# Let's Talk AI with Mike Hinchey

Mike Hinchey[1,2] and Barbara Steffen[3]

[1] University of Limerick, Department of Computer Science & Information Systems, Ireland,
[2] Lero-the Science Foundation Ireland Research Centre for Software, Ireland,
mike.hinchey@ul.ie
[3] METAFrame Technologies GmbH,
barbara.steffen@metaframe.de

***"We've made amazing advances in technology, especially in the last 20 years. Much of this is due to greater processing power, cheaper memory, and advances made over decades that now can truly be exploited. We need to ensure that AI is used to advance technology and for the good of everyone, without discrimination of any sort."***

---

## The Interviewee - Mike Hinchey

**My Personal AI Mission:**
To promote the appropriate, sensible, use of AI and to educate the public that they cannot simply rely on AI just because someone calls it AI, especially when much of it is merely automation.

---

## My Takes on AI

**Artificial Intelligence:** Doing by computer what would require intelligence of a human.

**Trust:** That reasonable faith can be placed in a computer system to work correctly, safely, and securely.

**Explainability:** Describing the steps taken by a system to justify the decisions or choices that have been made.

**Essential Elements of Human Capabilities:** Creativity, experience, culpability, sentience.

## The Interview

**Barbara** *Today I have the pleasure to talk to Professor Mike Hinchey. Could you please briefly introduce yourself and your relation to artificial intelligence?*

**Mike** Sure, thank you. My name is Mike Hinchey. I'm a Professor of Software Engineering at the University of Limerick in Ireland and I'm also the former director of Lero, the Science Foundation Ireland Software Research Centre, and former director of the NASA Software Engineering Lab. My work is somewhat on the peripheries of artificial intelligence because I don't directly engage in AI research, but I work in the area of autonomous space exploration missions which heavily utilizes AI techniques.

**Barbara** *Are there one or two research questions related to artificial intelligence which you are currently addressing?*

**Mike** Yes, indeed. We're looking at how we can develop self-directed software-based systems for space exploration. These systems need to be able to make decisions autonomously, primarily to protect themselves and maintain resilience during long-term missions in space. We're exploring how to build in a certain amount of self-management to ensure these systems make wise decisions and don't endanger the mission or anything else. This brings us into the area of explainable AI, which deals with understanding why decisions are being made, and of course, responsibility and responsible behaviour.

**Barbara** *Did you see any great changes and improvements in the last years?*

**Mike** Certainly, we've seen tremendous advances with large language models and the benefits of greater processing power, which allows us to do much more with big data and machine learning. We now have access to various data repositories. People still complain it's not enough, but we're making progress. My main concern is that the media often labels everything as artificial intelligence, when a lot of it is simply automation. It's many of the same techniques we've used for decades, just benefiting from better processing power, development tools, and data, which leads to improved results. But it's not artificial intelligence, it's automation.

"We shouldn't be trusting things where we don't understand how they do it."

**Barbara** *Can you elaborate on the distinction between intelligence and automation?*

**Mike** Certainly. Intelligence is something we find hard to define. Roger Penrose suggests intelligence requires sentience, so we need to be aware that we exist and be able to understand and think [5, 6]. That's true intelligence. Artificial intelligence attempts to replicate that in some way to provide its benefits and effects. But much of what we're doing and what the media reports on are things that we've done for many years. We're following algorithms, using data, applying techniques, and running computer programs to do things that don't involve

intelligence or even mimic intelligence. They just automate tasks much more efficiently than a human can. And that's not AI, that's automation.

**Barbara** *In your opinion, what role does trust play in the context of AI?*

**Mike** Trust is a significant issue. The media often portrays AI as something that's going to take over the world or rule our lives, which causes unnecessary fear. In reality, we have systems that are often mislabeled as AI being used to make critical decisions about who gets into college, what prisoners get parole, and who goes to jail. These decisions are based on algorithms written by humans, which means they have human biases, intentional or otherwise. We don't understand how these algorithms work or their details. If we had that understanding, we would have a greater degree of trust. This becomes even more important when dealing with machine learning. If we don't provide people with information on how decisions were made, we won't have trust in these systems. And yet, we're using them for critical things and trusting the decisions they make. It's like Asimov's laws - do no harm [1, 2]. We don't have that trust because we don't know what decisions are being made.

