Lazarus - PIR Ep. 412

26 MIN · YOUTUBE · HTTPS://WWW.YOUTUBE.COM/WATCH?V=XYQ2RREZPUW
https://www.youtube.com/watch?v=Xyq2rREZPUw

SUMMARY

Lazarus, an advanced document understanding company, utilizes large language models to extract and summarize information from complex documents, particularly in the healthcare sector. The founders discuss their pivot from a previous focus on AI for skin cancer diagnosis to creating a tool that can efficiently handle various document types, emphasizing the importance of accuracy and low hallucination rates in their outputs.

- *Lazarus specializes in document understanding using large language models, enabling users to query and summarize extensive documents.*
- *The company's model, Recai, is designed specifically for document processing, boasting a sub-1% hallucination rate, significantly lower than competitors like ChatGPT.*
- *The founders pivoted from a failed AI project in healthcare to document understanding, discovering their tool's broader applicability.*
- *The name "Lazarus" symbolizes bringing documents back to life, reflecting the company's revival and mission.*
- *Upcoming features include a medical record summarizer and semantic search capabilities across enterprise documents.*
- *Integration is straightforward via API, and pricing is based on pages processed with options for cloud deployment.*
- *The founders encourage potential users to test the product by attempting to "break" it, fostering a hands-on approach to demonstrate its capabilities.*

Speaker 1: Hello and welcome to Profiles in Risk Live. This is your host, Tony Canyons. And we are recording. When I say live, we're recording in person, but we're not transmitting this life. We are recording at INSURTECH Hartford Symposium 2023 at the Mohegan Sun Casino, an hour from Hartford. And today I have with me the guys from Lazarus. So I'll let you guys introduce yourselves 21st.

Speaker 2: Nice to see you again. So this is Ariel, he's our CEO. He's the brains of this.

Speaker 1: Awesome. So you're the look, you're, you're the looks.

Speaker 2: You're no less than the cool shoe I cannot match to your looks.

Speaker 1: So I'm trying, I'm, I, I, I, I'm gonna say. So what's your first, Ariel? Ariel. I do have to say, if you need to sell anything in a very conservative industry, hiring a long haired foreign hippie, in my opinion, it's one of the best strategies there are.

Speaker 2: You see, it pays off. I mean, it took me a while.

Speaker 1: By the way, for the listeners, we are recording with two cameras. We have my laptop over here hooked up to this microphone. And then we have my, my phone on a gimbal which I'm hand holding. I'm not sure how we will mix both videos, but that's why you have a multicam setup. Just, just so you guys know what this looks like. Okay, awesome. So, well, first of all, thank you so much for scheduling this. Normally I use a calendar to schedule podcast recordings and I don't accept any meetings at conferences because I forget.

Speaker 1: So not only did they go out of the way to schedule this, but they reminded me every half hour of the hour for the last two days. And so we managed to make this happen. So we always give the guests a chance to give the elevator pitch. What, what is Lazarus?

Speaker 3: Yeah, I can give you the rundown. So Lazarus is an advanced document understanding company. So we use large language models to help companies do document understanding. That's a whole lot of buzzwords. But what that effectively means is you can take any document, ask any question on top of that document or set of questions and get out answers. So you can turn 200 page medical record PDF into a five sentence summary with, you know, every condition, every medication, and every time the patient stubbed their toe in the, or listed out in the output or really however, whatever you

Speaker 2: want the output to be.

Speaker 1: Okay. And we, we are recording this on May 3rd of year one of our large ChatGPT. So, so, so given the tools that are now available to everybody, why can't I just grab, you know, that HIPAA protected document and throw it into ChatGPT API to my, to my guide wire or whatever system I'm using?

Speaker 3: You've answered half of that question. So because it's HIPAA protected, tools like GPT4, GPT 3.5 really cannot and should

Speaker 2: not handle that kind of data.

Speaker 3: Um, but the main stuff, number one, if you're working with like a 500 page PDF, it's way too big to fit in the context window for GPT4.

Speaker 1: Okay. Which is the largest, even the paid version.

Speaker 3: Yes.

