conversations, whether it's to better run the business and make better decisions by understanding insights, or whether it's about driving automation or providing assistance like suggesting answers to questions. And so when they acquired our business AI runs entirely through our stack. We own our entire AS stack. So we do all of our own speech recognition, which allows us to do things like real-time captioning or transcription. We do all of our own NLP work, which allows us to do things, and NLP would be naturally done with processing. So, for those that may not be aware, it allows you to do things with text for a computer. So I can start to understand sediments or I can infer customer satisfaction. We do our own semantic search that powers our own recommendation engines, which allows us to build features like suggested answers to questions. And then just a week ago, we announced our own large language model, which again, people get really excited about in terms of what does the next generation chatbot look like, or at least next generation capabilities within communication software. So we send this unique position. It's a long ramble. If we own our whole stack, we're really proud about it, and it's really about powering communications on any device, anywhere. And then it's about business understanding what's happening to you, insights, automation.

**Craig Smith:** 5:39

Yeah, okay, so I'm. We're on dial pad now. Are there features here? I see AI notes and on the menu bar. How do some of these things work?

**Dan O'Connell:** 5:57

work. Yes, so we focus on our meeting platform here. Right, we have AI notes at the bottom. You can click that and see a real time transcription happening. What we do at the end of a conversation is use a large language model, our own large language model, to power instant summaries. And you'll see some other things that can pop up in AI notes at the end of calls. When we generate summaries, we automatically capture the questions that are being asked and identify the topic and then again, what you'll get at the end of this day and I'll share with you. It's an instant summary conversation, which is great to have this long transcript. But when you suddenly realize, right, it matches it as match, it gets transcriptions and what you quickly realize is nobody wants to review a transfer, right, people want to get the short gist of a conversation. And then if you think of you know, this meeting platform is just one, one piece of what we have. But if you think about us as a contact center provider for businesses, so powering their service and support teams, suddenly we can go not only transcribe those conversations, we can identify the topics that are happening and action items and again generate those summaries and things and we write that information to the CR route. But the nature of these features and I can share some demos and features on things is the unique part for us is we can do all of this in real time on it, so that opens up these really unique opportunities to do insights, just to answer questions around a conversation. What is the exact unit?

**Craig Smith:** 7:28

Yeah, I mean you mentioned sentiment analysis. How would someone use sentiment analysis on this platform?

**Dan O'Connell:** 7:35

Yeah, we thought about it. You know, the interesting thing about such an analysis is our approach has always been thinking about it in three ways. There is sentiment that happens at the start of a conversation, and let's, let's, let's focus on sentiment around a support conversation system, recalling into a contact center. I think that's most relevant, that's what we see as most interesting. You know, you and I can get the sentiment of this conversation. I don't think it's going to be terribly informative for us. But you can imagine you, you care about customer satisfaction and so we think about sentiment in real time. Meaning you know Craig called into our contact center running, running our software. You would see a streaming sentiment of that conversation. You look at it in the 32nd window. So we know that Craig is getting more agitated or less agitated. Hopefully if you call in, you're pretty agitated, right? You may ask for an escalation to a manager and hopefully that agent is talking you off the ledge and will solve your issues, and so that sentiment is becoming more admissible and then leads to a call response. So one is you see sentiment in these 32nd windows, always machine Once the call ends. There is then the notion that you need to think about sentiment across the entirety of that conversation, right, and that ultimately evolves into customer satisfaction, which is what you called in, did we resolve the issue? And what happens for many businesses is call ends and you typically will receive an SMS or an email survey that says hey Craig, you know what? Would you rate your experience? The problem with doing those SMS surveys Is that very few people reply to them, and the people that end up replying tend to be in two camps. One, the agent did a really exceptional job and somebody must have seen your praises, which doesn't happen a lot, or it's that I'm really fed up with my issue wasn't, wasn't solved. I want to tell you how bad and frustrating it is. So it tends to be a little bit whole. The data tends to be color. The nice part of thinking about sentiment for the end of a conversation is suddenly you can measure sentiment in real time. It gives you an indication of customer satisfaction for every single conversation. So that's the second way that we do this and again it's kind of like at this magical unlock for these businesses because it's hey, you don't need to go buy new software, you don't need to send out surveys and, by the way, we'll give you far more data for every single interaction. So that's the second way. And then the third way is you want to start stitching together sentiment across multiple points of contact in the same business or multiple calls from the same person, and that gives you some notion of this of account help, which is it's not that Craig has called in once, it is Craig, it's called in three times in the past week, with increasing frustration right or decreasing customer satisfaction. And then that gives you the notion for these businesses to start thinking about proactive service to say I don't need to wait for Craig to call me the fourth time. So I'm pretty sure we will call the fourth time, it's gonna tell me to terminate the contract and you want some reason. But how bad we are. It's gonna go hop on Trustpilot or G2 crowd and tell the world that this allows you to start then getting to a place where you can help businesses be more, far more proactive and ultimately play offense. So those are just three examples of how we think about sense mix. I think you know one time we know a lot of people think sentiment is like this easy thing and suddenly it kind of brainstorms into these others. The last piece of my ramble I would say is we focus on the context of conversation Anytime. We get into sentiment analysis, a lot of times people ask us about intonation and tone as opposed to context, and we really focus on the context of the conversation and what our data scientists would tell you is when people call into a contact center, they're very specific around their needs. Most people don't call into a contact center and have sarcasm play around. They're very clear about what they're looking for and so we don't spend time looking too much at intonation and tone and we think you can get pretty much ready. You can build very direct, very accurate and directional models from just focusing on the context of where it's being set.

