ZETTA BYTES LIVE – Karim Lakhani of Harvard Business School – December 15, 2020

Mark Gorenberg, Zetta Venture Partners: Welcome. I'm your host Mark Gorenberg. I'm delighted to welcome you Zetta Bytes Live. Zetta Venture Partners started in 2013 where the first venture fund focused purely on AI B2B business model seed stage. Today we've made 40 investments in enabling AI platforms and applied AI. We've had six exits to date and one IPO, Domo. We were the very first investor market cap. Now at 1.6 billion and five acquisitions, two of them by Google when Bebop that became Google Cloud and also Kaggle, uh, and, uh, the largest data science community. Now about 4 million data scientists and Anthony Goldbloom was our last set of bites lives speaker hosted by my partner, Ash Fontana. So is it a bites live, does deep dives into state-of-the-art and AI platforms and enterprise applications with top industry leaders in AI, we believe technology powered by data will reshape business, as we know it, and we've set up a program to tell stories about that and we'll feature deep dives into the state of the art in AI platforms and showcase some successful applications of AI as we're doing with our guest today is professor Karim Lakhani from Harvard Business School.

Professor Lakhani specializes in technology management, innovation, digital transformation, and AI. He's the founder and co-director of the Laboratory for Innovation Sciences at Harvard, known as LISH. Co-founder of the Harvard Business School' Digital Initiative and co-founder and coach of Harvard Business Analytics Program. He's also taught extensively in Harvard Business School's MBA doctoral and online programs. And he co-chairs the HBS executive program on competing with big data and business analytics, which might be particularly interesting to many of you out there. He's published over 150 peer-reviewed papers now and in leading journals and is co-author of two books on open and distributed innovation models and on perspectives on free and open-source software. He's also coauthor of *Competing in the Age of AI*, which just came out this past year, a book published by Harvard Business Review Press.

His research has been featured well, frankly, almost everywhere: Business Week, Boston Globe, The Economist, Fast Company, MarketWatch, New York Times, NPR, Wall Street Journal, Washington Post, Wired, and many, many others. I'm particularly delighted because Kareem got his PhD at MIT and he's also on the board by the way of a number of private companies. And in full disclosure, we're on a board together at an MIT/Harvard spinout called VideaHealth, which uses AI to improve dentistry. So we invite you to submit questions using the Q&A button on zoom, which we'll get to in the latter part of this hour. And last but not least, I want to let you know, too, those of you who are here today, we'll be emailing you a survey. And those of you complete the survey. We'll be sending you a complimentary copy of his book competing in the age of AI second time, um, as a thank you for being here with us today. So Karim welcome to Zetta Bytes Live and, and thank you for being here today.

Karim Lakhani, Harvard Business School: Hey Mark. Thanks so much for having me here. It's a pleasure. Looking forward to our, uh, virtual fireside conversation,

MG: Our virtual fireside conversation, that's the way of the world today. To start off, whenever we talk with new companies, my partner Jocelyn always wants to hear the origin story. So in her tradition, if you don't mind, can we start with your journey? How did you grow up and when did you know you wanted to be a professor educator leader in data and analytics?

KL: And management? It's all a big accident. I always scratch myself and say, am I really here at this office at Harvard and so forth? I didn't have any intention whatsoever to be an academic, um, by that was, uh, uh, prior to me deciding to do a PhD, I told my advisor, like, I don't want to be writing obscure papers and it's your journal. And that's what I, these days, uh, that that's a better, better record Lumia. Um, so, uh, you know, I actually was born in Pakistan and then, uh, when I was born, our family moved to Canada and, um, uh, to do my high school, there went to McMaster University in a program called electrical engineering and management. Cause I always wanted to be at the interface of technology and business, which I had. This is all last century.

Speaker 2: (04:56)
And then I joined general electric, uh, in one of their, uh, their programs. So I joined the technical leadership program, which at that time in Canada was actually basically assigned to the medical system business, uh, the healthcare business, uh, and put me through a whole bunch of rotations in the U S and Canada, around new product development and marketing sales and so on and so forth. And so it was a fantastic lesson, spent four years there. And that's where I started seeing that the world was ending. My, you know, my first encounter with mosaic, the browser, wasn't at a radiology conference in Vancouver, I was in Vancouver and this radiologist who is a radiologist, tend to be the most devious of all the doctors have all the fancy toys. And so you showed me mosaic. I'm like, what I got I'm used to go for and, you know, texts.

Speaker 2: (05:45)
And this is a totally different view of the internet. And I just clearly remember seeing that and going, wow, the world is changing. And then through my experience at GE, you know, again, hanging out with the radiologist, I got to see the start of open source and it never made any sense because GE if you remember, jeez tag on used to be like you bring grits in select GE does. And all of a sudden these random people in the world that were bringing good things to life, it didn't make any sense. And I probably forgot about it, but somehow in the back of my mind was there. And then, uh, I went to MIT for a master's degree in technology and policy course. Um, and by, by sheer accident, I took a class at the so-and-so's law school, um, with Eric [inaudible] and he, uh, he said, uh, you know, you'd write a paper.

Speaker 2: (06:30)
And he was talking about, about user driven innovation and a hundred users going to be, and I wrote a paper saying, Oh yeah, it's not as simple than users is actually communities. And look at this open source example. Um, so at the time, you know, this was like my first year there and I was working that summer DCG, you know, some, some TA you know, graded the paper, gave me the name of the course. And then I think Eric then reads only the papers that get A's lazy professors. And so he calls me up, he goes, Hey, I read your paper. You know, I already got an

a from you, why are you calling you now? Because I want you to come work with me. I'm like, no, Eric, you can't afford me right now. I'm a BCG. So he was a persistent, and this is being like Eric was persistent.

