Hadley

Listen to Text with Envision App or Glasses

Presented by Ricky Enger

**Ricky Enger:** Welcome to Hadley Presents. I'm your host, Ricky Enger, inviting you to sit back, relax, and enjoy a conversation with the experts. In this episode, we discuss accessing visual information with technology, and our guest is Chief Technology Officer of Envision, Karthik Kannan. Welcome to the show, Karthik.

**Karthik Kannan:** Thank you so much for having me, Ricky. I really, really appreciate you taking the time out to talk to us.

**Ricky Enger:** Yes, and likewise, so happy to have you here. It's always a wonderful day when I get to talk tech and Envision just happens to be one of my favorite and most used tech tools. Yeah, a good day for me and I know it's going to be informative for everyone else as well. So, before we get into talking about the technical aspects of things, why don't you just give us a brief intro and tell us a bit about yourself?

**Karthik Kannan:** Sure. My name is Karthik. I'm one of the founders of Envision. I'm technically the chief technical officer, but essentially what I do at Envision is take a look at what the exciting stuff is happening in the world of AI and see how we can take that exciting research stuff and help it improve the lives of blind and low vision people all over the world.

It is the most fun job I've ever had, and it is hopefully a job I'll never have to give up ever. So, I'm very, very glad to be here. Very glad to be building Envision, which is basically a tool that helps people who are blind, or have low vision, to live more independently.

It is available as a smartphone app and also available in the form of smart glasses. It's basically a pair of glasses that you can wear. It's got a camera on it, and it's got a speaker as well. And what the Envision glasses do is they take pictures of things around you; it extracts information from those pictures and then speaks it out to you.

So you could be, for example, looking to read a menu at a restaurant and the Envision glasses can take a picture of the menu, get the text from it, and then speak it out to you. Or you could be out and about, and you want to know what different objects exist around you as you're walking down the road, for example, and Envision Glasses can help you do that.

Similarly, you can look to find faces of your friends and family members in crowds, the glasses can do that for you as well. And like I mentioned, Envision is also a smartphone app and it does a lot of the things the Envision glasses does. And it's a completely free app that is available on both iOS and Android for people to use.

**Ricky Enger:** Fantastic. And we’ll take a little deeper dive into all of those things as we get into things here. I think one thing that is really clear hearing you talk and having heard other interviews and presentations that you do, is that you really are passionate about creating technology that is going to empower people, that's going to give them the ability to be more independent. So, I'm always curious though, how people end up on whatever path they're on. How did you come to be working on this software for the blind and low vision community?

**Karthik Kannan:** Actually, a lot of people assume that there are folks who have blind OR low vision in my family or my co-founder's family that might have triggered us to start going down this path. But honestly, I think it was a chance encounter with a group of blind and low vision high schoolers in my hometown in India that got me on this path to begin with.

About seven years ago I went to a blind school in my hometown and I went there to give a talk to kids about what it is to be a designer, what it is to be an engineer, and I just told them essentially that a designer or an engineer is someone who wakes up and solves problems for a living. Quite a big to-do list of problems, and hopefully if I get to the bottom of that to-do list, I end up with something that is useful for people. That's essentially what we do as engineers and designers.

I posed a question to the kids after the conversation, it was a fairly innocuous question about what would they like to do or what kind of problems would they like to solve when they grow up? I was expecting something like, "I would like to go to the moon," or "I'd like to cure cancer," or some such big problem that young, idealistic kids usually go for.

But the answers that I got from this group really shocked me. It was about them being able to read books more independently, them being able to go out more independently, them being able to live by themselves or take a last-minute trip outside if they want to, and it's a nice sunny day. So, everything around their life revolved around independence.

What struck me a lot at that time was they would spend so much of their time and their lives overcoming some of the most basic hurdles that sighted people take for granted. And so much of that youth and their energy and their effort would go into overcoming these problems. For some reason that really got me thinking about how we can solve this problem or how we can improve independence for people with low vision or who are blind.

It was at that time I was working a lot in artificial intelligence, and I realized that a lot of tasks that humans were doing like being able to read text or being able to recognize spaces, artificial intelligence was starting to do just as good as humans or even better in many instances. I also realized, the world isn't really going to change a lot, or you can't expect the world to change a lot in order to make it more accessible for people who are blind and have low vision. You can't put a braille display on every single bit of visual information in the world.