**Craig Smith:** 11:48

Yeah, the large language model. So instead of leveraging you know one of the commercial models through an API, you've built your own. I'm curious about why you decided to do that. Or are you, when you say you have your own model, are you fine-tuning an existing model that you access through an API?

**Dan O'Connell:** 12:19

Yeah, so, yeah. So there's a couple of different angles to that. So one is as I share. One of our Q-Philosophy is about our businesses. We want to own our own stack and I think what's important around that and being a builder of startups is, if you own your own stack, it allows you to control your own destiny in terms of innovation and if you leverage other third-party APIs. It's my belief and there's no right or wrong way to build software. It's just my belief. You know this is right. If you outsource or leverage off-the-shelf pieces or APIs, you don't control that innovation. At times, the API might not do things that you want to be able to do and you've repented on somebody else to then go and make changes and change their roadmap. What we found in these initial foundational models so let's just you know we played around with ChatGbt, for example, for OpenAI, chat gpt is one where we got unpredictable outputs. So, and I think many people can experience this, when you go to ChatGbt, sometimes you ask it to do something, you may get a millisecond response, sometimes you might get a second response, sometimes it might take 10 seconds, sometimes it might tell you they were timed out on capacity of demand. We can't process this right now. Try again later. If you think about building up production, you know building a feature that should be commercially available, you can't have that experience. You need to be able to make sure that the user experience is really exceptional, and so having unpredictability in response is just one of the issues. Number two is unpredictably on. You know they had offered fine tuning as a. I think yesterday they made announcements, but if you think back to they've been working on this for four months there was no fine tuning right on those pieces, the GPUs that are needed to go and run these liberally large foundational models that are expensive at scale. And also, if you think about it, you know specific needs in the contact center. If you think of a large well, a foundational model is kind of like the library of Congress if it's got all of this knowledge. You don't actually need all of that knowledge, and so it's our belief of you can build a smaller domain, specific, largely, which model that needs less or less demanding computing resources, allows you to go scale, it allows you to have predictable output, allows you to reduce hallucinations and then is more cost effective, allows us to drive our own innovation road. So it's our belief that our philosophy is all of those benefits, but I will say there are places within our pack, in our tech stack, that these, that the large models, make sense and will leverage, and Google is wasteful Our trusted partner on that side. So Google Vertex is a foundational model that will have some specific use cases within our stack, but a vast majority power.

**Craig Smith:** 15:09

Yeah, and the model that you build how? How large is it? I'm curious about what I've been talking about this week about the cost of training large language models, not only of putting the infrastructure in place but the power generation and the power consumption needed to train and then, of course, collecting the data. So, yeah, yes.

**Dan O'Connell:** 15:36

So two things that we don't, we don't comment on. On parameter size, because I think one thing when you get into size is there is this belief that that larger is better, and I think what's true for many use cases and get it at software is you want something that, that and this is almost stopped after people which is faster, cheaper, better for customers. And so the fortunate part for us is, look, we do not have a model that is going to be the same number of parameters and our intention is not to build a model that competes with a large competition. It is. We are building a model that is designed to be smaller and is designed to be domain specific and use case specific, right, and so that is going to provide an enhanced experience that's going to be more predictable in terms of the output. And again, when you talked about trading and building up, more cost effective for us to go and manage right and scale the, the piece, when you, when you, when you mentioned that, that data set of what you train with which, unique for us as being the communication writers, we power all of these conversations and, and you know, our founder, five years ago, you know, had the foresight to say, look, if I can go again, power communications on any device and transcribe it. Those transcriptions, the training data that we can then go and leverage and build this large language, could power this large language model. So we focus on one option in user data. So we're clear around. The users opt in to show their data. We will take that data, we will anonymize it, we will strip it of any personal identifiable information. We're so hooked to HIPAA compliant, you name it. On compliance, we have very clear on and off switches within our software that allows a user to turn it on and off. And then you get into data retention. It's your data, you control it, you should be able to blow it up. And so, all of that said, you have access to this data that you can go and use. So one leverage to go and build this large language model. And you have the team of experts that sits in the house that you know has experience in NLP and speech recognition and machine learning. You have this capability. So, again, that's part of the reason we all wanted to do this in houses. We have that expertise and things that we do.

