



Vol.7 No.1 (2024)

Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin

Rose Luckin ^A	A	<i>Professor of Learner Centred Design, UCL Knowledge Lab, University College London, UK</i>
Jürgen Rudolph ^B	B	<i>Director of Research & Learning Innovation, Kaplan Singapore</i>
Martin Grünert ^C	C	<i>Dean, Academic Services, Kaplan Singapore</i>
Shannon Tan ^D	D	<i>Research Executive, Kaplan Singapore</i>

Keywords

Artificial intelligence (AI);
AIEd;
education;
ethical AI;
generative AI (GenAI);
higher education;
(human) intelligence;
machine learning.

Abstract

Professor Rose Luckin, a pioneer in the integration of artificial intelligence with education, holds the position of Professor of Learner Centred Design at the UCL Knowledge Lab, University College London. Her trailblazing research has profoundly deepened our understanding of AI in education (AIEd). Rose Luckin has authored over 50 peer-reviewed articles and key works, including "Machine learning and human intelligence: The future of education for the 21st century." As the Director of EDUCATE, she merges academic insights with ed-tech industry innovation. She is the co-founder of the Institute for Ethical AI in Education.

In our interview, Rose Luckin shares her educational awakening and her personal journey into AIEd, addressing gender bias and the unique challenges faced by women in the AI field. She delves into the ethical dimensions of AI deployment in educational settings, underscoring the Institute for Ethical AI in Education's pivotal role in fostering ethical standards. Professor Luckin advocates for AI's potential to bolster learner-centred methodologies and stresses the critical importance of forging robust partnerships between educators and technology developers. She evaluates the impact of generative AI on assessment, learning and teaching within K-12 and higher education. She provides insights into AI's evolving role in education and the imperative of lifelong learning. Emphasising a collaborative ethos among educators, researchers, and developers, Professor Luckin argues for AI's integration into education within strategically crafted ethics and governance frameworks. Our interview sheds light on AIEd's current landscape, highlighting the critical need for ongoing research and collaborative efforts in navigating its considerable dangers while seizing opportunities.

Correspondence

Jurgen.Rudolph@kaplan.com ^B

Article Info

Received 6 March 2023
Received in revised form 18 March 2024
Accepted 18 March 2024
Available online 26 March 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.27>

Personal experiences of school, further and higher education

Jürgen Rudolph (JR): You are such a distinguished figure in the burgeoning field of artificial intelligence and education. You are the Professor of Learner Centred Design at the UCL Knowledge Lab, part of University College London, where you have been an influential force for over two decades. Your pioneering research significantly advances our understanding of AIED.

You are not only a prolific academic, authoring over 50 peer-reviewed articles and several notable books, including “Machine learning and human intelligence: The future of education in the 21st century” (Luckin, 2018), but you are also a leader in applying research practically. As the Director of EDUCATE, a hub for educational technology start-ups in London, you have fostered innovation and collaboration between academia and the ed-tech industry. Your expertise has been recognised widely, including your appointment as a Specialist Adviser to the UK House of Commons Education Select Committee and your role as a Co-founder of the Institute for Ethical AI in Education. Before joining UCL, you served at the University of Sussex in various capacities, including as Pro-Vice-Chancellor (Teaching and Learning) and Director of the Human Centred Technology Research Group. Your administrative and research roles have been instrumental in shaping policies and strategies in both teaching and technological contexts.

To the best of our knowledge, academia was your second career. Prior to the birth of your first child in 1983, you had a career in Banking and Finance. You had been a Top 20 candidate in the UK for the A.C.I.B. Chartered Institute of Bankers Associateship Examinations. In 1990, you started to pursue a BA (Hons.) in Computing and Artificial Intelligence at the School of Cognitive and Computing Sciences at the University of Sussex (you were awarded a First Class Honours), and between 1994-1997, you pursued a PhD with ECOLAB on Explorations in the Zone of Proximal Development at the School of Cognitive and Computing Sciences, University of Sussex from 1994 - 1997. Could you please tell us more about your own schooling and university education?



Figure 1: Rose Luckin at a conference in China in 2019.

