chapter_one__hallucinations,_Hype,_and_the_realities_of_al_v1 (720p)

0:04
Well, now I'd like to introduce our last guest of the day, Daniel Huntenlocker.

0:07
Dan is the inaugural Dean of the MIT Schwartzman College of Computing.

0:12
He's one of the more influential thinkers out there when it comes to AI and is, among very many other things, the author of The Age of AI and our Human Future, which is co-authored with Henry Kissinger and Eric Schmidt.

0:23
Dan, welcome to M Tech AI.

0:25
Please join me.

0:30
Have a seat.

0:31
Have a seat.

0:33
Thanks for being here.

0:34
Hey, great to be here.

0:35
Thanks.

0:37
Your book came out five years ago, five years ago in November.

0:40
But yes, nothing's happened since then.

0:42
No, nothing at all.

0:44

I'm curious, like if so, if, if you, as you go back and you think about the process of writing that book and the things that you learned, if you were to write it today, if you were to go back or maybe to update it today, what would you, what would you add to it?

0:55
What would you, what would you change about it?

0:58
Actually, I'd change very little.

1:02
I I think that the most important thing to add would be all the examples we have today of the kinds of issues that we were identifying there.

1:13
Of the fact that AI is really this different form of intelligence from human intelligence.

1:21
But because the only form of intelligence we're used to dealing with, at least one that speaks to us and writes in our language, etcetera, is human intelligence going to be very challenging for us to figure out how to really work most effectively with this new form of intelligence.

1:37
And I think we're seeing that all over the place today.

1:42
There are many, many more examples.

1:43
You know, at the time that we wrote the book, there was no ChatGPT yet.

1:49
It was GPT 3 and then we did in the paperback, which came out about a year later, we had a forward that talked a little bit about ChatGPT, but that's just sort of for setting the context.

2:03
So a lot has happened now that, you know, it's hard not to view AI as having human like intelligence, but it's very different in ways that we don't understand at the moment and that are very important to making the most effective use of it in terms of benefits and also in terms of avoiding downfalls.

2:26
I'm reading a book about an octopus right now, which is very interesting, but also a different form of intelligence.

2:33
But I'm curious what you mean by that when you say it's a different form of intelligence.

2:37
Like how so?

2:39
So when you're used to dealing with a human, you expect, you know, somebody who has hopefully, although some of us are amoral, but somebody who has, you know, some deep sense of morality, has judgement that's come through human experience.

2:56
Various sorts of things that when you think about the ways that machine learning systems train, and I think it's inherent, it's not today's models.

3:07
These are things that will be with us forever.

3:10
You just end up with different forms of ability to reason.

3:18
These things can do a very good job of a peering to follow moral judgement, but the difference between appearance and reality becomes extremely important, for example, in any high exigency kind of situation.

3:34
But it's often important in the day-to-day, you know, and I think we see it in everything from the ways that, you know, the medical profession, which is fairly conservative about making change, is being fairly slowly, being fairly slow to adopt AI in clinical settings.

3:52
Although is really starting to now to, you know, areas like software engineering where it's completely transformed the way people code, the way programming gets done.

4:04
But it's still certainly apparent to me, and I think to many others, that this is not replacing the importance of human judgement and developing software.

4:15
It's fundamentally changing that role.

4:17
But if anything, I would say that human reason and human intelligence is more important today in the age of AI than it's ever been, rather than being sort of displaced by that.

4:29
And yet the narrative, the norm is all about displacement.

4:33
And why, why, why, why is that?

4:35
Like, what is it so important about human intelligence that we haven't been able to do with AI yet?

4:41
Well, I think some of this becomes a philosophical debate.

4:45
And so, you know, gets hard for, I try to tie it to sort of real experience.

4:50
But then, and in fact, the the book with Kissinger and Schmidt had a lot of philosophical elements to it at the time that we wrote it.

5:00
But I think many of us believe that living in a simulation of reality and living in reality are not the same thing.

5:08
And so it's really down to that same kind of question.

5:13
What is it about, you know, human decision making, human reason, you know, knowledge of the sort of consequences of your decisions for yourself, for your loved ones, for, you know, for your coworkers, all other humans with whom you have various kinds of relationships.