**Craig Smith:** 17:48

Yeah. Can you comment on the cost of training your model? Is it as expensive as everyone claims?

**Dan O'Connell:** 17:59

Extensive, and I'm not trying to dodge terms, I don't. I will, as he's having trouble seeing me suddenly, will probably yell at me if I tell you that you know the specific costs For us. Like we are not public business right or privately funded. We have to be mindful of our costs. And you see the smile, you see me, you know, smile, laugh, or I'm like I think it's a reasonable cost, that that again it comes down to. I think the costs are manageable. We think it's the most important thing for us. When we talk about having five pillars as a business, ai is number one on that pillar. So for us to be worried about the costs, to say that, hey, this is our most important thing, we're worried about the costs of it and we have to manage costs. As I said, as a private business, we have to focus on that. And then a two is we have the data set and then the in-house team of experts to be able to go and do that. The other thing with these, with these investments, is you know we're focused. A lot of our features are designing for recruit teams, for sales organizations and customer success and service teams. Those costs that we absorb ultimately are our soul. They'll pass to the buyer right. There's very new AI features that we do. That is just three features we do. Do you know, and I apologize for these large mammals? Transcription, for example, is a free feature. We don't charge for that. The reason we don't charge for it is, again, we do all of this in-house. We do it for basically one tenth of one penny in terms of doing the transcription, and then I'll P cost like we are far more cost effective and far more accurate in terms of the output than if we were to use a third party effort.

**Craig Smith:** 19:48

Yeah, so how did your dial pad exist before generative AI or the transformer algorithm came along, so how did that affect your product?

**Dan O'Connell:** 20:06

Yeah, yeah, we've been an AI first business for the past five years. I know a lot of businesses will say that they are AI first or AI power, but we truly, since our acquisition of showing up and we had built one of the first real-time speech recognition engines, we have been AI first business. I will tell you this really interesting moment in time. You know, I grew up in the heart of Silicon Valley and, unfortunately, two passions growing up were sports and computers. I look at this moment when chat TPT showed up and we literally, the next day had pretty much an all hands kind of executive staff meeting where it was hey, the game has changed. And we had, obviously, you know, our NLP team was, you know, I say initially concerned, because I think there were those reactions like, hey, do you need an NLP team in this large language model suddenly be able to do everything you've invested in overnight, and what does that mean for the competitor dynamics, for businesses that have never had AI before? Can they suddenly go and show up and do everything that you've done overnight? So we literally had the first time we had played around with it, with the, maybe, the, the, the, the oh crap moment, where you're like, can you better align and understand this and invest in it and start making decisions. And I think that's the nice part of being in a small startup that again cares about innovation is you can have those types of conversations very quickly and trigger out where the opportunities are, because we're not the cruise liner, we're still a small speed boat and so we can be pretty nimble. If we need to go, change direction and just wrap the red map, we can go and do that.

**Craig Smith:** 21:49

Yeah, and to that point about being smaller and nimbler, how do you know that I'm getting an alert? How do you compete with a company like Zoom, which is adding a lot of similar features themselves?

**Dan O'Connell:** 22:10

Yeah, I think for us, there's a couple of different ways that we would compete with you. We offer a single piece of software to get power, communications, to understand it, as I said, whether that's in the contact or sales side, whether it's for recruiting teams or just your back office, to let it be. For us, it's been heavily investing in AI and being able to tell that story again, having features that are going to drive insights, power assistance or provide automation, and I think speed matters at the end of the day, and I think that's something that can get glossed over at times by different people in markets, the way they look at it, but the reason that the largest businesses in the world constantly, over time, get disrupted by smaller businesses is that those smaller businesses push innovation. They carve out time for innovation, they don't have to manage their business to a public market and some of the pressures that show up there, and they understand that they're not the king of the hill. The leader in the market has everything to lose, and so when I look at Zoom, I think we can out innovate them in terms of the pace of innovation. We've demonstrated that. I think Zoom has had an incredible run up in terms of the opportunity through COVID, and I look at them personally. As you know, you're at the top of the hill and you have a lot to lose and we got everything to get, and so that means that we are not a public business. We get to run our business differently because of that. We get to take risks because of that, and our executive team has aligned on that, and AI is the key innovation. It's going to be the place where the battleground is kind of won.