So, we needed a tool that could help people access the visual world on their own terms without expecting the world to change a lot. So that bridge between the visual world and people who are blind and have low vision was or is artificial intelligence because artificial intelligence, it just needs an image. That was how we got started.

Initially, the idea was never to make this into a company. But eventually during the course of us building this as a project, it started to snowball and it started to become this organic thing where a lot of people in Europe at that time were just getting the basic beta version of the Envision app, installing it on their phones and using it. And with absolutely no marketing or no intention from our side to make this go viral, it started to go viral.

I remember waking up the next day, which was a Saturday morning and seeing my inbox had 4,000 or 3,000 unread emails. And it's when I actually opened it up, I realized all these people who are using the beta of the Envision app, writing back to me. And it was not a great beta, it was just a very ugly app with two buttons and would crash most of the time.

People were still loving it; people were still using it and that became the impetus for us to start this off as a company. And so, I moved from India to the Netherlands on a whim to basically start Envision. And yeah, we are on this journey for the last seven years. I hope we can do it for the next seventy, so yeah.

**Ricky Enger:** Whoa, that's amazing. And I think it really speaks to just how much this kind of thing is needed. Whether you're a person who has been blind always, like myself, or you're losing vision and you know that you're surrounded by all of this visual information and yet suddenly you can't access it anymore, now to have an app that says, "Hey, guess what? You can read the text on this menu, or you can take a picture of this document, or you can scan through your junk mail or recognize a product or a face." People jump on that. So, the features that I've just named are just a few of the features that are available in the Envision app.

When you started this with the beta and later kind of opening it up to more people, were you surprised by the kinds of tasks that people were using it for? Were you and your co-founder really good at predicting what people would want? Or did you have questions about could you make this feature, and you were thinking, "I never thought about that"?

**Karthik Kannan:** Oh, we were horrible at predicting stuff. We were really off the mark so many times that it's embarrassing to admit how many times we were off the mark. I think it taught us a lot about learning to listen to people, especially when we ourselves aren't the users of what we make. It humbled us a lot in terms of knowing that we can't consume things and we have to co-create whatever we are doing with this audience.

One very early example, which was very interesting to us when we started working on this feature in the very beginning. The early days of the beta of the Envision app was you take a picture, and it gives you a caption of the picture, it gives you a caption in actual language. We now have this feature called Describe Scene, which is what we put a lot of time into building.

The very first version of the Envision app had Describe Scene and Reading Text. And as an engineer, as an AI researcher, I was very, very thrilled to work on the whole image captioning feature because it's so exciting when you take a picture of something and the AI is able to give you a very nice description of the picture with the words and things like that.

Going into the first version of the app, we thought this was the feature that everyone was going to be wowed by. This was the feature that everyone's going to love and so on. But when we actually started seeing how people use it, people were using the app to read text more than any other thing. People would use the Describe Scene feature, probably once for every one hundred times they would use the Text Recognition feature. It was off the charts.

It's when we started to look into why people were doing that, is when we realized that a lot of the world around us is just text. Every product that you pick at a supermarket, the stuff that you have on your screen, everything is text. We spent probably just half an hour or 45 minutes working on that feature entirely, whereas we spent, I think, three or four weeks just perfecting this image captioning thing and making it learn all these different things. People didn't give two hoots about it, and everyone started using the texter. And that really surprised us in the early days, and eventually we shifted our focus to making sure that Envision became the best text recognition tool that is out there today, and I think we have achieved it to a fair extent.

The Envision app and the glasses support pretty much one hundred different languages out of the box. With the glasses, you're able to read handwritten text, you're able to, in over one hundred different languages, able to read text from curved surfaces like cans and bottles and so on. Ever since then we shifted so much of our focus into text and making sure that text recognition is amazing. But in the very beginning when it was just me and my co-founder, we made as many bad guesses as humanly possible.

**Ricky Enger:** We'll get to the glasses in just a bit, but still talking about the app, for somebody who is not so in love with technology, they sort of see it as a necessary thing, but it's hard to get excited about it, still. Envision is this free app and it's available for iOS and Android, including things like BlindShell. So, chances are that you may have a phone that is going to work with Envision. If you get this app, what do you think is probably the quickest and easiest task for someone after they install the app, open it, and just be wowed by, "Oh hey, this really is going to make my life a lot easier"? What's that first task that's going to get people excited?