Speaker 1: Okay.

Speaker 3: That's 32,000 tokens. If you can fit 500 pages, 32,000 tokens, 499 of those pages are blank.

Speaker 1: Okay.

Speaker 3: But the other component of it is so GPT4 and models like it are really good at understanding text. Kind of one token after the other, not very good. Understand text in two dimensions. So unfortunately, documents are in two dimensions. So you have tokens AK words and where they're located. And rekindle is built to understand that. RECAI is the name of our large language model.

Speaker 1: Okay.

Speaker 3: And so because of that, it works really well on documents. But given that it's designed exclusively for documents and pulling data out of documents, it cannot write a haiku about a medical record. So if you need to write a haiku about a medical record, other models might be better.

Speaker 1: I have to say, ChatGPT did a fantastic job when I, when I, when I asked it about writing a haiku about insurance careers. So I will include that here, here in, you know, when I want to edit this. Okay, so no haikus of the medical records?

Speaker 2: Nope, that's right.

Speaker 1: Is that coming in the next version?

Speaker 2: Might be.

Speaker 1: Okay, perfect if there are enough requests for haikus about medical records. Okay, so I can't write a haiku, but it can help my system much more easily digest that complicated record and HIPAA protected record. Yeah.

Speaker 3: And it'll tell you the truth. So it's got a sub 1% hallucination rate. It just doesn't make sense.

Speaker 1: Okay, okay, okay. So, so hallucination, I love how this, this word, I mean, we've used it like when it comes to people going crazy or using drugs or whatever. But I love how it's come into our space in the last like six months. So for the listeners who haven't been on top of the chat stuff, what is hallucination?

Speaker 3: Yeah, so language models are, at the end of the day, glorified autocomplete.

Speaker 2: Right.

Speaker 3: They identify what the next word In a sentence is supposed to be generated. And they do that when they get big enough. They can understand structures and they can understand patterns better so they can seem more human. Like the way that ChatGPT was trained and the way that our model was trained was through this process called reinforcement learning with human feedback. And what that just boils down to is a human said good job, bad job on all the outputs for the model up until the point that the model kind of was selected or self selected and was self trained into.

Speaker 3: In the case of ChatGPT, acting like a person chatting with you, answering instructions, filling in information, things like that. The problem with that is if you tune it to understand instructions and to act human and to give human like responses. So if your BART barrier for entry is human level responses that can trick somebody else, it starts to do things like lie. That's what hallucination is. It's not actively lying to you on purpose. It's just making stuff up that doesn't actually exist because of the way Recai was trained.

Speaker 3: It can't do a lot of those more advanced things like make a haiku, but it hallucinates at a sub 1% rate compared to GPT4.7% and GPT 3.5. I think 70%.

Speaker 2: Yeah. Wait until they're going to start tripping.

Speaker 1: So, so, so rick had those sub 1%. Sub 1%.

Speaker 2: Yep.

Speaker 1: So less than 1/7th the house hallucinations that off that a chatgpt does.

Speaker 2: Well, that's what makes it commercially viable. Yeah. So we're talking about like real business use.

Speaker 1: Okay.

Speaker 2: So solving a problem of extracting information from the documents.

Speaker 1: Okay.

Speaker 2: We really want it to be factual. We don't want it to be, you know, hallucinating, you know, this or that when you're making a business call.

Speaker 1: Okay, that's super interesting. How did you come up with the idea of building this thing?

Speaker 3: Okay, so we had to actually. So we used to be a completely different company. We used to make AI models for skin cancer, clinical decision support. Basically helping doctors figure out like, hey, is this molecular?

Speaker 2: Believe it or not,

Speaker 3: that entire industry got destroyed by regulation. And so when we had like three months of Runway left, we had to figure out a product that would keep us from dying. And what we did is we took this old dusty thing that we had off the shelf, which we used to like clean up biopsy reports so that we can label our data. And it was designed to work with any document type, regardless of structure because biopsy reports are not standardized.

Speaker 3: I suspect it's The a pathologist wanting to make my life specifically harder. No, but they.

Speaker 1: They suck.