Speaker 2: (07:11)
And then, uh, that fall, you might be able to his house super nice guy. And he said, you know, you gotta deal with thesis with me. Your master's is with me. I'm like, I'm doing computational biologists now, because I know this is a really important topic. So he finally convinced my wife, you know, uh, to say, I should switch over. So I switched with him. Uh, and I, that was like the first time I got the research bug, I got this taste of, Oh, I could go deep into something and really understand it and, and do a sort of, you know, forcibly research, like describe the phenomenon, understand some causality, explain it. It makes sense of it. Um, and Eric, wasn't a mentor. He sort of helped me do that. And so I did this study way back then in front of master's thesis about the most boring part of open source, which was support and support at that time, I'm using it, I'll use that. And Apache was being support and using that. And our thesis was if the most boring parts of open source are being taken care of, then for surely that software itself is actually pretty robust. So I did that. Um, and, um, and so,

Speaker 3: (08:13)
Yeah, that was fun. And then I, you know, I went to work at BCG, uh, and then Eric was again, very persistent and he would call me every three months. How's it going? How's it going? We've got a paper to write on your thesis, got a paper, write a thesis. So we wrote the paper while I was working at BCG and I finally got sick of consulting and he said, okay, she'll come back to school. I came back and did my PhD with him. And there is where it got interesting because I, I dove into open source, but I started to open my eyes about everything else that was happening around open innovation and did a whole dissertation on both open, open communities, but also open contests. Um, and that led to a job at HBS. Uh, and I started, you know, so it was an easy, easy transfer, just one subway, stop away from where I was living.

Speaker 3: (08:58)
Uh, uh, and then, uh, started to come here, started to teacher to research. And, uh, that's where things got interesting because HBS, you know, is like, okay, you gotta keep publishing in the top academic journals, but you also gotta talk to managers. You gotta be relevant to practice. So we write cases and we, you know, write HBRS and so forth and did some exec ed. And in one of my exec ed sessions, I was talking about all of these sort of distributed models for innovation. And this guy, Jeff Davis, uh, sort of comes to me because, and like, Oh, I'm from NASA. I'd love to hear more about blah gave me his card. Uh, and I was busy assistant professor. I said, director at NASA, okay, whatever, put the card away I got and persisted, Jeff kept calling me. He goes, I want you to come to Houston.

Speaker 3: (09:47)
I'll give you a tour of at JSC, but we really don't need to hear about what you guys are doing there. So then I looked at him closely. He goes, he's not, he's the director of the directorate. He runs the director of this place, life sciences. I'm like, okay. So he's at medical, he's a physician

she's medical officer for NASA and has a thousand person team that keeps astronauts alive in space. I'm like, okay, this guy cares about my work. I better go talk to them. And so, so that started this relationship, or we started to do open innovation projects for, uh, for, for, for NASA. And that's where things got so interesting because it was, uh, it was Jeff. It was, uh, Jason, Cruzan another partner at NASA and talent, Tom [inaudible], who was, uh, in the office where science technology policy in the white house.

Speaker 3: (10:35)
And so we started to run these experiments with NASA and we were getting these very high performing solutions going to top coder. This was the precursor to Kaggle. Uh, and while I'm running experiments and understand, you know, the mechanics of incentives and information and teaming, we're creating all this great code. Uh, and all of these were early machine learning prototypes that we kept seeing. And they were blowing pot past all of the solutions that were out there by NASA. So, so that's got me exposed to sort of, uh, the start of AI, the early days of, you know, machine learning stuff, uh, and code coming through. Uh, and so I drove diet, did a deep dive into that on the research side. Uh, but then also started writing cases about companies are powered by AI. And so those two things came together, like both the research on how the AI got developed through these contests and communities, but at the same time, uh, understanding how that was changing the role of companies.

Speaker 1: (11:33)
Wow. That's pretty amazing. So aircraft hippo for those in the audiences of God, at MIT who, you know, he has the original ideas on, um, I would say bottom up models, user models. So it's interesting that he had the theory and you were put it into practice.

Speaker 3: (11:46)
Yeah, yeah, it was, it was a great, you remember this and, you know, the thing that was that Eric's stuff, I also saw GE because a lot of our, our functioning novels stuff was actually being developed by doctors. We were good at making the magnet, you know, perfecting the algorithm, but like the, the, the novel utility was actually being invented by physicians in the, in the field. And that was Eric's theory that basically users developed functioning novel innovations manufacturer's three-dimension of merit innovations. Uh, and, and so I said, I saw that, Oh yeah, but this is happening at scale. It was not just, he had like a farming example, you know, some radio example, this and that, uh, industrial side. I said, no, in software, this is becoming the standard, but you know, this is 97, 98. So, you know, I tell my second Valley friends, this open source stuff, and they'd be like, laughing at me. They go, no, no, that's not how you develop software. You know, you need to be an office co located and sitting there and doing that. And so I'm glad it worked out.

Speaker 1: (12:42)
You were probably way ahead of your time, but, um, but, uh, you took these partnerships. So it's interesting this evolution to work with real industry, cause you took these partnerships like NASA, Harvard medical school road, uh Topcoder Linux foundation, and you, 10 years ago, you started the laboratory for innovation and science at Harvard. Uh, so tell us about the lab then

over the last 10 years, the digital innovation you put in place, and, and when you got the aha moment about AI being a big deal.

Speaker 3: (13:16)
Yeah. So, so I think, I think I, so the, um, the idea of the lab, you know, came from NASA, they said, look, you know, I, I wanted to most social science, econ, economics, sociology, psychology, most of us sort of, especially economics, sociology management. Most of us sort of get archival data because our Carvel data, we look at the archival data and then we try to draw some causal inference, portrait econometrics to get it, get it there. Uh, uh, of course the gold standard in any research is a randomized control trial. We're hearing all about RCTs right now with, you know, the, the vaccines that are coming through. Uh, but th they're they haven't been deployed as much in the social sciences as they are the norm and in the natural sciences or in life sciences. Uh, the one exception has been the field of development, economics development economics has really shown.

Speaker 3: (14:05)
And again, this is MIT story, MIT, Harvard story has shown how to do, you know, really careful studies that have massive impact on the wellbeing of people in, in, in, um, in, in poverty and so forth. And so that became like the thing, uh, could we do innovation experiments, uh, with companies? Could we actually get the same type of causality that the, uh, dev e-comm people were getting that for innovation had not been done before? So myself and my colleague, Kevin Boudreaux, good friend, co-author we both had gone to MIT for our, uh, PhDs. Uh, and so we said, okay, let's see if we can run an experiment in NASA. I know top coder and we ran the experiment, it worked great. We got machine learning code. We didn't care about the code. We're like, okay, like, we've got all this nerdy data about, you know, poppy plus a lag how, how much effort they do and so forth.

