

Hello, everyone. Welcome to another episode of Q&A about future of science and technology. Finally back, doing these after... All sorts of hiatus.

Well, let's see. I'm seeing a bunch of questions here, one from Arthur, one from Satyaki. about recent advances in science and mathematics with LLMs. Navier-stokes proof.

Future of science, future of scientists, and so on.

The, and another one here about, also asking about Navier-Stokes. Boy, that's been well promoted.

And, claims about misalignment by academic mathematicians who claim AI companies will stall mathematics.

All right, well, I can talk a little bit about that.

So...

I just wrote something, actually, about the future of pure mathematics research. I just posted, a few days ago.

some somewhat Elaborate thoughts about that topic.

I think the thing to understand is when people say, look, AIs are taking over math.

You have to kind of understand what they think they mean by math, because some aspects of that statement make no sense, because a lot of

what math is trying to achieve has to do with humans understanding things, which is not what's achieved by an AI doing things.

But maybe we start off with the Navier-Stokes.

result. I mean, I haven't studied it in great detail. I've been aware of the problem for a long time.

Maybe I'll do a little bit of a science minute about what that's actually about.

So, in, sort of, traditional physics.

that kind of was the evolution of what folks like Isaac Newton did in the late 1600s. One thinks about using calculus

to... and differential equations to describe things in nature. So, differential equations say the rate of change of this thing is given by this.

Bob Battalion, And that's... that, and so then you say, well, what is the actual value of the thing if the rate of change is given by that?

So, a lot of, kind of, equations in physics, a lot of the way that the sort of traditional physics is set up, is in terms of these equations that relate rates of change and rates of rates of change, and so on.

And.

One of the famous such equations is the Navier-Stokes equation, which governs the flow of fluids, like water and air.

in certain regimes. They can't be going too fast, too close to the speed of sound, they can't be... one can't be looking at them if they're too dilute, and so on. But in sort of the ordinary range of things, these equations govern the behavior of fluids.

There are many mysteries about how those equations work.

In particular, some of the more dramatic phenomena in fluids, like the phenomenon of fluid turbulence, when a fluid flows quickly, it kind of curls up and makes all kinds of random-looking stuff.

That phenomenon has never been really fully understood from the Navier-Stokes equations.

And it's not completely clear that it depends only on the Navier-Stokes equations. One possibility is that the randomness and turbulence comes by sort of amplifying little molecular details of

what's happening at a very microscopic level. And we just get to see that macroscopically through some mechanism.

I don't think that's what's actually going on, although people have believed that for a while. I kind of think that the randomness that one gets in fluid turbulence is more like the randomness one gets in things like the digits of pi. It's something that's essentially algorithmically determined, but nevertheless, even though there is an algorithm for it.

And this is, again, my favorite phenomenon of computational irreducibility. You kind of can't know what the next digit will be except by explicitly computing those digits. It will look random to us.

So, in any case, the, There are these equations which at least approximately govern the flow of fluids.

The question is, if you just start the fluid off with some, you know, nice sort of smoothly, it's moving at some speed, et cetera, et cetera, et cetera, could it be the case that at some moment there's at least a little point inside the fluid where the thing gets so agitated that those equations no longer apply?

that effectively, for example, the fluid boils, or something like this, that somehow, just by sort of sloshing around the water or something, there's at least a point in the water where the thing is no longer behaving like a fluid which obeys the Navier-Stokes equations.

So that's kind of the question. Can you set things up to do that? Now, is that a question that is ultimately sort of a physics interest?

To some extent.

The thing that is happening there.

is if that were what you concluded, then you would say, well, actually, the Navier-Stokes equations don't really apply completely, which is what we knew already, because fluids are actually made of molecules, and the Navier-Stokes equations are just an approximation in a certain regime of size relative to the size of molecules and speed of flow and all that kind of thing.

They're just an approximation, and this would say, well, actually, there are circumstances that seem like the approximation isn't going to break down, but actually it does.

Now, there's an analogous case of this that is probably even more dramatic, which is the Einstein equations for the structure of spacetime, which are, again, a set of differential equations that say the rate of change of this, it's then in terms of length measurements in space, and so on.

the rate of change of this with respect to whatever is equal to whatever. I mean, it.

Now, in the case of the Einstein equations, the thing that was discovered, so the Einstein equations were originally written down, well, by Einstein, and to some extent also by Hilbert in 1915. By 1916, it was known that there were solutions to those equations that had singularities.

That there were situations in which you could say, I'm going to have a solution to the equation, but there's going to be some place where something goes to infinity.

People were very confused about what the going to infinity was all about.

There was the situation in which you could have kind of some kind of spherical thing where there would be what, there would be this place where the sort of distances in space-time, kind of become singular, and.

At a certain distance away from the center of this spherical thing.

Well, it was very confusing. Until the 1950s, people were very confused about what this was, even up to the 1960s. Is this actually a singularity in spacetime, or is this just a feature of the way

that one has coordinates for these equations and so on? What became clear by the 1960s is that the event horizon

Place where kind of light.

things... nothing escapes the event horizon. If you... if you start light off inside the event horizon, it won't escape, and so on. That...

that is not really a singularity in spacetime, that's merely a singularity in the structure of the coordinates and so on that you use in the equations. There is, however, even in that same setup, there is a genuine singularity at the center of what is then a black hole. At the center of it, there is actually the quantities that one is looking at in the Einstein equations actually go to infinity.

The curvature of space-time actually goes to infinity, at which point the Einstein equations that say the curvature is the rate of change of things is equal to this no longer apply.

Okay, but so... so the Einstein equations, if black holes form, the Einstein equations imply that there are singularities in spacetime at which the Einstein equations no longer apply.

So people, sort of.

were wondering about that, and there was this hypothesis that there would be no naked singularities. There wouldn't be a situation in which there was an infinity in this curvature of spacetime that wasn't hidden behind an event horizon. So, in other words, some physicist could go and say, I found this singularity in spacetime, but they'd never be able to tell anybody on the outside that they'd done that, because any signal they sent

Would be pulled back by gravity and would not escape the event horizon.