Rose Luckin (RL): It's a very interesting question on many levels because my own schooling was quite disrupted. I didn't have a smooth journey through the education system. I had a lot of problems when I was in my early teens. I was a school refuser, though I was very good at sports. It was a very insightful sports teacher who sat me down one day and said, 'You know, you could be good at something else as well as running'. It's funny how one person can really have quite an impact because, in the UK, we have a system that involves selection at 11, called the eleven-plus [a competitive examination given between primary and secondary school at about age 11 as a means of determining in which of the three types of secondary school – grammar, technical, or modern – a child should continue their education]. I had failed the eleven-plus when I was expected to pass it.

My parents were very dismayed. I've let them down, and part of the reason that I was struggling at school was because I felt like a real failure. Obviously, my parents had confidence in me, but I felt too guilty for letting them down. So, it was very important when the sports teacher showed confidence in me. I've been very frank here, but I think it's quite interesting for your readers to realise that you can become a professor even when you've had quite a disrupted early education. Also, this disruption has impacted a lot on my research interests.

You can become a professor even when you've had quite a disrupted early education.

So, when I was encouraged to take learning seriously, I was very fortunate that my parents moved me to a different school to give me a fresh start. It was a shame because it meant I left the PE [physical education] teacher, but it did give me a fresh start. I then, of course, worked super hard. It was like going from one extreme to the other to try and catch up. In the end, I did get my formal qualifications, my O-Levels – now GCSE [General Certificate of Secondary Education] – and the right kind of grades. I did get my A-Levels, and interestingly, I ended up in the same school I would have gone to had I passed the eleven-plus to do my A-Levels. So, it came full circle in the end.

But even then, I didn't go to university straight away because I came from a family that had never had anybody go to university. Despite that fact, my father was a lovely man, and I loved him dearly; he's sadly dead now. He basically said to me that when I said 'I'd like to go to university': 'No, you need to get a job; that's what we do. You've done well, really well. But let's get a job now, my love.' This is the same man who, despite having no expertise in the area of my PhD, read every word of my PhD to check the punctuation [all laugh], so he is a lovely man.

But that was the way it was. So, I didn't go to university when I left school, despite having the A levels that would have enabled me to do that. So, I went to the bank because I needed to get a job. I was the only woman taking the exam in the regional branch where I was working. They were a bit surprised when I said I wanted to do the A.C.I.B. Chartered Institute of Bankers Associateship Examinations qualification. I didn't just want a job; I wanted a career. But they agreed.

Bless them! They supported me through college because I had to do it as a day-release student in a further education college. You would work, and then, every other week, you would have a day where you would go to college. But it was hard because you were doing a lot of studying at the same time as holding down a full-time job. But that was good for me.

I quite thrived on that concentrated effort. I surprised them when I did really well in my exams, finished in the top 20, and got a little medal. The branch was excited. They've never had anybody do this, so they were very pleased. It was a very interesting educational experience for me in the further education sector and definitely formative to me as an educator. The secondary school education I had was initially in a very poor-quality secondary school. So, when I had completed the banking exams, I stopped full-time work and had my children. At that time, my husband was working as a schoolteacher, so we didn't have a huge amount of money, but it was fine. I took full maternity leave, and I didn't go back to work other than part-time until both my children were at school.



Figure 2: Rose Luckin and daughter.

I became a tutor for something initially called the Rapid Results College, which was a correspondence college. Again, this was very influential on the research I later conducted. Before the Internet, I would receive students' work in the

post, mark it, and send it back to them with feedback. Then, they promoted me so that I was writing the course textbooks. The students who enrolled in Rapids Results College got a set of texts to help them through the banking exam content. I was writing the little course textbooks, pointing them to other resources and structuring the instruction in particular subject areas. That, again, very much informed who I then became.

I then – in order to bump up the family budget, to be honest, but also to keep my brain cells working – became a part-time tutor at the local further education college. So, after having benefitted from the further education system in the UK, I then became a teacher in the further education system. That was really interesting because I was engaged in teaching students, not just the banking subjects, which is obviously what I had been tutoring for Rapid Results College.