Speaker 3: (14:53)
And then NASA comes to me and says, okay, let's do more. I'm like, I've got enough data here to keep me busy for three papers. They go, no, no, you don't understand. We want to scale this across NASA. So I said, fine, if you need to do that and you need to support me in a very different way. And so that's where the be established the national tournament lab to say, okay, we're going to teach NASA how to do this. Uh, but also we're going to do with NASA, take their problems and put them on various platforms. But, but we're going to learn how to, how to think about the social science behind contests and how people work and so forth. So that got us going. And then very quickly our colleagues in the medical school said, Oh, if you can do this at NASA, why don't you do it with us as well?

Speaker 3: (15:31)
So we started to do that. And again, we saw the same kind of things, and that's where my paper streams got weird. Like all of a sudden I'm publishing in nature biotechnology, I've published in JAMA oncology, you know, uh, and, and, and, and so on, and like highly technical journals, uh, as I'm also publishing and management, science and range and real economics and organization science and research policy. So it was sort of this, this, this thing about like, Oh, I

can do super technical work with our team, but also we can then write about that as well. And that's where is it? It was through these contests on Topcoder that all this AI stuff started to emerge. And it was just like, wow, like this, this, because we were cheaper, faster, better consistently against elite institutions, right? The best people met at the med school, the best people at the Brode, the best people, uh, at NASA, you know, some kid in Estonia, right? Some kid in Paris, some kid, some policemen in Brazil is outperforming, uh, these guys, right.

Speaker 1: (16:29)
Well, you talked early on. I know one of the things that struck me in the book about your work at Lisha on AI included Dana Farber. So I guess Dana-Farber was one of your, one of your early companies to work on this intersection of healthcare and AI, and particularly around lung cancer. And you talked eloquently in the book about that. And if you want to talk about that.

Speaker 3: (16:48)
Yeah. So look, that's like one of the most proudest papers I've written with our team. And so this was with my colleague, Ava guidance, Ava, Ava is a radiation oncologist at Dana-Farber professor of medical school. And she has now more economics papers. And you're at your average economist because we've collaborated so much on, uh, on projects together. And so as we've done all this machine learning stuff, you know, she said, look, much of the machine learning stuff has been about image recognition. Right. Great. Uh, but it's been sort of diagnostic story. Like, you know, what, what, what is this ha is this cancer or not cancer and so on? And so has done this at scale. Uh, but what about treatment? You know, we need to also apply machine and into treatment. And so she goes, we have data on, you know, how you create treatment plans for lung cancer.

Speaker 3: (17:33)
Uh, and part of that involves basically fully delineating a 3d tumor, multiple 3d tumors and the CT scans. And is it possible for us to run some contests to get code for that? Um, and so we did, you know, we spent a good four months ahead of time just figuring out the problem, how would we measure, you know, effectiveness and so forth and getting the data, getting the data, getting IRB approval. And then we ran it on top coder and then, you know, eight weeks, 60 grand. And all of a sudden our code is better than all commercial software that's out there, but also as good as the average Harvard trained radiation oncologist. Right. And the best part is that the, some of the winners took software from autonomous driving, uh, projects and applied it here, this, this arbitrage, and you'd go like, what does autonomous driving have anything to do with cancer?

Speaker 3: (18:30)
You would never see those two fields on our side, but if you think about it, it goes, of course it has everything to do with that because autonomous driving is all about boundary detection, object, shape, recognition, right. We are the objects with rear shapes showing up and you need to know what's the end of that object and was whereas eroded is right. And so that's exactly the same problem in lung cancer therapy, right? And you gonna have this beam, uh, with radiation, the cancer to, uh, cells, but not everything else. How do you detect those cancer cells? And

then how do you create the right shape to do that is exactly the same problem. And that's what gave us a breakthrough. So it was just like a big aha moment. And then the fact that we got to publish in JAMA oncology was incredible. But also my joke is, you know, the last time I took biology was in 10th grade and I was off and here I am publishing in JAMA oncology on an AI project saying that, you know, like, again, the costs for us to innovate with AI have dropped radically that we can make these breaks.

Speaker 1: (19:25)
Well, you know, you train on a little bit of data and you can be as good as the average radiologist, and then you train on more and more data and you're better than any radiology.

Speaker 3: (19:32)
Exactly, exactly. Yeah. I mean, just say eight weeks, 50 grand and we got the solutions, so it was just unbelievable to me.

Speaker 1: (19:38)
That's great. That's great. So I wanted to know, how did you and Marco come to write the book and in your recent, particularly in your research along the way, what surprised you? What, what, what aha moments did you get in your, your journey to put this book together?

Speaker 3: (19:55)
Yeah, yeah. And so this book comes together because I'm seeing all this AI stuff 2011, 12, 13, 14, 15, and then Marco and I started this course, uh, in HBS called digital innovation and transformation. And what was fun about that course? What's great about, again, being a member at HBS is like we have intellectual hunting license and we write these cases. And what cases mean is that I go talk to companies, I call them up and say, Hey, I want to do HBS case on you guys. And I get access to these companies and I do tremendous vicarious learning from these guys. And we'd go spend time with them. Uh, you know, they share the secrets with us. We write it up, you know, that the cases have to be approved by, by the companies that can't be just published as a journalistic piece.

Speaker 3: (20:36)
We actually work with the companies, get them approved, uh, and then we're going to teach them. And so that just basically gives me many more reps in the economy with what companies are doing. So we're looking at the transformation story. And our thesis was Marco. Marco came from the Microsoft side. He did his early work looking at Microsoft and their platform and their ecosystem did early work on sort of Keystone strategies and ecosystems. And I'm coming from the open source. So the anti-market soft. Right. And we, you know, [inaudible], she hired me into history, right. And I'm like, are you sure you want to hire me because I'm going to try to destroy everything. Yeah. Now of course, Microsoft is like platinum sponsor of the Linux foundation. It's a world has changed in 20 years. Um, but what we saw in our work is that this whole notion of ecosystems and platforms both from open source, but also from Microsoft, was the dominant way to organize a software business, right.

Speaker 3: (21:29)
And, and so forth other a holder of the ecosystem or your participant ecosystem. But those dynamics were now basically because of the internet happening in every single industry, every single industry was basically getting this ecosystem dynamics. And, you know, there's a healthy notion of platforms and network effects and so on and so forth. And it was no longer just a software story, but an industry after industry, the same thing was happening. So we were writing about this, like, you know, from Andreessen software, eating the world, what does that mean? And how are companies responding? So we see this rise of platforms. And so our, our first set of courses were around, uh, cases were on that, and of course was around that, but then we're in the operations department. So we always wanna understand, like, but how does it work? How does this company work is the same way as GE is set up, right?