Speaker 3: So we had to make a tool that could work on all pathology reports. What we realized is later on, when we showed it to other people and they use it on their documents. Hey, it doesn't just work with every pathology report. It works with every document that we throw at it. And so we went down a little bit deeper. We replaced our main model, which was like a Bert Ish type model, with a large language model.

Speaker 3: And so In April of 2022, we started training it. It was finished training in November, like mid November, like two weeks before ChatGPT launched. And for two whole weeks we had to explain to people what the hell

Speaker 2: a large language model was.

Speaker 3: And ChatGPT launched. We didn't have to explain shit.

Speaker 1: That is amazing timing. Yes. And do you take credit for the timing your crystal ball or.

Speaker 3: We're business geniuses.

Speaker 1: Incredible luck.

Speaker 3: Incredible luck.

Speaker 1: Okay, well, it's luck.

Speaker 3: Plus the technology got to a point where we could implement what we needed to implement, and we just so happened to do it at about the same time that somebody else was implementing something else with similar technology.

Speaker 1: Okay, there is a word in evolution for this, for two things evolving. Parallel. It might even be parallel evolution, I don't remember, but something like that. Yeah, but basically two animals or whatever evolving the same advantages geographically separate from each other. So basically the same idea. Very, very, very interesting. Okay. Okay.

Speaker 2: So.

Speaker 1: So Lazarus, if I'm not mistaken, it is a biblical story of a man that has died and Jesus resurrects him. So why did he choose the name Lazarus?

Speaker 3: Do you want the marketing answer or the real answer?

Speaker 1: Both.

Speaker 3: Start with the marketing answer.

Speaker 1: You can start with the marketing answer. Exactly.

Speaker 2: We just simply bring documents to life.

Speaker 1: Okay. Okay. And what's the real answer?

Speaker 3: Okay, so when we started the company, we got a investor check before we had a company name. My co founder's name is Anish.

Speaker 2: His last name is Patel.

Speaker 3: Patel Inc. Didn't sound super cool. It was Aurovink. My last name didn't sound super cool, but the Latinization of my last name, Lazarus. So we picked that name and then we sure as hell earned it when we came back from the dead as a company.

Speaker 1: Oh my God, this is fantastic. This is absolutely fantastic. Okay. Okay, perfect. Okay, fantastic. Alex. So we met last year at itc. Okay. And at that time you were. I might cut this part out. At that time, you. You were a corporate. At that time you were with a

Speaker 2: very large corporate escort.

Speaker 1: Okay, let's make yes.

Speaker 3: You were in Vegas, right?

Speaker 1: So. So at that time you were at a. At a very large consulting company.

Speaker 3: Yep.

Speaker 1: So you had a dream job, like a great company.

Speaker 2: I started really doing what I wanted to do all along, which is getting back into technology and I was running the innovation team for financial services. And man, we've been trying to solve this problem for years.

Speaker 1: So why leave a job like that to join a startup?

Speaker 2: Kind of jumped the gun there, right?

Speaker 1: He has been emailing his boss asking for forgiveness for the last three months.

Speaker 3: Yes. He's only been gone for us.

Speaker 2: I'm still on, you know, sort of probation period. So. So when I left ITC after, you know, you and I met because I wanted to, you know, one up on you, on my socks.

Speaker 1: Oh, okay.

Speaker 3: Okay.

Speaker 2: That was the way we interpreted the first time. So I, I got reached out by Keith, our growth leader. And Keith said, listen, we're trying to solve this problem and we think we've cracked the code on it. And I said, you know, if you really think you do that, I'm absolutely interested. So, you know, give me access. I want to test this thing out. And I think it took about two to three weeks. And then I got access in early December and I started playing with it and I'm like, this thing is sort of unbelievable that it works.

Speaker 2: So I'm like, let's try to break it. Which, you know, I was kind of trying hard to break it. And, you know, I was getting on a phone with Ariel and Keith. You know, probably every two to three weeks I just kind of report back and everything. And then one day, you know, I just throw this thing, a document in Romanian. So that's originally where I'm from. And I'm like, this is not gonna work. And then I ask a question in Romania, I get the response back accurately and I'm like, what is this thing?