5:35
You know, how do you, how does your decision making, how does your thinking affect that whole network of other people?

5:41
For most of us, that deeply effects how we operate and what we do.

5:46
And those things are absent from non human actors.

5:52
That doesn't mean that non human actors aren't extremely powerful.

5:56
It's just that trying to shove us in the box of a non human actor or shove a non human actor in the box of a human actor, which is what you're doing if you're trying to say, well, these these, you know, these are the same things.

6:07
They're completely interchangeable, just sort of fundamentally flawed as an approach.

6:13
That seems to me like that doesn't stop us from doing it.

6:15
It doesn't.

6:16
No, it doesn't.

6:17
I think certainly that's a mistake you see people making is trying to shove it in the slot, right?

6:21
But that's that also seems to me that maybe what you're saying here is that some of the concerns that people have about what about the effects of AI in the workplace, for example, or just on society, are those are, are those overblown?

6:37
Are those are those are those too much?

6:40
Because it is, it does turn out to be difficult to shove an AI into a human slot.

6:44
It's, it's, it's a really good question.

6:47
So I mean, the, the, the way I think about it is that AI is such a big change.

6:52
You know, I mean, human reason was the only form of reason.

6:55
I mean, sometimes we would think our pets were reasoning, but come on, you know, like they, they certainly weren't talking to us.

7:01
And let's talk about hallucination.

7:02
We were really hallucinating the but, but with with AI it, it, it feels very human like.

7:11
And that is a really fundamental change.

7:13
We're not used to interacting and spoken and written language with anything except other people.

7:20
And, and so it, it, we're right to have that feel disorienting.

7:25
There's no other way to feel about it than that it's disorienting.

7:29
But I think that the, the, what we're attributing to it in that disorientation is actually in, in many instances, causing us to worry about some things that maybe we should be a little less concerned about and maybe not as concerned about some things that we should be more concerned about because in this different form.

7:52
And, you know, we're seeing it all over the place.

7:54
So like, you know, when you look at, you know, great software engineering, software engineers and software engineering teams that are making the most effective use of these coding tools.

8:03
It was just in the previous discussion you were having, you know, what you really want to do is pair one or a small number of your very best software engineers with these coding tools and you know, maybe an army of these coding tools.

8:22
But it's again, it's about this figuring out how to bring these two forms of intelligence, human and machine together that's delivering the best outcomes in big coding projects by far.

8:34
I mean, anybody I know who's, you know, a senior, you know, VP level or, or engineering manager or even individual coder on one of those kinds of projects, It's not like a completely green software engineering throwing these things loose is going to develop great production quality systems.

8:52
They can build stuff and sometimes that stuff is very interesting proof of concept.

8:59
But you know, anybody can write a novel today too with, with, you know, generative AI.

9:04
And you know, we're used to calling that AI slop today, right?

9:08
So it, you know, I think that to me, part of the difference between slop and, you know, really fundamental advances is the ways in which we're using this, the fact that we're recognizing that they're different strengths and that we're learning how to bring those strengths together.

9:24
And those are things both about individuals and how we're interacting with these systems, but also how these systems are being designed and developed, which right now it's not so well suited to sort of, you know, bringing together two different forms of intelligence and also how, you know, organizations and employers are bringing in these things to bear, right?

9:46
So it's, it's really, you know, a, a, a joint sort of mindset, responsibility, whatever you want to call it, of people using AI, organizations using AI and the, you know, those of us who are developing and, and, and deploying it.

10:02
I don't know how much of our previous converse of the previous conversation you're able to hear backstage.

10:06
But one of the things that I was talking about is a friend of mine who's ACTO at a company who had said to me something along the lines of I'm glad I'm not a young software engineer today, Glad I'm not an entry level software engineer today.

10:15
There are going to be short term challenges that people face, especially as they're trying to enter the workplace.

10:19
What are your thoughts on on how you know someone who is graduating just just out of school?

10:26
How should they think about working with AI?

10:28
Yeah, this is a questions how to tease this question apart in a good way in a short time because it's a very involved question.

10:40
You know, I think MIT in schools like us, our students often used to complain in computer science program and computer science classes.