Speaker 3: (22:14)
Or is this something different? And that's when we, like, we had these aha moments because we'd walk in, you know, you look at a Google, you look at a Facebook. I remember talking early days of Google. And, um, this is like, gosh, this is 2001 or two. Right. And talking to the two there, when they were talking openly about what they were doing, you know, uh, and I was talking to, uh, their data center people, I think they said, Oh yeah, we have 20,000 servers. And we have five people managing it. And I was like, or maybe eight, I'm sorry. Uh, I'm like what? That CA I, you know, I've seen all these other data centers and, you know, it's like armies of people that had met, how is it

Speaker 4: (22:53)
That you guys are doing? So, you know, these early

Speaker 3: (22:55)
Inklings as to how, uh, these platform businesses were actually running themselves and the big aha

Speaker 4: (23:03)
Was, Oh, they're no longer have

Speaker 3: (23:05)
The old process. He's the old operating model. Right. What they're doing is that they are sort of saying, let's automate our bottlenecks. Let's actually throw technology at it. And then rethink the ways in which machines do the work instead of humans doing the work. Right. And so the simple example I use when I talk to exact students is like, you know, there's a, there are a hundred million or more auctions running every day on Google. There was a little auctioneer sitting there saying, what's your bid, what's your, that kind of stuff. Right. It's fully automated, right. Pricing. Even to the point now that even the ads themselves are automated, right. Where the content is of the answers, right. That a system, that's a value capture mechanism that works at scale without human intervention, humans are the edges of it, but it's the machines that are doing that.

Speaker 3: (23:53)
Right. And then you go, and then you started looking at, Oh, like, you look at Netflix, right? Netflix at scale, can't be driven by humans. Just humans are on the edges, maybe making some content decisions and so on and so forth. But the operation of Netflix as a service is fully automated. Your Netflix is very different from my Netflix and so on and so forth. And I think that people in the audience understand that and know that, but that was the big aha that, Oh, like, you know how we think about operating model, which is how you achieve scale, how you serve lots of customers, how you achieve scope, how you offer them many different things and how you learn that is being infused by AI, that becomes your value delivery mechanism. And, but also your business model, how you create value, how you capture value.

Speaker 3: (24:36)
That's also been fully automated, right? And so that's our view was that there's a new shape of a company showing up organized very differently, uh, uh, which is in many ways, colliding, uh, colliding with existing companies. And that, that was a big aha. And in writing this book, you know, we tried to write this book three times and we feel like every time somebody else wrote the book, we want it. Right. And what's good about this book is that it includes all the homage to platforms and network effects that those are critical. But, and I think people know that, but it's not, that's not enough. It's really that combined with a change in your operating model, right. That drives automation. But then also has, what's what we call learning effects. Right? You could learn about your customer behavior and what's going on from the data, from all the interactions. And that makes your value capture on your value creation, even sharper. So that thinking, I think was like the big aha for us to, to put this together.

Speaker 1: (25:35)
So you basically become the center of the ecosystem without even trying you become the one that brings all these pieces together and the ones that are strongest to do that. And you talk about this in the book.

Speaker 3: (25:45)
Yeah, exactly. The connector, the connector is the same thing. Like, you know, like what Airbnb is doing is the same as what Merritt is doing in the end. You're, you're renting beds. Right. Airbnb operating model is very different from Marriott. Right. But, but okay. Go find it. Okay. They're trying to get, yeah, I get that Korean man, you know, there's like individual rooms and houses versus hotels. Like yeah, that's it. But if you just think about the investments they will make in, in data and analytics and AI are, or as a vetted more than what merit will ever do, because th it's not just their merit, isn't set up that way to do what Airbnb does. Right. So it's a different beast of a company doing the same, same thing as what a traditional company does. And that's the collision that I think is very interesting.

Speaker 1: (26:32)
So you, um, you coined the phrase AI factory. Yeah. The book talks quite a bit about that. Um, and we all look at this idea of data and algorithms, how we put that together, how we do sloops

et cetera, but I found it really fascinating that you actually coined a phrase for that. And it basically the autonomous AI factory is probably another step that you, you talk about there.

Speaker 3: (26:54)
Yeah. Can I, can I share a slide with you on that? So I think, yeah, please.

Speaker 1: (26:59)
So, so,

Speaker 3: (26:59)
Um, this is the, again, this is like a plug for my book, so I apologize. Uh, so, so this is this slide. This is the AI factory, right? So by the way, the concept, again, like, again, I I'm so grateful to be in the Harvard ecosystem. This notion of the eye factory actually came from my conversations with my friend [inaudible] who runs continual labs here in Silicon Valley. And he came and said, Oh, there's a data factory that many of these companies are doing. And, uh, and I go, yeah, it's not just data factory, it's actually AI factory. So his, his notion is that, you know, you basically have to build a data factory. And we took that and we built upon that and had this, had this, had this view. So, so let's contextualize this, this picture, right? So all firms have analytics at one point or another.

Speaker 3: (27:47)
And I think Bob has asked this question to us in the Q and a right. You know, most firms have databases, infrastructure technology in a, in a, in a, in a, in a, uh, in a, in a binary fashion and so forth. Um, and so, uh, so that's the segue it's most of, most of us, most enterprises have analytics, but then analytics is in many ways the Kraft model, right? Like what handcrafted analytics, right. You've got a database, you guys have infrastructure. You're going to call your buddy in it. You say, can you run the SQL query for me please? They'll say, well, next week, I'm by tight up next week, they'll send you like, okay, what was that query again, send me an email or you Slack them. Okay, fine. Then they'll send you a CSV file. Then you'll import it into your studio.

Speaker 3: (28:34)
And then you will sort of do some, some analysis and then you'll have a report. And then you'll embed that into a PowerPoint and you present it to your, to your manager. And there'll be like some retrospective view of the world. You know, ambassador might be like, Oh, lost yesterday or so forth. But that's a craft-based way. That's how we used to make machines. That's how we used to make shoes. That's how we used to make our podiums a craft based way of doing things. Uh, but the industrial revolution was all about automation, industrialization. And thinking about modularity, you thinking about way in heading parts, replaceable parts, and so forth, and the revolution and the economic benefits that all of the whole world has gained has because we have industrialized these prior craft processes. Now, of course, we have nostalgia for the handmade leather jacket from Tuscany, and we want that, but the majority of humanity needs manufactured stuff, and we need, we need that, right.