So until maybe 20 years ago, it was believed that there weren't naked singularities, that whenever, like, a star would collapse to form a black hole, that a event horizon would form, and that whatever singularity was at the center would be forever shielded.

It turned out

That there was some very elaborate construction that was produced that says, well, if you just add little puffs of matter, and just this way, in this kind of nested pattern of adding things, by very careful arrangement, you can set it up so there is a naked singularity. So there is a singularity in spacetime without the presence of an event horizon around it.

But the construction is complicated, and I don't think...

people believe that, in actuality, you know, you could do arbitrarily careful engineering, and you'd still never quite get it. You'd always be slightly off, and there'd be, you know, you'd off by some microscopic distance, and you'd miss the situation in which there is a true naked singularity. So it's kind of a... lots of angels on heads of pins necessary to set it up.

I think the Navier-Stokes... well, actually, to say about the Einstein equations.

The Einstein equations, people have tended to believe, are the ultimate story for the structure of spacetime, that it really is governed by these differential equations. I don't think this is right. I think that there is discreteness in spacetime, as there is discreteness in matter and fluids and so on.

And the true situation is that the Einstein equations are an approximation to an underlying discrete structure, so when you see a singularity, it's really just telling you, wait a minute, this description of things as a continuous material with continuous spacetime isn't really right. You have to go to the lower level, just like in a fluid, you might have to go to the molecular level.

Okay, so the Navier-Stokes story is a story of trying to do the same angels on heads of pins thing for the Navier-Stokes equations, another set of differential equations, very different structure from the Einstein equations, though some similarities.

And the thing that I have not gone through the details of this, that it's claimed has been shown, is that there's some way of arranging the angels, so to speak, on the head of the pin, so that with this definition and that definition and that definition, you actually end up with sort of continuous things that you're adding in, and it's all very smooth and so on, and then suddenly something singular happens.

And that's kind of the claim.

The,

The question of whether... so that's an interesting kind of thing, because in a sense, what you need to do is find a configuration of a fluid which has that property.

And that's a thing that, to some extent at least, once you've found it, you can kind of verify that it really behaves that way. That's not so trivial to do.

And it's not self-evident that that's a thing that one can just immediately say, oh, look, it's not like if you factor a number and you say, here are two factors of the number, well, you just multiply them together and see if it gets the number again. It's a more complicated story, but I think it's a story that is fairly straightforward to verify.

Now, you know, people have been looking for these kinds of configurations that lead to funky things happening in Avistox equations for ages. It is certainly an achievement if one has really been found.

And that's cool and interesting, and I think the way to think about it, and the way to think about a lot of these kind of LLM sort of breakthroughs in science.

is there are 3 or 4 million papers that have been published in the literature of mathematics, and what LLMs are good at

is kind of taking those things, and they've sort of ground them up and digested them in their training, taking, you know, a result from a paper here, a result from one there, a result from one there, and sort of mashing them together, gluing them together to say, and that means the following is true.

So, I see it as much more, and that's certainly how I use these kinds of things, as much more of a super way to search and fit together existing literature of mathematics than as something where you're sort of figuring things out from scratch.

And there's a lot of mileage to be got from sort of putting together a thing here and a thing there. You know, no human has read millions of mathematical papers, so if it's something which has to sort of span many, many pieces, that's a thing for which the AI has a tremendous advantage.

I mean, I remember back in the late 1970s, I was first using, sort of, computer referencing systems, where you could go search the abstracts of all science papers in an online system.

And I discovered a bunch of things by doing that. I never really thought until now, actually, that that was kind of an important secret weapon that I'd had in doing science, but I realized that it actually was, and it allowed me to make a bunch of progress in a bunch of kind of interdisciplinary areas that people hadn't really been able to tackle before, just because I was able to find things from different areas. I mean, the fact that other people didn't use those tools.

It's just one of those bizarre anomalies of history. People often don't use the tools that are readily available. I mean, I'm always, having spent much of my life building tools, partly for myself to be able to use, I'm always, if other people aren't using Wolfram Language or whatever else to do their work, and they're like, oh, well, I can put this together in this way and that way and that way.

But... and it's all very difficult, and it's like, you could just type in one line of code into Wolfram Language, and you get this done. And they're like, oh, that's too hard! And I'm like, yeah, that's silly.

But in any case, people, you know, one of the things I've noticed in doing science and other things.

is, you know, the people with the best tools tend to win, and tend to get further than anybody else. And I've taken that principle to heart, because I've been building this kind of tower of tools for the last 45 years or so, that I certainly find super useful, and I think lots of other people do too.

In any case, I think the thing... so to just say a little bit, and I talked about this at length, actually, in talking about the piece that I just wrote a few days ago, but, the, you know, sort of one of the pieces of the big picture is sort of what is science, what is math, you know, what are we trying to do with these

things. In... in science.

A lot of the story is, take what is out there in the natural world, and try and get sort of a human narrative for what's going on in it. Try and take whatever details are happening in the natural world, and try and set them up so they kind of fit into finite human minds.

That's the typical mission of science. Find natural laws, find things you can say, principles you can have, classifications of things, all those kinds of things. That's the... that's the thing, to go from what's out there in the world to what we can understand with our minds.

Now, there's a different objective, which is, for example, try and predict things for engineering or technology.

where you can have much more black box kinds of ways of looking at the natural world, you can say, look, I can predict

what this is going to do, even though I can't understand... I don't have scientific laws that humans will ever understand about what's going on inside there.

In science, that black box method can be useful, particularly for doing technology, for doing engineering, and so on. In mathematics.

It's not clear why that's useful. So, in mathematics, it's the same kind of story. There's sort of this... this very detailed world of sort of axiomatic structures and so on that's out there.