10:52
You know, you're not teaching us enough about the practical things so we can run out and be a, you know, productive coder tomorrow.

11:00
And we said correct, because we're trying to teach you how to think computationally for an entire career where the ways that these things develop are going to change multiple times and going to accelerate during your career.

11:18
This is probably the biggest proof point in my career of that kind of a change where you go from, you know, coding tools that help you develop more code to coding tools help you develop less code but have more code produced, which is a very different take on the problem.

11:39
So I think the the the education and the programs that have long thought about how to educate people to think deeply about the objectives of building a large scale software systems and how to get those done most effectively what we sometimes call computational thinking.

12:00
I don't think that I think that those jobs are going to get better, much better.

12:05
They are getting better today.

12:06
They're getting better.

12:07
Talk to some of those developers, whether they're relatively junior or super senior, who you know, are now working in an environment where they have the support from their management and where the tools are getting better, you know, very rapidly.

12:21
Nobody's telling them.

12:22
Just set this thing loose and don't care what it's doing right there.

12:25
They're stupid ways to use these tools, those people.

12:30
It's exhilarating.

12:32

You can in a, you know, a team of 6 good software engineers in a few weeks, do something now that might have taken, you know, a team of 50, which involves a lot more of those boring coordination meetings that software developers hate.

12:46
Might have taken a team of 56 months a year.

12:51
But you're not going to get that by not understanding the different roles of somebody who really understands how to think at the big systems architecture kind of level of a good season software developer, someone who understands broader computational thinking and what I'll call a code jockey, somebody who's good at producing a lot of high performance code.

13:11
Those are different things.

13:12
The latter 1 is much more subject to what some of these tools today are able to do.

13:19
So I think, you know, there are a number of people who trained more to write code.

13:29
That's, that's probably a less good career path.

13:33
There are people who really trained and have the experience.

13:36
And this doesn't just mean, you know, great university education.

13:39
I know people who don't even have college degrees who are unbelievable senior software architects, right?

13:44
This, this comes from some mixture of experience and the right mindset.

13:49
If you don't have the education, and even if you do have the education, you didn't have the right mindset, it's still a problem.

13:55
This is the golden age of that.

13:58
So there's a big bifurcation happening there.

14:00
And I, I think frankly, it's going to happen everywhere that those I, I really appreciate that answer.

14:05
That was very thoughtful.

14:06
I'm starting to get some questions.

14:07
The chat that was a long answer.

14:09
It was a long answer.

14:09
It was very interesting.

14:10
Thank you.

14:12
Starting with some questions to remind, if you do have questions, you can come to the microphones in the room or you can ask them on the chat online and we'll, we'll get to them here.

14:19
You, you mentioned that, you know, maybe we're thinking about the wrong challenges.

14:25
What, what are the, what are the challenges we should be thinking about in terms of AI adoption as it becomes, you know, as it becomes more fully integrated into society in the ways that we're seeing it happening rapidly?

14:37
Yeah.

14:37
So there's one that's sort of at the center of not just me, but a number of us at MIT have been talking and thinking about, you know, at least since the the the last five years.

14:52
And certainly the history of this at MIT goes all the way back to JC or Licklider or like, as he was called, a faculty member here in the, you know, 1960s.

15:06

And it's really about how do we focus on collaboration of human and intelligent machines rather than just focus on intelligent machines and what they're able to accomplish on their own.

15:20
And when you think about the whole early development of AI, you know, the Turing test or Imitation game as it was called, that was all about replicating human performance, right?

15:31
In fact, it was about, could you tell if this was a human or an AI or not?

15:34
Well, that's long gone.

15:36
Usually if it seems a little less clueful in like, you know, some kind of test you would give somebody, it's probably human because humans aren't good at pop tests, right?

15:45
So, but that really oriented us toward not thinking in the whole design and development and deployment process toward this human machine collaboration.

15:58
And it's a harder problem.

16:00
So if you even just think about the challenge of, OK, I have some benchmark that I'm benchmarking, you know, a machine learning model against, it's all kinds of challenges and sort of deciding is the benchmark really reflective of like reality of using this this large model?