"[...] we do need to look at this explainable AI. We do need to understand why decisions are made, or how decisions are made."

**Barbara** *Do you see any essential measures to ensure ethical AI adoption?*

**Mike** Absolutely. In fact, the IEEE is working on an AI ethics standard. It's crucial that we behave ethically in everything we do in computing. We're affecting humans in various ways, and as we're giving decision-making power to machines, we must ensure that it's done ethically. This means that the people who are involved in building the original systems must behave ethically, and the system itself must behave ethically in the sense that it makes ethical decisions.

**Barbara** *As for the future technical capabilities of AI, on a scale of 1 to 10, where 1 refers to the artificial intelligence systems known today, such as ChatGPT, and 10 refers to general artificial intelligence that surpasses human capabilities. What level will we reach in the future?*

**Mike** In the short term, perhaps the next five years or so, I think we could reach two to three. Maybe in the next 10 years, we might get to five. I don't ever see us reaching 10. I think that's a pipe dream, but it will be marketed as if we've gotten there, because we're already marketing that we're close to it and we're nowhere near that.

**Barbara** *Do we still lack definitions that would allow us to distinguish between what makes us human and what is necessary to be considered AI or AGI? Or should we rather observe and compare the performance of humans and (AI) systems in specific tasks? This would allow us to monitor the progress of AI and separate hype and fiction from reality.*

**Mike** It would be helpful if we had some sort of common understanding of what various terms are, because at the moment, a lot of things that are just automation are being promoted as AI. While they do great things and much better than a human can, and certainly robots can do wonderful things, they don't understand what those things are. So there's no intelligence, but they are using various techniques that are very, very useful. And they do things better than humans in many, many contexts. But again, sometimes that's overhyped. For instance, a recent survey showed that when doctors looked at x-rays of lung cancer, and an AI system did something similar, the AI system did pick up the positive cases, but it had about four times as many false positives. So it's a very good tool to assist the human, but it's not replacing the human and we're putting too much emphasis on the fact that it can do that. So anything that's labeled as an AI is suddenly trusted. And that's the big issue. I think that's the ethical issue because we shouldn't be blindly trusting anything where we don't have detail. We shouldn't be trusting things where we don't understand how they do it. Not saying that the person in the street has to understand, but there have to be people who do understand. And if we agreed to use the proper terminology that if there is something where a system has the ability to make some dramatic decision that could affect people, that would be important. But I think we're overselling something as AI because it's a buzzword when it's just automating things. And yes, we're doing it better now than we did. We have much better techniques. A lot of research has gone into AI techniques over 50, 60, 70 years, and all of these have produced great things. And they're all useful and when we put them together, they're obviously more useful, but they don't necessarily supplant everything and they don't necessarily involve anything like AI when you put them all together. They're techniques that are very, very useful. And if we could come to a common terminology, it would make an awful difference.

"[It is important] to realize that AI is a tool, it's not taking over from us. It's not going to be our boss. It's going to be our servant."

**Barbara** *Do you think there is an interest in driving this AI hype, or is it more of a misunderstanding and confusion that leads to this narrative?*

**Mike** There are some people in the AI community who overhype things. Unfortunately, the AI community has done this since the Dartmouth Workshop back in 1956 [3, 4]. Every 10 years or so, they come up with a new claim that "we're taking over the world, we've built these wonderful systems." And they've done great things, but they haven't built a machine that can think. And I don't think that will happen for anything like the foreseeable future. But a lot of the problem is the media reporting things that they don't understand, claiming that things are intelligent when they're not, claiming that things are going to be revolutionary when they're basically only a step up from the previous version because of better processing power and some new techniques. But some of the people in the community are definitely to blame for allowing this hype to go ahead and say these things. And when we have calls like, "Let's stop doing research for six months because these machines are going to overtake us," I mean,

that's just madness, absolute madness. And of course, the media love that. That sells newspapers, it sells magazines, it gets people reading the websites because they're in a panic. I've had lots of people ask me, people who are not involved in computing, "Should I really be worried?" And I'm like, "Why? Why?" Yes, you should be worried that people are using algorithms to make decisions about you that they shouldn't, whether it's your credit rating or your job or whether you're going to get into college. You should be worried about that because you don't know why and how they're making those decisions. But we're not going to have robots stand up and take over the world, which is what some of the media are claiming. It's madness, absolute madness.