And the challenge is to take those detailed things and make them something for which there's a human narrative we can understand, make it into something where we can kind of fit it in our brains, fit it in our minds, and think about things. And that's kind of the main mission of mathematics, as mathematics has been done in the last century or so.

And, the, so, you know, to make something that has a sort of human narrative about what's going on.

Now, the other thing you can ask is, I just want to work through the math.

That's kind of applied math, more than it is pure mathematics. In applied math, the goal is just get an answer, because we're going to use it for something else.

you know, calculate where that spacecraft is going to go, calculate where... how some fluid is going to behave, those kinds of things, where it's a question of just... just figure it out, by all means necessary.

And that activity is one where you can... mostly, the biggest thing you can use is just computation.

You know, maybe you can speed that up a bit with AI, maybe the AI can figure out some regularities that were not already included in what you were doing, and let you jump ahead a bit.

You can find kind of a pocket of computational reducibility to sort of give you a bit of an acceleration of figuring things out.

That doesn't tend to lead to math of the kind of... that people are most interested in. It leads to this kind of, yes, this is true, but we don't really know why, kinds of results.

and You know, even when it comes to... sometimes it's just worthwhile to know what's true.

But much of the real sort of grist of pure mathematics and pure mathematics research has to do with the understanding of this sort of abstract structures and concepts and so on, that sort of show you this kind of humanized version of what's out there in the mathematical world, and kind of give you ways to understand why things are true, and so on.

I mean, I, I, I think, so...

you know, what I expect is that there's a fair amount of low-hanging fruit to be picked by just sort of connecting different things in the existing literature of mathematics and applying them to things that

had been sort of languishing for a while. The thing to realize is that the, sort of, the zeroth problem in mathematics is to figure out what question you're asking.

And that's often what distinguishes really great mathematics from just kind of okay mathematics, is was it a really great question? Was it something formulated in a really interesting way that really connected different kinds of ideas and so on?

Or was it something where it's kind of like a puzzle? Is there another number that satisfies this thing?

But just sort of absent any kind of bigger picture.

And I think the thing that, is... is really a very, sort of, human kind of question is, well, what should the right questions be? You know, what is the intellectual structure that we want to build that we want to sort of construct that increases, that sort of builds these sort of towers of abstraction and mathematics, what is that intellectual structure? And that's something that is really up to us humans to decide. It's, you know, we can build a tower in my work in studying simple programs in the computational universe. I've discovered many, many, many things.

Which are kind of out there in the computational universe, but not really, and they're kind of... yes, that's kind of interesting in its own right, but it isn't connected to kind of the main trunk Of what one thinks about in an area like mathematics.

I mean, there's a lot more to say about this, and I think I talked about it at some length just a few days ago, so let me not go into more detail about that here. But I think, as I say, there's, you know, the real thing that's most valuable is knowing what questions to ask.

The things that... the sort of the abstract construction of things that are for human understanding is really the essence of what good, pure mathematics tends to be about.

the, you know, the sort of connect the dots of things, certainly there's mileage to be got from that, and possibly there are things that are of important intuitional value that can be deduced by just joining the dots in the right way.

All right, let's, look at another question here.

So another question is from Anon is, what do you think is going to happen in the future of academic publishing?

That's an interesting question.

You know, academic publishing has... had quite an evolution. I mean, Back.

In the distant past, people just wrote books. When they thought they had something to say, they would write a book.

There were lots of books written, many of them were no good. Some of the good ones have survived.

In the 1600s, there started to be things like Proceedings of the Royal Society, where people would come and, I guess, hang out in, actually, a building in, I think, probably even the same building for a long time in London.

And, would have, kind of, people would, would...

Kind of give talks about things they'd figured out.

And that became something that got printed, and the things that were printed in the Proceedings of the Royal Society read a lot like a typical blog post might read today. Quite chatty, quite a lot of, kind of, whimsical content.

Of, you know, oh, I ran across this thing, and that made me wonder about that thing, and so on.

Well, then.

By the 1700s and 1800s, particularly 1800s, there were starting to be much more kind of buttoned-down academic journals, and a lot of notable people, like in mathematics and other places, sort of started journals and published things of their own and other people's.

And then... by the...

By the 20th century, particularly by the time of World War II, there have been... there were lots of, sort of, fragmented pieces of, you know, publications by the such-and-such Academy of Sciences, publications of so-and-so's journal, this university, and so on. Lots of different venues where people published, kind of, academic material.

And, I think it was shortly after World War II that.

People started, I think a chap called Robert Maxwell, who was somewhat notorious for many reasons, started doing this particularly, just bought up all of these kind of obscure journals and then went around to the sort of emerging and growing university.

library system and so on, libraries, and said, if you want to be a self-respecting university library, you need to be a subscriber to all these random journals from this place and that place and the other place.

And it was a good business. And so then a bunch of other people got into the act, and there started to be this sort of huge expansion in the kind of academic journals that existed. And then there started to be pecking orders for those things and so on.

And at some point in the middle of all of this, well, by the time I was doing science, at least in physics, by the mid-1970s, there was a whole very active

ecosystem of preprints, where people would just print up, you know, a few hundred copies of a paper, and send it around to all the standard universities that mattered, and then

they would... people would send out preprint request cards saying to the authors of papers, saying, could you send me a copy of your preprint? Because there were aggregations, online aggregations, for the field I was involved in, particle physics. There was one run by the Stanford Linear Accelerator Center, one run by CERN.

And these were things where you could

This was by 1976, for sure. You could log into the ARPANET, which was the predecessor of the internet. You could access these systems, on these, at Slack and at CERN.

And you could see what preprints had been, had been sent in. I think there was also a physical mailing of the list of preprints.

So then, if you were a kid like me, when.

I was in high school, I wrote a bunch of physics papers, and they were on all these lists and so on, and so I would get

a large collection of the preprint request cards from all over the world. I was at boarding school, and so all the other kids were super amused to see all these cards coming in addressed to Doctor, which I wasn't yet, you know, S. Wolfram type thing, asking for copies of these papers that I'd written. I mean, I got hundreds and hundreds of these cards, which I've kept, it's kind of interesting.