**Craig Smith:** 23:50

Yep, I'm trying to figure out where this noise is coming from. Is that from Dialpad? I'm getting a? You don't hear it on yours, I don't hear it. Yeah, I don't yeah. And where do you see this going? Not only Dial Pad in your market, but the spread of generative AI and large language models.

**Dan O'Connell:** 24:20

Yeah For these. I think the things that get me excited is obviously, I think it's readily apparent that these digital agents, or next generation chat bots that are actually capable of let's just say, I think the chat bots that we've gotten in the past have been good at password resets and probably but I think it's readily apparent that we're going to have these digital assistants that can guide us, whether that's guiding us through music theory, whether it's guiding us through how to play piano, maybe it's guiding us through the workouts that I should be doing and designing a workout for me, maybe it's having, again, the digital agent that's writing my first draft of the blog post. I use large language models a lot for generating ideas and first drafts, so I think that's very evident for that, for our business in the communication collaboration space. For me, what I think that the large language models open up is the ability to completely disrupt the CRM market. And then I constantly think about you know, how does a business like Salesforce get disrupted? And if you and I were to build a CRM business today, we would build it the exact same. The CRM business today has always been reactionary. You know, you and I have a conversation, I go into a database and they write my notes in there and help play it. But if we were going to go build that, knowing having the technologies that are available today, you would probably give away communication I would want you to work on. You know, power that, whether it's SMS or team messaging or voice or video, whatever it might be, I want to understand those conversations by nature. I put a large language model on top of it to do the automation and categorization of that data and then write it into the database that I own and I would completely say you can't write into the data. No person can write into the database, right? We're going to use the machines to write it, so it's always structured. You know, that's the perfect vision that I think we get to and I think that business has not been built yet by anybody and I think it's one of probably the biggest, best opportunities out there in all of software.

**Craig Smith:** 26:26

Okay, is there? I'm running out of questions here. Are there aspects of Dialpad or of the integration of generative AI into this industry that I haven't touched on? What can we talk about?

**Dan O'Connell:** 26:49

Yeah, I think, like one of the ones that you know. I think that the one that becomes most apparent to people is, you know, the next generation chat bot, and I think people get very excited about that because they can see the type of natural experience conversational flow is largely which models are capable of, but really I think what captured people's imagination of chat GPT just sounded like you were talking to a friend. What I think is really surprising and interesting, which doesn't get as much attention, is the ability for these large language models to categorize there. And I think, when you think about all of the time and wasted energy, that that every business in the world they sit on a bunch of data and then they have different teams try to go through to categorize it, to get it structured, and I think there's just this massive opportunity for the large language models to do that categorization and get unstructured data in the structured ways and then to also then tell you things about it. Right, I've seen some really incredible demos where somebody can load up their data into a large language models and ask sit things, and so I think we're going to see the whole way we think about data analytics and data warehousing and the visualization of that is going to get completely upended by large language models and I get really excited about that and I don't think it's had as much fans there as you know that the chat butts of the future of the mutual assistance that are showing up and I actually think that other pieces more is more viable and interesting and probably valuable, at least enterprise type, that in terms of like enterprise.

**Craig Smith:** 28:24

Yeah, and are you doing that? Are there projects that use Dialpad that you're working on? In data categorization, I mean presumably in this conversation a large language model could go through and pull out names.

**Dan O'Connell:** 28:43

Yeah, yeah, we take it a step further. So it's not just about capturing, you know, things like the actual items that become rail, hair and topic model, but it's also about outcomes and purpose. So, again, if you think of what we do for our summarization in our contact centers that we generate and again I'll share some examples for you we automatically summarize that conversation in the four or five sentences. We immediately identify the purpose. So what was the reason that this person is calling in? Right? Might be a refund request, might be technical support, you know, for example, we identify the outcome. So was it resolved to follow up? Then you have any of the action items and again, all of that instantly happens through the power of a large language model. And that's just the tip of the iceberg of these that we're working on.

**Craig Smith:** 29:36

That's it for this episode. I want to thank Dan for his time. If you want a transcript of this conversation, you can find one, as always, on our website at eyeonai.com. That's EYE-ONAI. That's it for this episode. I also want to thank our sponsor, Mind Studio by UAI, which is giving creators the opportunity to build and deploy generative AI apps for profit. Uai has an emerging AI marketplace and Mind Studio is the best way to build apps with generative AI. Anyone can do it. Mind Studio uses conversational language to program incredibly powerful AI tools. No coding knowledge is needed to start your AI business today. Check them out at [youai.ai/mindstudio](https://youai.ai/mindstudio). That's Y-O-U-A-I. Dot A-I and start building your AI app today.