**Barbara** *Did ChatGPT take you by surprise?*

**Mike** I was very impressed with it. I watched a BBC chat show on BBC News (which I only ever watch in a hotel room). Wendy Hall was on it. And they asked her, "Isn't ChatGPT going to take over?", etc. She said, "All it is is predictive text and great processing power." (or words to that effect). And she's absolutely right. That's exactly what it is. But it's very, very impressive. It does some great things. The first time I played with it, we were in a group and we had a task to advertise how IEEE won an award for a conference. And we were in Limerick at the time, so we decided to write a Limerick. So we tried ChatGPT and we gave it the parameters to write the Limerick. And it did quite a few good ones, but it never got the meter right. So obviously, somewhere it didn't pick up that the meter was needed. So the rhyme was good. The text made a lot of sense. A few of them were a bit silly, but a few of them actually made an awful lot of sense. But the meter was never right. So it's got a long way to go. I'm sure it'll improve dramatically. But yes, I'm very impressed with what it can do. It's a great tool. It's a great aid. I have a friend whose first language isn't English and she writes a lot. And so she's using it to improve her English and it does a beautiful job. It really does. But it's not intelligent. It can't write it for her; it doesn't understand what she's writing. It can only improve what she's doing. But yes it's really impressive what it can do.

"My main concern is that the media often labels everything as artificial intelligence, when a lot of it is simply automation."

**Barbara** *Do you think it's important that we educate users in what it actually does and what it can't do to ensure that they have a better understanding of how to interpret what they get out of it? Because currently there might be some confusion.*

**Mike** Absolutely. They need to know that it's a tool that will help them. It's not a tool that's going to take over. It's not intelligent in itself. It's doing smart things. I always prefer "smart" to "intelligent". It's doing smart things that are very, very useful. It's producing very nice results. It's like I say, it's helping people with their writing and there's worry that it's going to help people with their cheating. That's a different story. But I think people need to know that it is

not intelligent. It's going out, it's doing huge searches, it's essentially performing predictive text writing based on material that it's collecting. And so it's not generating something new. It's bringing together something in an essential and useful way. And people need to know that it has its limitations and that it's not thinking about us and it's not going to take over our world and it's not going to turn around and turn on us and try and remove us. It may get rid of some jobs, absolutely, because it's going to be doing things that people will no longer be needed to do and it will do them faster and better. So yes, there will be jobs lost, but there'll be lots of other jobs created in other areas.

**Barbara** *Looking into the future, there are many different scenarios discussed, basically ranging from dystopia to utopia. Where would you position yourself personally?*

**Mike** I guess fairly towards the middle because I don't believe in dystopia. I wouldn't say it's utopia in the sense that it's not all perfect and it's not wonderful, but it's very useful. So I would say somewhere nearer to the utopia end, but near to the middle.

**Barbara** *AISoLA is an interdisciplinary conference where different mindsets and perspectives around AI come together. Was there a specific insight, for example from a different discipline, which was particularly interesting to you?*

**Mike** I don't know if there was a specific insight, but I do think that being able to bring together techncial people, computer scientists – some of whom are AI experts, some interested in using AI in the future – together with philosophers and the like is a great opportunity. We don't normally meet such people at our events. It is great to engage with them, along with sociologists interested in the societal implications of AI, and lawyers interested in the legal aspects and legal implications of using AI in various domains. Bringing together the technical, non-technical, the socio-technical, is a really, really great opportunity.