But that was a whole ecosystem of preprints. And then what would happen is, you know, something would be circulated as a preprint, and then eventually some journal would stodgily accept it after six months or something. If it was actually an interesting paper, it was always hard to get it published in a journal. That's why, in the end, I became an editor of a bunch of journals, and I started my own journal, then I stopped publishing academic papers in 1999. 1986, because I decided the whole ecosystem was kind of broken.

But in any case, the, the, at.

So, preprints were a way of distributing information, that existed certainly from the 1970s in areas like particle physics, and that's kind of spread. My friend Paul Ginsburg started Arkiv. When was that? That must have been sometime in the 80s, as a sort of preprint distribution system that was an online preprint distribution system.

And then, you know, things happened in journal publishing, like... like, a lot of the journal publishers were... decided to sort of make the prices of their journals go up and up, which was kind of a...

Kind of a crazy thing, because

Back like in the 1970s, when you would send a paper into a journal, the journal had to do real work.

You sent in a typewritten thing, maybe even handwritten in some cases. They had to typeset it. They had to do copy editing, et cetera, et cetera, et cetera. And they would send it out for peer review.

to random, anonymous reviewers. I have to say, after a while, I insisted on not being anonymous, because I figured if I was going to go to the trouble of commenting on somebody's paper, I might as well have them know who commented on it, and I wasn't shy about saying, I don't think this paper is much good, the authors should do this instead of what they've done, type thing. I think I was quite helpful in that respect, but I didn't really believe in this kind of anonymous hiding behind kind of anonymity peer review mechanism.

In any case, and that mechanism became sort of deeply corrupt over the years.

It's, but in any case, the crisis right now is about none of these things. The crisis right now is that AIs can write papers which are pretty good relative to human papers. I think the, really the issue is that the low-end

of human papers, human written papers, definitively overlaps with the high end of AI-written papers.

And what does one do under those circumstances? Not least because the number of AI-written papers can be very large.

You know, one has to really zoom out and ask, what's the point of this academic publishing enterprise? You know, what is one trying to achieve with it? What does... and I should explain a few other things. I mean, one of the horrors of the academic publishing system is that the people Write papers, and then nobody reads them.

And you know the typical paper.

it's just not read by more than a very tiny number of people. Often, nowadays, there are lots of authors on a paper, and I'm skeptical that all the authors have even read the paper, let alone anybody else.

I will say that for the things I write.

You know, I just post them on my writings website, and, you know, an awful lot of people, it can be 100,000 people, will read some... something there, and which is cool.

But it's an unbelievable factor larger than the distribution channel that you get in journals that go to university libraries and so on.

I mean, I will say, back in the day when I was a kid, it was really cool to be able to write a paper, you know, when you're 16 years old or something, you write a paper, it gets published in some journal, you can go to any university in the library in the world, and you can find your paper. I thought that was really cool.

Of course, in today's world, you can put up a website, and from any random, you know, device anywhere in the world, you can see a website. So that thrill has kind of transitioned to something different now.

But in any case, the question is.

In a world where it's possible for AIs to generate paper-like things, what do we kind of do with academic publishing?

I will say another thing.

At any given time, in... there are people who are...

fully trained and themed academics and so on, publishing papers in academic journals and with a whole club of people interacting in some particular field. And then there are kind of amateurs, trying to contribute to these fields. In some fields, it works better than others.

In kind of advanced physics, it works quite poorly. And you know, it's a it's kind of an unfortunate thing that, well, in some ways it's good, in some ways it's bad, that I don't know, somebody like me will get at least five emails a day of people sending me kind of theories of physics.

It's not completely clear what I'm supposed to do with them, other than you have a theory of physics, so I want to send you my theory of physics, which doesn't really quite compute as far as I'm concerned. In any case, in the past.

Those documents were very... it was very clear what kind of a thing they were, and what they could and could not achieve.

Because mostly they were using, kind of, high school-level knowledge of physics.

Which... is fine.

for studying some kinds of everyday physics. It's not fine if you want to study, kind of, foundations of physics, because high school-level physics doesn't include everything that got figured out in the 20th century, basically, about physics. And a lot got figured out in the 20th century. So if you are trying to base your theory of fundamental physics on kind of high school-level physics, it's just not going to work.

If you're trying to base your theory, which unfortunately is not usually what you see, but if you were trying to make a theory of kind of how, I don't know, different sizes of marbles settle in something when you're shaking it around, things like that, that could be high school-level physics, and you'd come up with a conclusion.

But you're not going to come out with a conclusion about the foundations of general relativity if you don't know what general relativity is.

And so.

Back, you know, only a few years ago, you could kind of easily distinguish Kind of the... it really doesn't have a chance from the, oh, that might be interesting. by the kind of almost texture of what was being said. It's kind of high school-level physics, not gonna work, etc. Nowadays, your average LLM can happily write what looks just like absolutely fine, professional, top-of-the-line physics.

You know, I think I was mentioning.

It was just an event, oh, a few days ago, about formalization of math, where it's kind of a hackathon-type event, where people were presenting, kind of, things they'd done to solve math problems.

The thing that was really interesting was not a single person there was a mathematician by profession, and people were presenting these things that looked super impressive.

It might have actually been quite interesting, but they were saying, well, you know, the AI produced this, I really don't understand anything it said.

That's a very funny situation.

Because, in fact, some of those things might have been interesting, things one should read in preference to things that some humans have produced. But what do we do with this? I mean, in other words, do we... do we just generate

you know, encyclopedias and encyclopedias' worth of things that seem kind of interesting, and then wait for humans to show up and read them? You know, is that the right model? I don't know. I think that... that.

It is something where, somehow.

There has to be some kind of human commitment that somebody really looked at this and found it interesting. It's kind of like, do you want to be flown by a pure AI, or do you want a pilot in the front of the plane flying you? It's good to have some level of human commitment that somebody is there with you on the plane, rather than just, oh, you know, some company set up this AI and it's flying the plane and good luck to you, type thing.