Speaker 3: Like, how can he do that?

Speaker 2: Has.

Speaker 1: Was it trained in Romanian at all?

Speaker 2: Not my knowledge.

Speaker 1: That is insane.

Speaker 3: The fine tuned one.

Speaker 2: So there's a.

Speaker 3: So one of the moss is 100, has a 40 languages and the other one is 140. 140 language model. Definitely knows Iranian, but that's not the

Speaker 1: one that Alex had access to.

Speaker 3: Nope.

Speaker 1: That. Okay, that is very scary. As in, like, I really hope that we make it to. To ITC Vegas for one last time before we inevitably get enslaved or. Or annihilate as a species. But by generating reality.

Speaker 2: You just have to be kind to us, Tony. That's all it does.

Speaker 1: Okay, so. So you threw in the dog with Romanian expecting it not to work. Ask the question in Romania. It worked.

Speaker 2: It worked. Then I'm like, all right, so I gotta talk to these guys, like something's going on here. Then we kept talking and then I'm like, well, there's this, this thing has so many different possibilities to like redesign, you know, how we do, you know, our own sort of client related services, but you know, the industry at large. So we've been just exploring ideas for a while and then, you know, I think Ariel one, one day telling me, you should come and work with us.

Speaker 2: You really should. So I'm like, you know what? I think this is one of the things that I really wanted to do. And what could be better than, you know, know, like taking all the sort of industry experience I had and just doing, you know, what I was doing at PwC, trying to solve this problem, and now I have the opportunity to do it. So I said yes.

Speaker 1: So there's a great lesson here for any other insurance founder. So there are basically two paths to finding fantastic people. Path number one is you can blow their minds, but with the product being incredible, or you can hire Tony to find you the right people. Path number two is a lot easier. Path number one especially the product isn't quite ready. So that's, that's a really cool story.

Speaker 2: That was a very subtle plug for you.

Speaker 1: Very subtle, I know, but by the way, so Alex and I have been kind of doing the same conferences for the last two weeks. So we were at ITC Latam last week in Miami, the first inaugural item. And I've never been very impressed by anybody in sales. In the middle of the closing day party, Alex is, is giving a product demo on his phone in the middle of the Havana nights party. So like, if that's not hustling.

Speaker 1: Yeah, I don't know what is.

Speaker 2: Plus, you got to do what you got to do, right?

Speaker 1: For real?

Speaker 3: For real.

Speaker 1: Yeah. See that's the advantage of, of hiring long haired foreign hippies for, for sales jobs. Okay, so we just invented a new profile there, right?

Speaker 3: Am I like a second, second level removed? My parents are both immigrants.

Speaker 1: See, but since, since you are, since you are the, I'm guessing you're an engineer by trade of some sort. Computer science or.

Speaker 3: It's actually neuroscience background, but yeah, data science.

Speaker 1: So by the way, for the listeners, normally I'm recording the episode with a 49 inch screen where I have their LinkedIn pulled up, the company website pulled up and like I can really dig into their background. I Have nothing right now. Okay, so, so, so, so, so you qualify more as like the crazy mad scientist long hair more. More like like the, the, the guy from Back to the Future. The Christopher Lloyd Doc Brown. Doc Brown, exactly.

Speaker 1: More that than the hippie long haired foreign salesperson. Which. So we're recording this at Insurtech hard for symposium. I will get it out fairly quickly probably early next week. So. So the listeners will be watching this, let's say May 5th, May 6th. What other conferences will you guys be at? Where, where can people go to to to get a, a, a demo on. On the phone or a demo at the booth?

Speaker 2: So I'll do the demo on the phone when we go out for drinks.

Speaker 1: Okay, perfect.

Speaker 2: Off to the comforts. But, but you can find us at InsurTech Insights.

Speaker 1: Okay. InsurTech Insights in New York at the Javits Center. And I believe that's June 7th and 8th.

Speaker 2: June 7th and 8. And I'm working on a interesting approach, a contest for people to try to break it.