**Karthik Kannan:** That's a good question. I think the first thing that will get people excited is the Instant Text feature on the Envision app. So, it's the very first button that you land on as soon as you open the app, and that is for a reason. Because with the Instant Text feature, as soon as you hit the button, it starts to speak all text that's there around you. So, if you have a stack of envelopes that you need to get through, you could use the Instant Text feature to read that. If you're at a supermarket and you want to read the products on the shelf, you could use the Instant Text feature to read that. Or if you are at the train station and you want to read what's on the display or if you want to read the timetable, you could use the Instant Text feature to do that.

It works completely offline, works in thirty different languages out of the box and can also work with digital screens as well. If you are at the ATM or coffee machine, one of those really annoying touch screens, they're everywhere and I hate them because they just are the least accessible things on planet Earth. I don't know why people just stop making buttons anymore. You could use the Instant Text feature to read all of that stuff.

I think after that, a very, very close second is the Scan Text feature, which is incredibly versatile. So, you could import a PDF, you could import any file format that is popular today, like a Word file or even an EPUB file. If it contains handwritten text, the app is also going to be able to read that as well. I think reading text, both in terms of short pieces of text offline, or if you want to scan documents or import documents into the Envision app, it's incredible. And those are the first two buttons that you land on as soon as you open the Envision app and I think that's what I would suggest people would love.

**Ricky Enger:** Yeah, I would agree, actually. I have to say I love the Handwriting feature because as cool as it is to have the text read anything around you, whether it be a document, a piece of junk mail or the back of a cereal box, one thing that was severely lacking in any other technology was that ability to read a Christmas card or if somebody leaves you a note and kind of, oh, right, you can't see that note. Now, how are you going to read it? And so, the ability to read handwriting, and I will say even not great handwriting, my mom has apparently the worst handwriting. And my son would be reading cards and he sort of loved getting cards from his grandmother, but also dreaded it a little bit because it was so hard to read, Envision did a great job with that. So yeah, love the feature.

Let's talk then a little bit about the glasses. We have made reference to them on a few occasions throughout this and now we'll dive right into that. I know a lot of people when they think glasses, if they have low vision, they're thinking, "Oh, so this will make my existing vision better." And if they're totally blind, they're like, "Well, there's nothing there, so glasses are not going to help me." But the good news is that when we say glasses in reference to Envision glasses, they aren't actually helping with your existing vision, they're providing a different way to see things around you. As you mentioned, it has a camera, you can wear it, the camera will be above your right eye and you can even have it in lenses, which is really, really cool. So, with this kind of hands-free approach as opposed to what you can do on your phone, are there places where that really shines? Are there places where having the glasses on your face, as opposed to holding the phone, just makes more sense?

**Karthik Kannan:** It makes a lot of sense. In pretty much every aspect you can think of, everything you can think of doing with a phone, the Envision app gets infinitely better when you actually use the glasses. The main reason, it's a hands-free experience, so you don't have to essentially hold your phone in one hand and say, for example, the document in the other hand and so on. It's essentially having your hands free to do more things.

And two, it becomes a lot easier to point with the glasses than with your phone. Especially if you're someone who's not very tech-savvy, you're someone who has come to smartphones a little later in life, you probably feel more comfortable pointing with your head than with your hand. That's what we've noticed. That's the main reason people tend to prefer the glasses, they have their hands free and then they can go ahead and look better with the glasses.

Now, to give you a very specific example. When you're trying to go ahead and scan a document with the Envision app or with the smartphone in general, you usually have to put the document somewhere on the table and then find a way to align the phone with the document itself and then take a picture or have the app take a picture for you or whatever. Now, that's one experience.

With the glasses, the way it works is you could just hold the document pretty much in front of your nose a few inches away and the glasses will guide you on how to move the document, which is a much more natural form of doing things than having to move the phone around. And once you move the document around a little bit, the glasses will basically capture a picture for you automatically when all the edges of a document are visible and then they can start reading it out. That's one thing.

With the glasses, you can also make video calls to a friend or a family member without having to take out your phone. Again, if you're making a video call, usually on FaceTime or WhatsApp, you have to point. And if you want to get some help to pick an outfit or you want to get some help with cooking, then you'll have to point the phone around and move around a little bit and so on.

With the glasses, you have your hands free, so you can maybe hold the document or hold the object that you want to get more information about more easily. The glasses also have a much wider-angle camera than your regular smartphone camera, so it's able to capture more information for the video call than what you could do with the phone.

**Ricky Enger:** Yes, that makes sense. I remember being very excited when I got my pair of Envision glasses and I was wearing them and just moving my head around and being astonished at just how much print there is in the world. If you're sighted, you probably overlook it because it's everywhere and you only look at the things that you want to look at.