So I think there's some level of, sort of, the need for human commitment in getting you, the reader, to commit to spending time, kind of, reading the thing. Like, if no human touched it, no human committed their time, why should you commit your time to reading it? Now, you might miss some great thing that an AI discovered that it told, you know, was out there.

And it would have been super interesting to you.

I tend to think.

Bobby R. that in that case, that the right way to think about that is then you should custom get the paper written. That is, if you have something you're wondering about.

then you can just tell the AI, write me a paper about this, and then it will do whatever it can do in that area. Whereas the things where it's just like, hey, this is an interesting thing, are things where one sort of, I think, perhaps needs a human commitment to justify one's effort in actually reading the thing.

And... but, but...

You know, then there's another problem, which is, in the world of academia, kind of what publications you got and what you wrote and so on, that's important for assessing, you know, hey, is this a really good academic or not a good academic?

And, in fact, after I was in academia, there started to be all of these kind of metrics for, you know, not just how many papers did you publish.

But also, what was the quality of the journal? What was the impact factor of the journal? Which is roughly determined by how many people cite papers in that journal, and so on. Then there

started to be alternative metrics, which is like, yes, well, nobody read the paper, but they watched the podcast about it.

And that's sort of an alternative metric for seeing what the impact of what you were doing was. And I don't know exactly how it's worked today, but, you know, it was the case that in some, you know, like, universities, it's like, is this person going to get promoted from, you know, assistant professor to associate professor? Well, let's look at their metrics.

You know, how many citations did their papers get, etc, etc, etc. And of course, citations are a whole complicated mess, because you can cite yourself, or you can just have a clique of friends cite you, and then you have a little citation loop, and you can kind of spin up lots and lots of citations to something.

You know, is that less worthwhile than a very broadly distributed citation network?

I mean, these are the same kinds of questions as one might ask about websites, and that was kind of, well, that didn't... things like the PageRank algorithm for, sort of, assigning credit to websites, tried to deal with these questions about citation and so on. It was an early methodology used by the Google folks, for example, in assigning kind of rankings to websites. I don't think it's used anymore. There are lots of other signals that get used instead, but it's kind of a conceptually interesting thing.

of sort of how many people link to this site. Well, if a lot of people link to it, it must be a, you know, a good site, so to speak. In any case, so, you know, so part of the thing is, in a world where people can just sort of splurt out AI-generated papers with wild abandon.

What does that mean for, you know, who who gets to say, I'm very impressive and some other person doesn't?

John Downton, Now, at some level.

If you have two people, one of them's generated 200 AI generated papers and none of them are any good, and another person's generated 200 AI generated papers and some of 'em are kind of interesting, then probably the person with the kind of interesting papers did something better than the person without them. Or maybe it was just the luck of the draw, the random numbers that the AI came up with at the moment when it was generating those outputs weren't as felicitous.

For one person as the other.

But...

You know, I think that the, I mean, one sort of bizarre possibility is that it's kind of the, you know, people generate gazillions of AI-generated papers, which are read by AIs, and those AIs rank those papers, and so on.

I mean, another thing that people... so one of the challenges is things like the Archive preprint server, I think, is getting deluged with AI-generated papers.

and

As I say, the low end of human papers and the high end of AI papers probably overlap. So it's like, well, what is the criterion?

for, you know, having this paper there, or not having this paper there, because it could be that it's actually fairly interesting, even though it was written by an AI. Could be it was written by a wonderful human, but it's completely boring, and nobody's ever going to want to read it.

So...

you know, I think that's some... but the one practical challenge for any of these places is what has been the case, is that humans have been trying to assess these papers and say, hey, do we

want to put this here, do we not want to have this here, et cetera, et cetera, et cetera. As I say, that system became deeply corrupt.

So I will... here's a funny story about Archive, actually. So, I was uploading a bunch of things that I write to Archive, you know, just turning them into PDFs and uploading to Archive. And, you know, we'd sort of routinely been doing that.

And I wrote something about the foundations of the second order of thermodynamics.

And, you know, we routinely uploaded it to Archive, and it got bounced.

It's like, what the heck is going on? This is a, you know, it's just a physics paper, it's not, there's nothing particularly... I thought it was pretty interesting, it's something I thought about for close to 50 years, and so it's kind of like, you know, it's got real content. But it's like, oh no no, you know, it was sort of bounced by the system.

So, upon talking to the people involved with Arkiv and trying to track things down and so on, I finally figured out what had happened, which is really quite amusing.

which is

In one of the types of things that sort of nobody believes in are perpetual motion machines.

People say, oh, you can't have something which will go and just, you know.

turn heat into mechanical work forever, or sort of just generate energy from nothing. Just impossible. You know, we proved that, people think, in the 1800s.

Well, actually, no such thing, but that's another matter. In any case, so...

The thing that is supposed to sort of establish that perpetual motion machines of the second kind are impossible is the second law of thermodynamics.

So, needless to say, by the time I'm writing a book-length description of the second law of thermodynamics, I talk about perpetual motion machines. And, you know, I explain the extent to which they're impossible or not impossible, and, you know, what it all means.

And... but the system that, you know, Archive have been.

kind of overrun with submissions, and so they had keyword searches that were kind of, sort of bouncing papers based on keyword searches. And so my...

very staid, sort of, basically monograph about the second law was kind of, oh my gosh, it's talking about perpetual motion machines. Help! We can't have perpetual motion machines here.

So, now, the unfortunate part of that story is because there's sort of no humans involved in the whole process, that never got resolved, and I still haven't uploaded, I still didn't restart uploading anything to archive, so... I don't know whether I'll restart it or not. It doesn't really matter much.

I think people can find my stuff perfectly easily.

But, I thought, sort of as a courtesy to the community, I should do that.

But that was kind of a, I'm not going to bother with this.

But in any case, the So... Now the question is.