16:18
But now if you think about how do I benchmark how much better this AI is together with a human than a human is alone or an AI is alone on this problem.

16:30
Just even trying to frame that as a benchmarking challenge, much less than how do you sort of adapt it to a particular work scenario is is much more challenging.

16:41
This is a harder problem.

16:42
It's not like we're sort of not working on something that would be low hanging fruit.

16:48
But I think that makes it all the more important that in addition to developing systems, the way we're developing them today is stand alone forms of intelligence that there's much more focus on the integration of, of human and machine intelligence.

17:01
So as you mentioned, we, we, we, we build AI by design to be human like it's, it's trained on human data to do human tests or to succeed at human tests.

17:14
Are there alternatives to building an intelligence?

17:16
What what alternatives are there?

17:19
Well, so a couple of things.

17:20
So one is you could, I guess the first thing I would say is actually one of the big advances in large scale machine learning in the last couple of years is less and less training on what humans do and more and more training on other things, right.

17:34
So when you start, you know, so like self play in any sort of situation that you can set up multiple AIS playing against each other, a lot of training now is using that and not using sort of, you know, human reinforcement learning type feedback, but actually just these systems judging each other or, or, or, or interacting with one another.

18:00
But also when we do think about any sort of use of AI in the physical world, more and more, it's it's sort of, you know, raw sensing data completely uninterpreted by humans, you know, images, you know, audio, any, and I don't mean of people talking like, you know, just large scale sensing data listening to the machines running in a plant and you start hearing noises that are indicative that something's going to go wrong, right?

18:25
I mean, it's just, it's a lot of this kind of data going in directly in a way that that that's not encumbered by the human perceptual system.

18:34
It's encumbered by whatever system is recording is again a difference between the machines and the humans.

18:41
The human visual auditory systems are extremely powerful in ways that, you know, cameras and microphones and so forth aren't, but they are also limited in ways that those other things aren't.

18:52
So it's maybe a more concrete version of this.

18:55
These are different systems, different machine learning algorithms, different sensing capabilities.

19:01
And so it shouldn't be hard to understand that at least in principle, if we could harness the two of those together, we're going to get better outcomes.

19:08
Like if you and I think exactly the same way, we're more likely to come up with a less good answer than if we don't.

19:14
Right.

19:14
So it just it's it this is not just about humans and machines.

19:19
And so that was sort of the the the first piece that this is all all already happening.

19:23
And now I need your question again, because what are the alternative?

19:28
Like what are the alternative approaches to, to developing an intelligence?

19:30
Oh, yeah.

19:30
Yeah.

19:31
So that.

19:31
Yeah.

19:31
So other approaches.

19:32
So then I think the other thing is that we really need to work more and more.

19:37

And I think this will start to happen in the software development arena because of this partnering of experienced engineers, bringing different kinds of knowledge to the table than these machines is starting to become.

19:49
So valuable companies, at least that are at the forefront of doing this, are going to start to learn how to do that more effectively.

19:56
But for instance, most of what AI does is relatively uninterpretable to people right now.

20:01
It's a bit of a black box.

20:03
Well, if you really want human machine interaction to work in a reasonable way.

20:09
These things have to somehow be legible to human beings what it's doing.

20:13
This is not necessarily the same as, you know, explainable or justifying.

20:19
They're just, there has to be some things that I get some understanding of what that thing was doing.

20:25
And in fact, you know, one of the complaints that I do still hear from, from from senior software folks is, you know, if there's something wrong with some AI generated code, some of some subtle bug, just throw it out and have a human do it.

20:44
Because it just, you know, part of debugging subtle bugs is you try to get yourself in the mindset of the developer.

20:52
Why did they make that stupid mistake in this absence that there is no sort of mindset, there is no sort of legible kind of more abstract understanding of what these systems are doing today?

21:04
And I think those kinds of advances will be very powerful there and will be powerful elsewhere.

21:08
But it's a different framing of the problem.

21:12
Like, I think when people have talked about interpretable AI, they've more right now been worried

about sort of, you know, sort of moral and ethic, you know, like making more ethical AI, you know, sort of is it being biased?