Speaker 3: (29:31)
I mean, the super rich people can afford that the rest of us, you know, just go to target and buy a jacket and move on. And, and the, and the thing is the same thing is with data and analytics, right? Most of us have these craft-based models, this AI factory, which is at the core of, you know, from our sense of, you know, Netflix and Facebook and Google, and, you know, what, what Domino's trying to build a model based company, right. Says, we're going to have a process that ingests a ton of data. And then we invested in our data pipeline on gathering, feeding, normalization integration. We have algorithm development, either it's off the shelf where we do some tweaking infrastructure, develop an experimentation platform and we prioritize and deploy. And the AI factory basically does three things, right. And three things. One is, it does predictions, right?

Speaker 3: (30:20)
The second is that it does pattern recognition. And the third thing is that it does, uh, uh, is, um, does, uh, process automation, right? So the three piece, right? And if he's not start thinking about any enterprise, right, there's a ton of predictions going on in enterprise, right? What students are gonna come to HBS, who, which will do well with students should get promoted. You know, we should be tough that doesn't take a year off. Those are prediction tasks, right? Which customers are gonna come, what should my pricing be? What should my inventory be? What machine is going to break down? What machine needs repair, all those things are prediction tasks. So you can model your enterprise as predictions. Uh, and if you do that, there's a ton of scope for applying AI, right? Same thing about pattern recognition, right. Just think about how much time is spent by managers and executives doing pattern recognition, but it's all intuition based. It's all sort of like pre Moneyball scouting reports, right? And then in fact, you can actually, again, drive some rigor and show you patterns that don't exist. And then finally, uh, you know, there's processes and, you know, AI BI factory can do process automation. So that's the notion that came to us. And as we dug into examples like ant financial, right, Ali Baba, you know, so on and so forth, all of these companies have, have AI factories at their core. And that was like a big aha for us, uh, to put together.

Speaker 1: (31:47)
Interesting. So the, the, the, aha, you talked about the aha here. I mean, this is a sort of product, but they build the factory. They can build a product and over, after a while the technology becomes relatively straightforward. But you also talk to in your book, uh, about something that I found fascinating, which is, I think you said the line that, uh, basically technological changes, easy organizational change is hard, which maybe gets to the heart of why this would be coming out of labs at Harvard business school instead of just coming out of other types of labs. Uh, but it seems like the wrong company implementing this is just not going to get there. There's so many other, tell us about the organizational challenges and who got over those organizational challenges that you saw.

Speaker 3: (32:35)
Yeah. Yeah. Look, I think, I think this is, I still think technology is easy as a technologist, you know? Um, and, uh, it is the organization. It really is. Uh, you know, my, my, uh, uh, again, being

in this, this dense ecosystem here at Harvard and MIT, uh, you know, we, we, there's a, there's a nice, uh, density of sort of economists and sociologists of innovation and studies going back 70, 80 years as to how companies adopt technologies and so forth. And one of my colleagues, uh, two of my colleagues here at Harvard business school wrote the seminal paper in 1990 pre-internet anything. Um, and, uh, this paper was, is known as architectural innovation. Uh, and, uh, so it's, uh, Rebecca Henderson, she was a PhD student, Harvard then went to MIT. And then I was back at Harvard, uh, amazing colleague and Kim Clark, who used to be our Dean.

Speaker 3: (33:36)
Um, and the question that we're asking is like, why do incumbent firms fail? Right? So Kim Clark had this, had this program of research that said, why didn't come as from fail. And so there were two streams of thought. There was, uh, Rebecca came for us and the architectural innovation, and then the person who got super famous by asking the same question with a different answer, complimentary answer was clay Christianson. So clay was also Kim student and he saw disruptive innovation. So do something that innovation story, everybody knows, right? You have a technology that's lower price over performing it, it opens up a new niche. New niche comes in, right? I know what time it gets better and better and better than it takes over the existing technologies that are out there. And that, that, that took that, that shook Silicon Valley to its core, right? The famous picture of him and, uh, uh, and Andy Grove, right? Unfortunately magazine, unfortunately, both of them have passed away. It's just so sad. Now, Rebecca,

Speaker 2: (34:31)
Prior to place, as I also see

Speaker 3: (34:35)
Incumbent, successful incumbents, not adapting, but her view was that in fact, it wasn't new technology that was killing them. It was new architectures. And the way to think about that is a technology has, has two things. It has components, all the different widgets that come together and an architecture, right? So a first ever cell phone had both new widgets and a new architecture, right? Because you need an antenna as you need batteries, your base stations, so new components, and it was put together a new way. Um, you know, uh, uh, then you can think about incremental. Like if you think about iPhones two or three, four, five, six of all incremental change components, didn't change your architecture. Didn't change. If you think about it got iPhone, right. Uh, you know, sort of touch ID, face ID apps, all of those are modular. You just take out a component, put a new component in, but do the same functionality. Architectural innovation is when you take existing components, but you organize them in

Speaker 2: (35:29)
A new way. And her

Speaker 3: (35:31)
View is Rebecca Rebecca. His view was that it was the, it was the failure to recognize new architectures that were killing companies, right? Because they will see the components that go,

we have all the pieces, right? Remember Nokia said, this phone is a crappy phone. We have touch screen. We have an app store. We have a music phone, we have a map phone, right? What do you, what kind of, what do you want? I've got them all from whatever desire you have, right. That wasn't the point. It was the rearchitecture that mad mattered. And so architectural innovation is critical. And so Rebecca, his view was that architectural innovation is both hard to see, but also what we see in industry is that the architecture of the technology represented mirrors, the architecture of the company. There's a, one-to-one almost mirroring of that. And it's been talked about quite a bit in the literature.