The... I could tell another story.

Oh.

person... not very satisfactory person, who was working with us for a while, was, had done the heinous thing of, producing a document that was intended for outside distribution, which they just wrote with ChatGPT.

And it was terrible, actually. And I was sort of really annoyed about that.

And then this person produced another document, which had very weird wording in it.

And I'm like, you didn't write this one with ChatGPT as well, did you?

And the person's like, no, I promise, I really, really didn't write it with ChatGPT.

And, I was like, okay, can you prove that?

And said, well, you know, this kind of turn of phrase, which is very odd, was something, you know, I used it in my PhD thesis, so I opened this person's PhD thesis, and sure enough, the very first page I see has that weird phrase on it.

So they, they escaped that, but if you'd had a kind of, you know, me as an LLM detector failed in that case. I thought that document was written by an LLM. It actually, presumably, wasn't.

Since the thesis had been written long before LLMs existed.

But in any case, the, so then...

sort of, there's this question of, you know, what... how do you... how do you deal with, sort of, the human aspects of, you know, the humans writing papers, and what are they really doing, and so on. And, okay, let's talk about a different issue.

Another issue is, what tools helped you write your paper?

So for the last 40 years, lots of people have been helped by Mathematica and Wolfram Language.

Sometimes they mention it, sometimes they don't. The conventions are different in different fields. In things close to biomedicine and so on, there tends to be always a kind of methods section where they mention things, even when it's like, I don't know quite why you mentioned this, because it doesn't make any difference. The answer will always be the same, however you did this.

But in any case the people

You know, it's like, people are using tools, do they even mention what tools they're using?

And does it matter if they mention what tools they're using?

So, it would matter if they said, you know, oh, we used Wolfram Language to do this, and then it turned out there was some horrifying bug in Wolfram Language. Thankfully, I don't think anything like that has ever happened, not that there are no bugs, but something where somebody published some big result, and it was wrong because of that. I don't think that's ever happened.

But, you know, that will be a reason to say, well, I used... I did this, because then people can kind of go back and say, oh, if that was invalidated, then this.

The real thing that people don't do is they write a paper. They say, yak, yak, yak, you know, we did this, we did that.

How do you know if this or that?

you know, really follows. You really have to reverse engineer the paper and redo the computation, or whatever it is. If there's a physical experiment involved, that's a yet different story. But even at the theoretical level, it's like, you have to go in and reverse engineer what was said in the paper, and go and sort of do the computation again for yourself.

Now, we've had a solution to that forever, which is actually include Wolfram Language code, and we've had notebooks for... since 1987, and we've had, you know, cloud ways of doing that for, oh gosh, more than a decade, well over a decade by now, where you can

You can... actually, more like two decades, where you can have something where the full kind of computational language code, self-contained computational language code, that does the thing that you talk about in your paper is right there.

So when I write things for the last many years, any picture that's in what I've written, any computer-generated picture at least, you can click on that picture, and you will get code, Wolfram Language code, that in a self-contained way will regenerate that picture.

So that means that you can kind of take whatever I've done and build on it, because you have this kind of solid kind of representation of what's being said, and you can kind of build on that. And

as a practical matter, at our summer programs and things like this, people all the time are taking sort of code from things I've written and building from that to do the things that they do. So, I've, so this is another dimension, is kind of the, the sort of, the backup code that exists. Now, in physical experiments.

You know, sometimes if you're out, you know, in the field, and you see a lemur doing some exotic thing, well, you know, maybe you took a video of that lemur, so that's a... that's something, but, you know, there isn't a repeatable lemur trick to be done.

In other cases where it's a lab experiment, you could, in principle, repeat it.

And, for example, there's a company I've long been involved with called Emerald Cloud Lab, which has sort of developed this... this automation for hundreds of kinds of biology and chemistry experiments, where you just give a Wolfram Language script, and it will go and, and actually execute

the experiments you specify in a lab with machines, robots, some humans, etc, and carefully record every detail of what happened, so that, essentially, you have sort of complete reproducibility of the physical experiment.

So in addition to having, sort of, the computational language on the computation side, you can also have, sort of, reproduction physical experiments if you use that cloud lab.

So that's a... that's a way that you kind of build up science in a more solid fashion than just, oh, I wrote it in a paper, and it's just the written words that carry what I'm trying to say.

I will say also that I've been a big proponent of the idea of computational essays, where the computational language code is part of what is carrying the message.

So in addition to text and math, there's also computational language, and that's part of sort of explaining what you're talking about. And I think that's another important thing.

that, it's a bunch of work to do all these things that make your paper kind of really clear at that level, clear computationally, and able to be reproduced. It's work.

And the ecosystem of academia does not particularly reward that work at this point.

It's better to publish another paper than to make the paper you have be one that can be reproduced. It's just, you know, there have been many government

Sort of initiatives going back decades, actually, to say.

you know, oh, you know, you really should upload your raw data, you should do this and that and the other, but it really hasn't worked. And, you know, people will upload terribly cruddy spreadsheets where nothing can be extracted from them and so on. And, you know, it's more like we're doing it for regulatory compliance than we're doing it because we actually want people to use this.

I think also, there tends to be a dynamic that if somebody has written a bunch of code, gotten a bunch of raw data, they're like, I want to write another 10 papers with this. I don't want to spread this information around for other people to do this. I want to keep it for myself.

I have to say, my approach to this tends to be that I try to... well, first of all, I end up doing very open science, but these days, but... but, has always tended to... my approach has tended to be, given a particular domain. I have good enough tools that I can kind of mine out

the most obvious stuff from that domain, and so I'm not... I'm not into the writing a series of 10 papers. I've never done that. I really wouldn't like to do that. It's much better to have this one integrated thing that can communicate what you want, and it's one place to read, and so on, rather than the sort of serial publication of lots of different, oops, I did this now instead of that.

Now, I have to say, the limiting case of that was my book, *New Kind of Science*, which took a decade to produce, which was, as far as I was concerned, from a human point of view, a bit too long.