Speaker 1: Awesome. So. So are you ready to officially announce this? Rules, price, all that or we have.

Speaker 2: We're still deliberating on the prizes. So you know, like maybe you know, lunch with, with me and Ariel or is that valuable? Or instead of cards, I don't know. We're still thinking popcorn maybe it's on the list.

Speaker 1: Okay.

Speaker 2: Okay.

Speaker 1: Okay, perfect. How long will people have to try to break it?

Speaker 2: So they will get three questions and if they can essentially get the get us to consistently produce the wrong output without any sort of very easy to do what we call prompt engineering. And you're just asking the question a different way. Yeah, they get a winner Price.

Speaker 1: Okay. I love that approach. Love that approach. So in terms of insights, I'm guessing you don't know your booth number yet or you don't have it on hand?

Speaker 2: Not yet. No. I'm, you know, I'm supposed to get, get to my friends at ITI to to give me a boot that's not, you know, at the very end of the, of the conference space. So we'll see if I can convince them to do that.

Speaker 1: Okay, perfect. For those that are not making to the conference, what is the website where they can learn more about Lazarus? Yep. You can just go to lazarusai do lazarus AI.com okay. Lazarus AI.com when this goes live, I will tag you both on on, on on on LinkedIn and I will tag the company also what is important. Okay. What? Actually let's, let's back. Let's back out a little There's a couple questions that, that I forgot. So number one, what is the integration like?

Speaker 3: Oh, we're an API so virtually non existent. You just plug it in wherever you need it.

Speaker 1: So what if we're not API ready?

Speaker 3: You should be, but, and much better

Speaker 1: now that like seven years ago when we started recording the podcast.

Speaker 2: Next question, Tony. What if we don't have money?

Speaker 3: So if you're not API ready, you can literally just like pull in something like UiPath or another one of those RPA platforms. Stick it in.

Speaker 1: Okay, perfect. So integration is pretty easy.

Speaker 2: Yeah.

Speaker 1: Okay, what is the pricing model?

Speaker 3: So we charge per page and we have a minimum annual contract value. Okay. And it depends on like which model you're using. So we have like lighter weight models, we have heavier models. It also depends on whether or not you want to run it on our cloud or yours.

Speaker 1: Oh, okay.

Speaker 3: So yeah, okay.

Speaker 1: So there's, there's customization, there's a lot of options.

Speaker 3: Yeah, yeah.

Speaker 1: On how to integrate. Yeah. But, okay, fantastic. What, what is next for Lazarus? What will, what do you expect if you're able to announce it already? What are you expecting we'll be able to do, let's say next year that it can't do yet this year?

Speaker 3: Oh, I can't answer that. Okay, let's take it like week by week. Maybe

Speaker 1: we need to get used to the new rate of advancement that AI is seeing because it truly, it seems that we've hit kind of the elbow of the exponential curve when it comes to how quickly AI is advancing.

Speaker 3: You called me mad scientist. Alex has got some mad science ideas himself.

Speaker 1: Yeah.

Speaker 2: So I mean, I think one of the things that's already started to circulate on LinkedIn that people should expect to see as a product from us is the medical record summarizer, which I would invite all of them to take a look at the actual demo that we have there. It's mind blowing. 50 page medical record summarized in about 42 seconds to run. And it organized information by domain. Again, it really, it doesn't matter what language it's in because it's also translating it.

Speaker 2: So it could really be used, you know, across many different puzzle pieces in terms of like travel services, for example, you know, get, get these types of documents in many different languages, many different formats and they have to make sense of them all. So, you know, that's one product that's, that's really pretty much introduction already with the clients. So you know, we're going to start making a big, a bit more, you know, public noise about it.

Speaker 2: Then, you know, on the roadmap, we'll have semantic search at scale across all the documents in an enterprise. So you can, you know, it's not perfect to say you don't need to digitize it, but that's. That's really truly what it means. You can ask questions of, you know, any of your old documents, and you can, you know, compare, you know, sections. You can ask questions like, all right, what is this clause, like, you know, for these contracts submitted between such and such a date?

Speaker 2: How does it compare? And you get, you know, a real good answer to work with.