Speaker 3: (36:24)
And when the architecture of the technology changes as a new dominant architecture, then the organization also has to change. But this aura change we know is tremendously difficult, especially as large companies, except for exactly for the earthquake. So this change is what is so difficult. And so they look at the architecture, but they focus on the components. They say, my component X is better than your component of why this kind of stuff by saying the customer value is shifted, the architecture. You're just not seeing it. Right. And that's exactly again, what Amazon did to retail, right? Brand new Ark, take existing pieces, new architecture. My favorite example of these days is Peloton, right? You look at Peloton, every single piece of it has existed out there, right? They just put together a new package or new architecture, and also changed the business model. Right? Because now what you could do is you can say, Oh, I can put a subscription business.

Speaker 3: (37:21)
I don't need to pay per use. I can do a subscription business. I can, I can build the best spin experience for you in your home. You don't have to go to the, yes. Some, some people like going, uh, pre COVID times and sweating together and coughing and breathing on each other. That's all gonna change, you know? Sure. Vaccination, uh, but, but what matters was a not touchscreen internet, you know, broadcasting of, you know, there's been live streaming of fitness content forever, but in this new architecture with music was not done this way before. And so I think this new architecture is what you know, and it's the same Pasek story, right? They went to soul cycle. They said, let us stream your classes. And so it's like, it says no frigging way. We were all about what happens with the 50 people in the studio. We don't care about a thousand people in, you know, wherever doing our class.

Speaker 1: (38:12)
So pizza, if you think about the company, does your, the best example perhaps of rearchitecting is Medina return it today? I mean, if I pulled the right journey, I got their FDA approval today. Um, you wrote this funnel here, I'll show this, uh, it's all scratched up cause I've been dogging it and writing on it. But, uh, in September you wrote a case study on as the true architectural change a company that is a technology AI technology platform and not a bio discovery drug discovery company. So they turned this on its head over, uh, over about seven to 10 years. Can you talk about this case study and this exciting company?

Speaker 3: (38:53)

Yeah. So part of this was case study was like people saying, okay, Kareem, great AI, AI, blah, blah, blah. But you're just talking about consumer giants. Like, come on. Like, I can't be Facebook. I can't be Google. I can't be this. I can't be like billion customers. Like I can't like, this is only for those guys. Not for me. Right. My problems are too exotic. I'm too specialized. I'm in this whatever. So I've been on a hunt. And if there's people in the audience that have examples of, you know, real world AI and real world, you know, uh, full of atoms or cell settings, please come to me and let me know. And I'd be happy to do a case on you guys. Um, so Madonna, uh, again, a great again, you know, pitching Alma about our great Harvard MIT collaboration story, as well as its fantastic, uh, new bar.

Speaker 3: (39:37)
And you know, who's been on our faculty right now, uh, teaching teaching with me as a, as a co-founder, um, comes up, comes about, uh, you know, reducing Cummings as it comes about with their view as Lasher ventures to say, we don't want to be in the white space. We want to invent new space, right. They go and they always ask, what if questions? And part of this was a conversation that Newberg got into saying, could you use MRI to make drugs inside the body instead of it being done exhausted asleep. Right? Uh, and there was some early work elsewhere that was not so promising. And then they went ahead and figured it out. Now, the thing about them is that in many ways, this is the first information molecule going into your body, right? This, this is not a molecule that is going to respond to our cells and so forth.

Speaker 3: (40:31)
This is design, it's a programming instruction. Emrani is programming assertion to go into your cells to produce a specific protein, which then tunes your, um, your, your immune system to then, uh, be responsive to the COVID virus. So on this, on its head already, it's an information molecule, right. Very different paradigm. And then they, and then what they realize is, Oh my goodness, like, we don't know if our first, um, uh, if you said we're gonna do disease X or disease, Y or Z, that it's gonna work out. We just don't know because there's so much uncertainty in science. Right. And so we're going to do a platform. This information molecule has many different applications across the board, right? I'm just gonna do 20 programs. So as a, as an early stage biotech to come in and say, I wouldn't do 20 programs, you know, vaccines, cancer, vaccines, and this and that is crazy.

Speaker 3: (41:31)
But they were able to show really good data to convince their investors as well as partners to do that. And so, so, so this is part of their they're big, they're a platform company where the same technology is being applied to a whole bunch of different diseases and they can succeed with that. But also what was amazing is that Stephane Bancel the, the, the CEO of a company had this, this big moment where, uh, he walks into his scientist's office and sees a wall of Excel sheets. And his scientist was very, very proud. They sort of coding every single Excel sheet, right. And every single cell with the code ons, right. With all the ACTG. So it was sort of laying out the MRNs sequence. And he's like, if there is one slip up, anywhere in this process, that's six months to a years worth of experiments down the drain.

Speaker 3: (42:27)
That was his big aha moment. He goes, we can't run like the way we used to do pharma, which is like, take notes, figure things out, passing it over to somebody else, go back. This whole process has to do digitized and asked to be automated. So they have invested a ton on basically automating the clinical processes, their research processes and their manufacturing processes to the point now where that somebody can design a drug on their web browser. And then the next day, the factory makes it and gives it back to them. I mean, that level, the Corona virus side, the actual time for them to create the molecule was actually four hours.

Speaker 1: (43:09)
The rest was four hours, but normally it takes about 10 years. Yeah.

Speaker 3: (43:13)
Four hours to make the molecule. And because they had done this before with other works with the NIH on SARS and MERS, before the news about the spike protein, they knew the safety profile already. They knew all that information. So this is the, again, the platform learning we've this molecule that we have is, is, has all these capabilities, right. And that allow them to basically four hours to build the rest of us, to prove out safety, this and that for the regulators. And then basically, you know, they saw that they saw the, the, the, and they never actually ever touched the virus. They had the, the sequence of the virus from China showed up to them, and then they use that to then create the thing, and then it made this admission. So just a remarkable, amazing story of using AI at scale. It's just, it's fantastic.

Speaker 1: (44:02)
So I highly recommend everybody to read the case study. It's, it's quite amazing a well-written case study. And I think with that, maybe we'll go to our audience questions. Uh, uh, so Todd, if you want to, if you want to queue up, I think, uh, we have max Hoffer. Who's gonna ask us a question. Max is a PhD, uh, and lecturer at EPL, Kareem in Switzerland, and you're the expert in open-source Kareem. So max is your perfect question. Hi, professor Connie. Thank you so much for these exciting insights. It's really fun listening to it. So what you're saying, I've got two related questions. The first one being when you're accumulating and developing these assets, especially these intangible assets that you've mentioned, um, what took firms pay particular attention to them? And the second one, um, in a world of open source technologies, um, how confirms build and sustain competitive advantage. Great, great question.