But

In any case, the idea of, you know, it's more valuable if people can reproduce it, that idea hasn't really taken root. I mean, the systems that exist for kind of rewarding things in academia don't really reward that properly.

So

Okay, so the question then is, well, okay, if AIs are writing all these papers, how do we know, you know, what do the humans even do? You could say, well, send your prompts.

But of course, one of the problems is that whereas with Wolfram Language, if you type a piece of code.

then... and then run it, you'll always get the same result, assuming it doesn't have explicit random numbers in it. Then you can seed them.

But in any case, whereas with an LLM, that won't happen. You feed it the same prompt multiple times, you'll generally get different results.

And that.

So it isn't even good enough to just say, hey, give me your prompts, because you couldn't reproduce something from those prompts alone.

So then...

You can say, well... I was thinking about this, actually, of what can you do? And actually, there is one answer.

Which is something that I suppose I've been involved in quite a bit, which is, you can livestream what you did.

Or you can at least record what you did.

And I've done a lot of livestreamed live science, so to speak. I've also recorded a whole bunch of unbelievably boring, in many ways, video work logs of just work that I'm doing by myself, typically, just figuring things out, doing research, writing things, and so on. I just record my screen.

And we upload those video work logs, and, you know, sometimes they're useful. It's like, how did... how did you get, how did I get

To some particular result. Well, I can watch and see the actual second where I typed that word or that number into the document. I could see what what was going on. Why did I do that? We can get that information.

It's very similar to what's happening with, like, Emerald Cloud Lab, where, you know, every kind of little motion of a pipette is captured on video, and you can kind of rewind exactly what happened. It's kind of the, for, for, kind of.

you know, FDA-style folk, it's kind of a dream situation, because you really get to see everything about how something was done. So there's a question of, could you do the same thing for kind of theoretical science? Could you do the same thing where you're just sort of recording, the, you know, recording everything?

So it's kind of like, you want to know what happened? Here's what happened. Here's what the AI said, here's what the human typed, here's everything that happened.

Now, you can imagine a world where people do that, and where then, if it's like, well, how, you know, how impressive was what this person did? Well, you get an AI to go in and review the

video and come to its conclusion. I mean, in a sense, this is similar to what we see happening in, kind of, student

kinds of things, where it was the case that, you know, you can have a take-home homework or something, or take-home exam, and that doesn't really work anymore, because people just type the things into AIs, and sort of everybody could get a perfect score. The only people who don't get perfect scores are the people who are actually learning things by not typing everything into AIs. So you get this ridiculous inverted situation.

But...

So, you know, one thing people do about that is they have, you know, in-class exams. Another thing people are starting to do is to have oral exams. Well, they've had oral exams for a long time with people. Oral exams with AIs is kind of a new thing.

Because, you know, assume that you actually have the live human in front of you.

then knowing whether the human knows this thing, if you just ask them and they respond in real time, then you have something going for you. I mean, of course, you can imagine weird situations where people are freezing the software in some way so that they have the extra 5 seconds or 2 minutes to go look something up. I don't know what kinds of craziness could happen there.

But that's kind of a...

a, a trend there. I think in terms of the, sort of, the creation of science, maybe this sort of crisis in academic publishing, maybe it has a good feature, because, let me mention another issue.

Which is that in fields like mathematics, there's a tremendous tendency to just write a paper that just makes a statement. Doesn't say why you're looking at this, doesn't say why anybody would care, doesn't say how you came to think about this, what your thought process was in coming up with a conclusion. It just starts

You know, my favorite line sort of starts, you know, "Let F be a field."

And then let this be that, blah, blah, blah, blah. And it's just pure kind of it's very inhuman, it's very kind of just very abstract, and it's like, and if you ask me, why did you do this piece of work?

If you ask them as a human that question, they'll probably tell you something that's much more interesting than what they said in the paper, where they just stated the facts, so to speak.

So...

it wouldn't be a bad thing to have people do more in sort of explaining why they came up with something, and maybe how it relates to things that have gone before. I mean, people are very keen on... just put those citations into the paper, and people... one of the things that.

It was kind of a pet peeve of mine when New Kind of Science came out. This book, people were like, oh my gosh, it has no references!

Well, I didn't put in a few 10,000 or so, you know, citations to papers that would have required somebody going to a library to find them and so on and so on and so on. Instead, what I did, which took a hell of a lot more work.

was I wrote, kind of, detailed histories of the various things I was talking about, which often involved going and actually doing primary interviews with people to find out what had happened, and so on.

To me, that's a much, much more worthwhile thing, and it's something I've continued and expanded on in the writings that I've had in more recent years. The idea that you're really contextualizing things with the history of what happened. You're explaining how did these ideas get to be come up with.

And you know, and who were the people? You know, hey, sort of kudos to so-and-so for having come up with this thing.

It's important for the ecosystem, so to speak, to recognize when work has been done, because it's kind of part of the kind of part of the payback for putting all the effort in to do something cool, is people know it was you who did it.

But.

You know, that's, so...

In any case, the, you know, I'm not such a believer in the, you know, just the raw citations. I don't think that's the useful thing. I think the useful thing is the actual history and context and narrative behind whatever was discovered.

giving, sort of, credit to the people, kudos to the people who did things, but also explaining how things fit together. Because quite often, among other things, even if you... even if all you want to do is kind of, quotes, give credit to people.

It's a question of how that really works. I mean, like many, many, many times in my life, I've sometimes somewhat offhandedly suggested ideas to people which became big things.

that, you know, they sometimes said thanks in the paper they wrote, and sometimes they didn't, sometimes on purpose they didn't, which is always annoying, although that's, I'm happy to say, the minority case. But skull... you know, there's skullduggery there. But.

The, but mostly, you know, that doesn't really have there's no real mechanism for saying, hey, this person offhandedly at lunch with the people who wrote this paper suggested the idea of this paper to them.