But I can remember sitting somewhere and noticing that there's a help wanted sign, this restaurant is hiring that I'm sitting in, and I've just read the menu with the Envision glasses. All of these little things that are a part of daily life, it was a lot of fun and not to mention incredibly useful to have the glasses to do those kinds of things.

So, I know you can't get too deep into what is going to happen in the future, you’ll talk about that when it's ready. But is there a direction that you're going with Envision, some things that you're looking at that you think are going to be helpful or even maybe you're predicting some improvements in existing things that are going to help the blind and low vision community with either the glasses or the app?

**Karthik Kannan:** Yeah, definitely there are some directions that we're very excited about. I think what we've seen over the last year is this whole explosion of new technologies happening, where the idea is that you don't essentially do things like pressing a button or having to be very specific about the things you want to do. You could talk to a computer, in a very natural language and the AI would also respond back to you. And we have been working on that kind of interface for quite some time, where all that you have to do is have a camera or feed open and you can ask questions of the world around you and the AI would be able to respond to you.

In fact, we implemented this in the Scan Text feature of the Envision glasses. It's called Ask Envision, where you can scan a document and earlier you would be able to scroll through the document with the touch screen, but now you can essentially ask questions of the document you just scanned and the AI would actually be able to fetch the answers for you.

For example, you could be sitting at the restaurant and you have the menu card in front of you. Now, earlier you scanned the menu card with the Envision glasses and let's say you're interested in the desserts, which usually is at the very bottom of a menu card. And so, you have to scroll through the appetizers, the main course and then get to the dessert and then see what's there. Now, with the Ask Envision feature, you could just hit a button and you could say, "Hey, can you tell me what dessert options are there right now?" And the glasses would actually go ahead and scan the menu for you, understand your question, understand the menu, and then go ahead and speak out the dessert options that are there in the menu card.

What we're also working on is a version where you could just take a picture of anything and ask it a question. So, you could take a picture of a scene outside your house and then you could ask it questions like, "What does the sky look like?" Or you could take a picture of the interior of a room and ask it questions like, "How many people are there in front of me?" Or "What is the color of the floor?" So that is something that we are working on right now.

Of course, the other direction that we've been working on for quite some time is making more and more features of the glasses without having to need an internet connection always. So that is the second major direction that we're heading as a product.

**Ricky Enger:** Yeah, I feel like Envision has really grown to encompass a lot of different aspects of daily life when you're living with blindness or low vision, and you want to access all of this information around you so that you can make decisions about what to do next. "Is this the right train number that I'm getting on? I always go back to junk mail, maybe that's because that's the bane of my existence, but "Is this a thing I can throw away or is it important?"

So, all of these things are wonderful and we so appreciate the work that you all have put into making this possible. For people who want to know a little more, they want to learn more about the glasses, or they want to learn more about Envision in general, what's the best place to find out what is going on and to get more information about what you're doing?

**Karthik Kannan:** I think the best place to find out more information about the Envision glasses is on our website. You can go to letsenvision.com, that's L-E-T-S-E-N-V-I-S-I-O-N.com. So that's where you can go to find more information about the Envision glasses. You can also request a free demo of the Envision glasses. So, depending on whether you want to do it online or offline, we'll be able to go ahead and give you a demo of the glasses, where you can just ask any questions you have about them and so on. And if you are happy with it, you can also purchase the glasses on our website.You can try the glasses at home for 30 days and then if you like it, you can keep it, but on the odd chance that you don't like it, you can go ahead and return it as well.

We offer every customer of Envision glasses a free one-on-one onboarding, so we help you set up the glasses, set up the app. If you're someone who's not generally very comfortable with technology, you could still go ahead and get a pair of glasses and then we will help you set it up and make it good to go. So that's where you can find more information about us.

**Ricky Enger:** Perfect. Well, I have so enjoyed this conversation, Karthik. I'm so glad you could take a little time and just tell us about what's going on with Envision and what we can look forward to next. Thank you again for stopping by.

**Karthik Kannan:** Thank you so much, Ricky.

**Ricky Enger:** Got something to say? Share your thoughts about this episode of Hadley Presents or make suggestions for future episodes. We'd love to hear from you. Send us an email at podcast@hadley.edu. That's P-O-D-C-A-S-T @hadley.edu. Or leave us a message at 847-784-2870. Thanks for listening.