Speaker 3: (44:59)
So force is data, data, data, data, data, right. I think, I think I can't tell you enough. I, you know, the, the, you know, I think, I think most sort of digitally native or AI native firms get data and the value of data and building infrastructure, but incumbents are just terrible. Right. And so I tell executives, you know, like invest in the janitors, invest in your data pipelines, invest in cleaning your data, do that first, because we know garbage in garbage out, you know, and talking to folks at these big tech firms where they try to work with their partners in other organizations, they always run into the data silos and data cleanliness, all that kind of stuff. So if you're going to be playing in this game of AI factories, which I think everybody has to, right. AI factories, my view is

going to be necessity either for your own operations, or if you're gonna be sitting, talking to Amazon, sitting talkative to Google and so forth.

Speaker 3: (45:50)
So investment in data and infrastructure. And then by the way, that also means organizational change because right now data is, are siloed and hidden and co and, uh, and not shared across organizations. And so you've got to work your off on that. And so I would say, don't hire the MIT PhD yet, right. Go or the ETH, or if UFL PhD, right. Fix your day, I think, and then hire them because that's the only way. And if you hire those guys, they're going to be spending clean data. So you all of a sudden taking a high-end PhD guy and then making them do sort of data engineering work, which is a total waste of time for their skill level. So I think, I think data, data, first thing, second thing, which is very interesting, you know, I'm on the board of Mozilla, uh, open source company with, for a browser.

Speaker 3: (46:32)
And this is one of the reasons why I joined Mizzou's board because in many ways the open-source software, uh, Wars have been won, right. Microsoft now is premier funder of the Linux foundation, right? When, before they were saying it was an IP destroyer, right. All of a sudden there are members of that. There's been a whole sea change of that. My view is that in fact, uh, and you know, Google, Facebook, Amazon, everybody does open source, right? It's not as if they don't do open source. Yes. We have fights about how open source are they really, but, but sort of the belief of sharing the code and so forth is out there. Um, I think the game has shifted to, in fact, how you build a complimentary assets around the algorithms. So algorithms themselves, aren't, aren't gonna, you know, before we would be secretive about our algorithms, but now if TensorFlow is free, right.

Speaker 3: (47:23)
And Google's business model is for, for, for them to have the best hosting service for TensorFlow, they're going to keep making all that stuff for free because they just want usage and consumption of their services. So then I, us, uh, who are playing in the eye space, we have to think about our business model. We have to think about value creation by the capture and how we sort of do this, uh, in a way that, that, that allows us to have competent Coventry assets build network effects, big build learning effects. That's super interesting, so much. Appreciate it.

Speaker 1: (47:57)
Nice. Thank you. Uh, Cod, I think a new show shows next, is that right? Yeah. Well, anyway, um, a new show. Hi, how are you? Um, welcome to zettabytes slide.

Speaker 5: (48:17)
Thank you. Thank you for the opportunity. And thank you for having this conversation today. Uh, my question was on legislation. How do you think legislation in the us is going to well as customer data is being used more and more, and also maybe if you have a perspective on the customer sentiment answers, I would love to hear that.

Speaker 3: (48:33)
Yeah. Look, I think, I think this is, this is a great question. Who's young. So a lot of your assets, I think, I think, you know, uh, we spent a whole chapter talking about the ethics of digital scale scope and learning in our book because, and that was the toughest chapter for us to write because we're like, we're B school professors and technologists. What do we know about ethics? Right. Yeah. We, we also were sat right down to go across the river to the law school and the philosophy department. Right. And so, uh, and the religion department. And so, but as we sort of looked at our work, we said, look, just as I can scale in an exponential way, the benefits of AI and networks and value creation, I can also scale the harm. Okay. I can also see the harm. So the bias in algorithms from data, the fairness questions and so on, and this is, this is, uh, you know, um, uh, a primary question of concern to business, not just after you built the systems, but beforehand.

Speaker 3: (49:28)
And so my view is that sort of the work being done in computer science, in philosophy and law around AI ethics actually have to be important also to the business school world and to business training and business education and for businesses. And we can't any more sort of differ systems designs just to engineers and say, oops, we didn't realize our data was, was biased. We didn't realize our labeling was biased or non-representative and so forth. So I think the first dream is that the, this legislation actually is the impact of that. You're more broad-based. So w one of our incoming colleagues at HBS is Seth Neil. Uh, he didn't see us and, and, um, uh, patient in CS, uh, and he asked this question in his job talk, he goes, do we want to be, to be fair on average? Or do we have to be fair to everybody?

Speaker 3: (50:20)
Machine learning was say, if I'm fair on average, we're doing good statistically, right? The law says, no, you have to be fair to everybody. So there's a big disconnect. And in fact, you know, I think he had said, don't quote me on this, but it was like a large portion of algorithms today being used by firms today are probably breaking some law because they haven't thought about fairness this way. And in fact could, could have these consequences. So, so all this is pushing into your question about the data, because I think data quality, the data who has permission to it who has access to it. What is the value of the data? Those are open, open questions. Do, does Google really need to save my search from 10 years ago? Probably not. Because what they care about is what I'm searching in the moment, right?

Speaker 3: (51:04)
Because their business model is based on showing me the ad, the relevant ad in the moment, not what I was searching 10 years ago. And the 10 years ago, search is probably not that predictive of what I'm searching today anyway. Right. And so there's a whole bunch of questions that are being raised for all of us to say, what is the value of data what's perishability or not. And then, and then what do consumers want, you know, in our Mozilla work, we know we talked, we would talking about privacy for a long time, but consumers say they care about privacy, but then they act in a very different way. And my colleagues at MIT and also here at HBS have shown the privacies of weird, good people don't know how to value it. Um, some of the work that Leslie

John has done has shown that in fact, you can get people to reveal amazingly crazy stuff, amazing and crazy private stuff, simply by the way you show the website to them.