**Barbara** *In your opinion, is there a specific question or topic we should address from this interdisciplinary perspective?*

**Mike** From the interdisciplinary perspective, we need to look at the ethics. Ethics are absolutely essential. I'm not saying that we have any simple solution for it, and that's why it's so important, but we do need to look at the ethical aspects, and we do need to look at this explainable AI. We do need to understand why decisions are made, or how decisions are made. Again, not the person in the street, but experts need to understand why specific systems are biased in particular ways, and we need to try and either counteract that bias, or at least know that that bias is there.

**Barbara** *Basically, addressing the question, 'should we trust this AI application'?*

**Mike** No, we shouldn't. But we should address the question, yes, you're exactly right. We should be addressing is this trustworthy? Is this saying, is this doing

something that we can allow it to do? And do we have confidence in that it's acceptable to allow it to do that? So yes, trust is the big issue.

**Barbara** *From your personal perspective, what should be the AI vision?*

**Mike** I think the AI vision should be to have machines, whether they're software, hardware, combinations, that will improve society for us. I mean, this is the big opportunity to make life better for all of us. I think we have a duty not to scaremonger and have people panic that these things are going to take over our lives. As I said, it was the same thing in the Industrial Revolution. It was the same thing in the '40s and '50s when computers first came out. People were so naive back then about what computers could and couldn't do and what they would do. And I think we have that same naivety right now about what we can achieve with AI. People believing hype, people panicking because they believe the hype, or they believe that it's part of the hype. And we need to get to a situation where we have AI-based systems that are improving our lives, making life better for all of us, doing the tasks that are boring and mundane and dangerous, that we don't want humans to do anymore, making life better for humans, making healthcare better for humans, basically improving humanity in various ways. But realizing that AI is a tool, it's not taking over from us. It's not going to be our boss. It's going to be our servant.

**Barbara** *Building on that, how do you see the likelihood of using these technologies for bad purposes?*

**Mike** We want to use these techniques, so it's important obviously because of the amount of processing that's needed in things like large language models and in any real deep learning. It's only big players who have the ability to do that, have the processing power, have the financial resources to do it. Even when it comes to smaller applications, it's university level. It's not something somebody can do it in their own home. So we have to consider that there is a divide between those who have access to all of this material and who can control it and who can make money from it. We need to make sure that we benefit all of society, that it's not just a few large companies make billions and everybody else suffers. We need to make sure that people have access to this and can improve their lives.

**Barbara** *Is there anything else you would like to add?*

**Mike** I would ask people to be very skeptical about what they read and what they're being told and what they believe. I think AI has produced a lot of great techniques over the years. It will continue to produce a lot of great techniques. We've seen a step change in the last two, three years of what can be achieved and we'll do more definitely, but we don't need to worry that these systems are going to take over the world. What we do want to do is make sure that they're used ethically, safely, reliably, and that they improve life for everyone.

**Barbara** *Is there a specific place, people or outlets where I as a user could inform myself sufficiently?*

**Mike** I think if you want to get information online, make sure you're taking trusted sources, that you're not reading blogs and postings from people who are just trying to cause hysteria and get attention for themselves. Go to the sites of recognized engineering or software organizations, whether it's the IEEE or ACM or IFIP or the British Computer Society or Australian Computer Society or whatever, but trusted sites that will give you proper information and don't believe things randomly on the internet. But of course, that's true for absolutely everything, not just AI.

**Barbara** *Perfect. Thank you very much for your time and insights, Mike. Enjoy the rest of AISoLA!*

## References

1. Isaac Asimov, "Runaround", Astounding Science Fiction, March 1942.
2. Isaac Asimov, I, Robot, Gnome Press, 1950.
3. Ronald R. Kline, "Cybernetics, Automata Studies and the Dartmouth Conference on Artificial Intelligence", IEEE Annals of the History of Computing, October–December, 2011, IEEE Computer Society Press.
4. John McCarthy, Marvin Minsky, Nat Rochester and Claude E. Shannon, A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. `http://raysolomonoff.com/dartmouth/boxa/dart564props.pdfÂăAugust,1955`
5. Roger Penrose, The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics, Oxford University Press, 1989.
6. Roger Penrose, Shadows of the Mind: A Search for the Missing Science of Consciousness, Oxford University Press, 1994.