Speaker 1: So we talked mostly about health, which is weird because I'm not a health guy, but so this can be used in pnc, it can be used in life. Oh, like just off my head, like, is this likely covered? Or what are the changes between the old policy and the new policy? Or. Right, like the life insurance policy that we wrote 55 years ago that is now actually being collected on where, like, nobody alive remembers how, like, what was

Speaker 2: this around in those companies to, like, you know, even have an idea of what was going on back then? Right. So the. These old systems, the mainframes, and they have a very small percentage of the information that's in the actual contract. And so, you know, to really go ahead and try to bring all that information back into your systems. And nobody knows if there's an ROI there. So now it's possible you just ask questions of it.

Speaker 1: Okay, so you said you can run it.

Speaker 2: You can.

Speaker 1: We can run it on your cloud. We can run on our cloud. So will it run on our AS400, though, that we've been putting duct tape on for the last 26 years.

Speaker 3: Maybe not.

Speaker 2: It depends how long you live. It might take you a couple lifetimes to get it response back.

Speaker 1: So for. For the listeners at, at really traditional companies, if you still have Cobalt programmers who can program that thing and Fortran, then maybe it's a maybe.

Speaker 3: If you can figure out if you can find somebody who is still alive at your organization who knows how to do it, okay, then you could. It'll work, but it'll be slow.

Speaker 1: It'll be slow. Okay. Okay, so. So we would ask it today and like, come back next Wednesday kind of thing. Yeah.

Speaker 2: And it's still. And it's still printing.

Speaker 3: It's still running.

Speaker 1: For the listeners, this is a lot of fun. I. I'm gonna do more like recording at the conferences. Thank you, Alex. This is a wonderful idea. It's something that, that Nick Lamparelli did years ago at InsurTech Hartford Symposium at the first institute symposium about three years ago, I think three or four years ago because we do go way back with, with IHS and at that time like Nick's significantly more formal than I am. It wasn't this, this fun but, but, but since then because I've been the main one running the podcast and, and, and I love the conferences but never think of doing something recording like this in person.

Speaker 1: I want to do more of this at the conferences. Thank you so much. So, so for the listeners, if you, if you fall in love with it with, with this format, you can thank Alex and yeah, my pleasure to have you guys. Thank you so so much for, for coming on the show. Look forward to see how things advance normally. What what I tell people is we should have you back in a year to, to see how things have advanced.

Speaker 1: Given the, the, the the how quickly things move in in AI. I, I feel like we should have it on the schedule to, to record an update three weeks from now. It has been a lot of fun. And also I really have to say, say while the stats say that our listeners are majority carrier people and broker people, I do imagine that we have a significant amount of insurtechs and especially future founders that listen to this.

Speaker 1: And I love how candid you were, how honest you were with the company was dying, the old model didn't work and the pivot to like this one tool we had built for this one very specific, this very specific purpose, all of a sudden we realized that that is the real valuable part of the company.

Speaker 2: And that philosophy honestly transpires into how we also talk about the product and we get people to like try it on. Like you can't really buy a car without like trying it, right? So you know, we're inviting them come in, try to break it, play with it. You know, throw it your real life examples at it.

Speaker 1: So at icloud time there was one company, I remember who it was that had a big thing that, that said the world's first free insuretech and basically like I interviewed them like the first hundred grand that you pass to the platform are free and then after that it's a percentage. But if you never get to 100 grand, you never pay anything. So you guys should have a big sign that we probably don't have time to make it in time for insights unless you hand make it.

Speaker 1: But a big sign that says something about breaking come break our product. Come break Lazarus. Something like that. You have T shirts that say come break Lazarus or. I don't know. There's a lot of really fun stuff here.

Speaker 3: And here's the fun part. If you break it now, we know what doesn't work, and we can fix it so it'll never break again that way.

Speaker 2: Which.

Speaker 1: Yeah, no, no, that makes sense. Makes perfect sense. Thank you both for your time.

Speaker 2: Tony, thank you.

Speaker 3: Thanks, Tony.

Speaker 1: Awesome.