Speaker 3: (51:59)
Uh, and so, so I think, I think we're at the beginning stages of understanding privacy, consumer preferences. I, you know, I think people, it's kind of like everybody wants to eat healthy and exercise and they say, when you ask them, do you care about exercise and eating healthy? Yes. Do you actually do it? No, I'm happy having my cheeseburger again. Right. And so, so, so I think that's the dilemma we're faced with, and I think there's, you know, there's utility in the data there's utility and the learning is aggregate learning at the same time. There are these concerns and we have to figure this out. What's also interesting is that there's a ton of research being done about how I do, I do edge device learning, edge based learning instead of going to the central cloud. And so, so anyway, I think this is important business concern. Policy. People need to understand this, uh, I think much of the legislation lags behind practice or doesn't understand consumer behavior. Um, but I think it's a board level issue for companies because they actually have to understand what's at the guts of their systems and are they actually being following the rule of law in the first place above and beyond everything else that we care about. So I think, I think it's a rich area for people to go create companies in rich area for people to help companies with and understand.

Speaker 1: (53:16)
Yeah. Sure. Thank you very much for your question. Uh, Todd, we have, um, a bunch of questions in Q and a, but I'm wondering if you can let, um, anybody in, or if I should ask the question specifically, one of Kareem's colleagues here, uh, from Harvard, uh, Tommaso, di Davey, uh, asked a question, uh, from the urban stack initiative. Um, I wonder, uh, if you could let, just, just let Tommaso in or if I should, uh, go ahead and ask his question, let me answer it. Well, Kareem, I think you can actually see it up there. Yeah,

Speaker 3: (53:50)
Yeah. I can see it. So can everybody else see the question too?

Speaker 1: (53:55)
I don't, I don't think so. So maybe if you want to paraphrase it and then answer it.

Speaker 3: (53:59)
Yeah. So Tim is saying, uh, he's a co-founder of the urban stock initiative at the school of design, which is just across the river from us and they built this urban stack. Uh, and, uh, what they're trying to figure out is, uh, you know, how do we account for the intelligence interactive being and connection into our urban stock. Right. Uh, and do I think AI will have a big impact on the built environment? Um, I think, I think it's a really good question. So, uh, one of my favorite examples, again, these joint MIT, Harvard examples, I get I'm in this ecosystem a lot. So, you know, I have fewer Stanford examples or more MIT, Harvard examples, uh, is, uh, the street score project that got developed a few years ago out of the media lab. Uh, and there, what they were able to do is use crowd labeling to look at Google street images, to tell you if this feed

image looked safe or, or, you know, rated out of, out of 10 or something about how safe this is refill to you.

Speaker 3: (54:58)
And then they had this labeled data that they built algorithms across, uh, sort of the, uh, the, the Google street, uh, infrastructure to then predict, uh, you know, gentrification predict values, a real estate predict a whole range of things. So I think, I think becoming smart about the built environment and also sets refining and you know what, uh, mean Zang from Alibaba's is data flying. Our infrastructure is actually like a really huge opportunity. And you can imagine, again, I, we get a bit of this already. Like, you know, we have a card access system into our buildings, so that's doing some location level tracking of me and so forth. How do we use that for good? How do I alert you that you're too close to somebody now in the, in the pandemic, how to restrict access based on that? How do I, based on the oxygen I'm breathing in and out, uh, you know, in my, in my rooms come up with some predictions about health and so forth, uh, you know, uh, Alan out of the, uh, out of, uh, out of, uh, the school public health has done a bunch of work on healthy buildings.

Speaker 3: (56:01)
Healthy buildings will rely on centrification on AI. And just as Google can run a hundred million auctions a second, right? The buildings should be able to run a whole bunch of calculations, figuring it out, what, what, where to direct airflow, where to vent air and so forth. Right. And I think that's, that's, I think what you're trying to do, I think an urban stock, I think it's a pretty, pretty cool thing.

Speaker 1: (56:24)
It's really, really fantastic. We have so many questions, but unfortunately we're out of time, you are such an amazing speaker, such a gifted speaker Korean. We're really an honor for you to be here today. And I know you were out on your book tour when COVID hit, uh, you had to Hightail it out of Europe and get back to the U S and, um, but COVID has made your book even more.

Speaker 3: (56:46)
Yeah. Yeah. Look, can I show one more graph? Let me show one more grants, uh, to contextualize all this screenings down. So the thing we say in our book, um, is, um, that there's a collision happening. There's a collision happening between what we call sort of digital businesses that had this, uh, convex curve of value creation is exponential curve, uh, set up in a very different way with a traditional product business. That's, that's sort of set up in silos, right? And we see this collision over and over again in a range of settings. Uh, and that this is what happens is that the digital business takes a while to get going and to accelerate. And so the traditional business looks down upon the visual business and says, why would I invest? Why would I invest? Why would I invest? Uh, uh, and, um, uh, uh, and so I thought the coaching happens to like, right, so this concave curve, uh, intersecting with this convex curve and this collision happening, this is by the way, the COVID attack on our healthcare system COVID is an exponential disease, right?

Speaker 3: (57:58)
Hitting with the flat capacity healthcare system. When you say, find the curve, what we mean is right, make sure that this collision doesn't happen. Right. And what happens with COVID like it's too today for tomorrow 60 after it's an exponential process, it's slower the beginning. So we get fooled and the infection rate, the death rate, lags, the hospitalization realize infection, which, which then lives for the lags, the death rate. So it takes us a while for this to get going. But if it gets into a runaway process, like it is today in the U S it's too late, it's too late. The collisions happen and we're all in trouble. That's the COVID versus the healthcare. But the same mechanism that incumbents face with these digital businesses, right? They don't understand exponential systems. They don't understand how these things take off. And then the collision happens and then they wonder what happened, right? It's the same story. It's the same story. So I didn't know about this. I mean, this was done, you know, three years ago, four years ago. And then we sort of packaged in our book, but as we were like, as Marco and I were going around the world and going, Holy, what's going on? And then we're like, Oh yeah, this is exactly the COVID crisis, uh, in our collision breath. Anyway, uh, there's a, there's an awkward,

Speaker 1: (59:15)
That's amazing, quite an amazing analogy, I guess that tells us that we need a new architecture for our government, but that'll be, that'll be for a whole nother hour. But the most important thing I do want to say right now is get your surveys in. But a lot of people on this call today, we'd like to give out, we'd like to give out a lot of books and Kareem, we can't thank you enough for being here today. You were truly amazing. Have a really great holiday. Smart. Everybody else was coming. Really appreciate it. Take good care. Everybody take care. Bye.