21:23
Is it Those are also important things, but it's a very different domain and not one that's oriented toward effective collaboration where you could start to measure the collaboration effectiveness.

21:35
So I think that those are some of the kinds of things that we need to start to add to the stable of the ways that we're developing.

21:42
It's not like we should stop doing what we're doing.

21:44
It's it's paying off a lot.

21:45
But it's that these other things are harder and therefore probably going to take longer and are going to offer, you know, even more impressive set of things that we can do.

21:57
Appreciate that.

21:57
Yeah.

21:58
Hi, My name is Ashmita Gupta, and I'm a computer engineer by education.

22:03
And we talked a lot about for, you know, graduates coming out of school or going into school, what should how should they be thinking about it?

22:14
My question is, how should educational institutions be thinking about adopting their curriculum in the light of how AI is now changing, you know, once the kids graduate?

22:25
So what's your thought on that?

22:27
Yeah.

22:30
Again, many layers to this.

22:31
So I'm actually going to start with academic content about AI and how it works and its systems, which is very different from say, how AI might be used pedagogically.

22:47
So I'll take those as two separate questions that they sort of meet in some places.

22:52
So I think when it's understanding about machine learning and other AI approaches, it's extremely important that the that gets more and more integrated with every academic discipline.

23:05
And that's something that partly by luck and partly by design, with the Schwartzman College of Computing at MIT having started in 2019, which was, you know, ahead of this gigantic tsunami that we're living through right now.

23:22
We've put in place a whole educational set of educational offerings that we call the common ground for computing education, which integrates the forefront of computing, both CS and AI with disciplines across MIT.

23:37
And these are classes that are Co taught with 20 different departments across MIT.

23:41
So you get faculty and engineering departments and you know, in Sloan, in, you know, in humanities and social sciences all across MIT working on how to engage and sort of like thinking about machine learning in their discipline.

23:59
So I think that's extremely important when I think it comes to the, I think when it comes to the pedagogical side, it's extremely important to, to figure out how to use AI effectively in your teaching in any department, any field, any faculty member.

24:19
Like, I think back to like, I'm old.

24:21
So the calculator, you know, replacing the slide rule is something that I still remember.

24:27
And I remember all these faculty, I was a pretty young student at the time who, you know, banned programmable calculators in their classes that was really going with where things were going.

24:38

The, and you know, the banning of AI is it just like it's another we have to figure out this is ubiquitous, it's in people's.

24:46
You can't take something that people use all the time in every aspect of their daily lives a lot more than engineers use slide rules back in the day and say, no, you have to, you may have to say pause a little bit while we figure out how we can guide you because this is something that can wreak havoc with your ability to learn or it can actually position you to learn more effectively than you currently learn.

25:12
But so we don't want to just unleash it without thinking about how to do that and how to guide students and how to be sure that they're, you know, following that guidance.

25:21
But those are those are the things that I think are really, really important there.

25:24
Thank you.

25:26
I'm going to do one more question if we can and then think we'll be out of time, Sir, thank you.

25:31
Right now, my understanding about like AI is about data algorithm, computing power.

25:37
But as long as you talk more, I just feel a little pause.

25:41
Do you think after five years or 10 years when AI keep developing, if lease pattern are going to be changed or switching to a new pattern, maybe having like a human brain connected to like algorithms more?

25:57
Or you think lease will still layer data algorithm and computing power?

26:03
Thank you.

26:05
I mean, I think data algorithms and computing power are going to be true for a long time.

26:10
You know, I think algorithms and algorithmic advances can change the computing power picture a lot.

26:18

And sometimes people are not paying that much attention to that right now because the algorithms have not been advancing as quickly as they might at some point in the future.

26:27
And, you know, the, the form, the, the mode by which humans and AI interact, you know, whether that's going to become more direct than than written and spoken language, I, I don't know.

26:38
It's kind of it's, it's certainly happening at some levels these days, but I I think that those things will remain very important.

26:45
So we still can use in these three element to like a thing.

26:48
Yeah, yeah.

26:49
Thank you, Dan.

26:50
Thank you so much.

26:51
No, thank you for having us.

26:52
Great to see you.

26:53
Great to see you.

26:54
Thank you all.