And, you know, it's kind of great. You know, I'm... I probably enjoyed the lunch being able to suggest it, and they had a good time figuring out the consequences of something. But, you know, there isn't a good mechanism to say, hey, extra citation, you know, plus one citation, so to speak. For that, you know, for what might have been an important idea there.

But in any case, so I mean, I think

This question of, so what did the human actually do?

Is a question that, in the end.

If you livestreamed or recorded every detail, every keystroke of what was happening, you really know what the human did. Then the question of assessing whether that was significant, that's a different issue, but at least you then have the data.

I think also this idea of live streaming science, for example, is an interesting and important idea, because it kind of opens up something which has been a very, you know, and the scientists go off and discover this, and they're wearing white coats and nobody knows what they do, type thing.

That's... I think that it is good for science

To have a bit more visibility about what science actually consists of, and how people actually come up with discoveries and so on.

And I think that's, so, you know, having that as a mechanism, that perhaps is a mechanism that's introduced for the sake of dealing with, you know, who wrote this in AI or human, but has these good side effects as well.

I don't know what gets to happen.

to the traditional academic publishing venues, including the journal I started back in 1986, Complex Systems, to publish things in that sort of then emerging area.

You know, I think That the issue is, again.

Like, you can put anything out there, just like you can put up any website, but the issue is, is anybody going to actually go there?

And that's the thing, I don't know what the... what the appropriate sort of ecosystem is for having that happen, and to what extent it, you know, does it matter that the paper you just read was written by an AI if it told you useful things? You know, do you.

So anyway, a few thoughts about that. I think there's more to figure out about this, and I'm hoping to be able to do that.

And maybe, maybe we'll even come out with some, some product related to academic publishing, we'll have to see.

Let's see.

Those comments about,

Jamie is asking, what parts of the traditional journal model are still worth preserving? It's a reasonable question.

You know, something that is sufficiently... Curated.

is going to be useful. The... if somebody gives me you know.

A newspaper that has 10 headline stories.

is more useful than a data dump that says a million things happened yesterday, at least to me.

Now, I might not agree with the newspaper's picks.

But it is useful to have that curated list.

curated by, you know, an entity where I think I understand where they're coming from, so to speak, and maybe it's aligned with things I'm interested in. So I think that the sort of the curation aspect is useful. I'm not sure that the...

You know, the thumbs up, thumbs down function is, I think, so corrupt at this point that I don't even know if it's worth... I don't know how worthwhile that is. I mean, remember, if you have curation as a primary thing that's important, you could say, I put out this thing, and it's on 30 curated lists.

Or it's on one curated list, or whatever. And that's kind of a proxy

A more visible proxy for kind of the peer review mechanism, because the curators of those various lists are going to be, do I want this on my thing or not?

And it's a much more even-handed kind of market, because the curators want to put things that are good on their curated lists, so people keep coming back to those curated lists and don't just say, oh, that person just puts their friends up there, and they're no good in the end.

So I think, you know, and on the other hand, you know, it clearly is a benefit to be in the curated list, so I think that's a more stable thing, and probably a good idea in that case.

you know it.

Things that journals used to do.

Bobby Reddy, involved sort of more microscopic sort of tweaking of things, like typeset that formula, then it's easy to read. You know, I always liked the fact that when I was typing papers on a typewriter.

And then, you know, and elaborately moving the, you know, things up and down for subscripts and superscripts and occasionally writing in a letter because I didn't exist on the, you know, on the electric golf ball typewriter thing or whatever else it was.

You know, I always thought it was a big uptick when the paper got published in an actual journal with actual typesetting, all looking beautiful, so to speak.

That went by the wayside when, you know, things like tech and so on made it sort of routine to get nice-looking typesetting. There is still a frontier there, which is if there's code in the paper, there's a question, does the code work?

And it's something I tried to persuade publishers of for probably 35 years or more.

that a place where they can add value is in doing QA for that code. So, in other words, if you find a paper that is published by such and such publishing organization, it's like, okay, I can trust that the code is going to work. Because it takes effort to get the code to work, and to keep the code working, and so on.

But I don't think... I feel like it's one of these things where... where, sort of, publishing is an old industry where it's kind of, like, what's the least amount of work we can do and still be able to charge for this journal and so on? So I don't think that's been properly developed.

Oh.

Let's see.

I think Alan here is commenting on something I was talking about a while ago here, saying when you get a number turning out to be infinite.

Watts have been finite.

shouldn't you admit that there's something wrong with your equations, rather than hope you can get a good theory just by doctoring up that number? I mean, there's a long story of these kinds of things. Usually.

you know, infinities from coming from equations usually are signs that there's some thing that isn't really governed by those equations, and then you have to ask, well, what else is there? And that becomes interesting. So, yes, that's.

It's a reasonable clue that there might be something wrong with your equations if their infinity is produced. There are cases, like the idea of renormalization, where, it's kind of.

there's sort of infinity minus infinity to get something real, and there's really are infinities and so on, and it's a slightly more complicated story, which maybe I shouldn't talk about today.

on.

Well.

Let's see.

Q. Hug. I'm not sure.

Hag is asking, can you reverse engineer the code to detect whether an image is AI-generated, saying particularly relevant in the UK sovereign AI defense sector right now?

I think it's basically impossible.

to tell whether some collection of pixels was AI-generated or not. There may be clues.

The a particular model may imprint some watermark.

on, or the harness around the model may imprint a watermark to tell you this was AI-generated. I think the efforts to make things be visibly not AI-generated mostly operate by picking up other attributes of

From the camera, so to speak, when the picture is taken.

I think that after the fact, if you didn't keep that additional information, I think it's pretty hard to tell where the picture came from.

Except possibly with some heuristics.

But I don't think, oh.

Let's see.

Well, a bunch of other interesting comments here, but I think I need to go. Unfortunately.

Well, a lot of interesting questions. So I look forward to addressing these another time.

But, but now... It's,

Bye from me. Look forward to doing more of these live streams.

Talk to you another time. Bye.