But I was also asked to teach adults with special educational needs. I have no training in doing this at all. This was incredibly informative to who I then became as an educator: to be set the challenge of 12 very significantly disadvantaged young people, trying to help them use a computer. In one instance, the student didn't even recognise that when he pressed the letter on the keyboard, he got the same letter on the screen. We're talking about serious disabilities in learning terms, so it was very challenging. At that time, I realised that because the legislation around banking changed in the UK, I wasn't going to go back to a career in banking because it wasn't the same job that I had left. A lot of the more interesting activities in local banking had been moved to central offices and taken out of the regional branches, and I didn't want to end up in some kind of service centre. So I thought, 'Well, ok, I don't have a degree, let's now go to university, let's have a look'.

Because I had small children, geographically, there was only one university I could get to. It was quite easy to fill out the form, and I had the A levels and the banking qualification. I thought I'd probably better do something financial or economic. I applied to do economics. Then, I looked through the student brochure and saw this subject called artificial intelligence and computer science. I had never heard about it, read through it and thought, 'Oh, this looks interesting'. They recommended some reading, so I went and bought the books. One of which was "Gödel, Escher, Bach" [Hofstadter, 1999; first published in 1979], which is an incredibly difficult read. It was a very intense introduction to AI philosophy, things I hadn't engaged with previously but fascinating, and I loved it. I thought: 'Nothing to lose. I'll just put it down on the application form and see what happens.'

I got rejected from economics and accepted for artificial intelligence and computer science. When I became Pro Vice Chancellor, the person who was the vice chancellor was the man who rejected me from economics, and I never let him forget it. [All laugh.] He rejected me because he thought I had already done it. He thought my banking qualification would have influenced my thinking. So he said, therefore, I wouldn't approach the subject fresh. But he did me a favour because I loved computer science and AI.



Figure 3: Professor Rose Luckin.

I was one of two female students on my course. Everybody else was male. I went to Sussex University, which encouraged mature students. There were other people who were more of my age because I was in my 30s at this point. But they were male, and at one point, the other female dropped out of the course. So, I was the only female in the cohort. I had a very different experience of the university and degree study because I couldn't join the social things, as I had two small children.

I had to get home, and I had a very intense work phase. I would work in the middle of the night because it was the only time I could get quiet when the children were in bed. We lived in a very rural village with no street lighting. So, I would often sit in the bedroom working away and look out. It'd be completely black outside. It felt quite alone trying to work out this computer programme that wasn't working. A lot of the time was spent debugging.

Anyway, that's how I got into computer science and AI, and it was interesting being female. But, in the end, I had a wonderful relationship with the other students, even though most of them were younger than me. Once they realised that I could actually do the work and that I was a serious learner, we got on great. I'd help them. They'd help me, too. It was a really lovely learning experience, and it was a very interdisciplinary course that has very much influenced the way I think, too, because it was quite the early days of AI. This was 1990, so it was the very early days of AI, and

it was very much about what intelligence is. So, we were looking at psychology and philosophy, and of course, we did theoretical computer science and programming and all of those things. But we also did linguistics and lots of contextual courses, and I really enjoyed it. I loved studying, and I loved the course. That was a very long answer, but I hope it gives you some useful context that your readers might find interesting when thinking about how you get into a subject like AI.

Gender bias in academia

Shannon Tan (ST): You have researched AI since the 1990s. As one of the early female pioneers in the field of AI and education, could you share your experiences regarding the challenges you faced and overcame? We suspect you encountered glass ceilings that you have successfully shattered during your stellar academic career. How have you seen the gender dynamics in this field evolve over the years?

RL: It's very complex. I definitely experienced challenges both as a female and as an older female. Because of my previous experience, I initially regarded some of those challenges as being my own fault. So, for example, I would never use my children as an excuse for not being able to do something because I didn't want to draw attention to the fact that I was female and had children. Of course, everybody knew I was female, but I didn't want to use that as an excuse for not being able to meet a deadline or not being able to do something. I would make a huge effort not to have to do that. I don't know whether some of the challenges I faced, particularly in the early part of my career, were to do with me being female or to do with me being quite successful quite quickly. It's hard to know why people behave badly towards you because there's a lot of competition in academia. There are always people who want to get one over on you. I wouldn't want to blame gender for things that might not be gender-related.

But what I would say is that things have improved in terms of gender politics and the glass ceiling. It's still an issue and a challenge. I don't think we've solved it. In fact, in later years, I would say that the challenges that I experienced came more from other females than from males. Unfortunately, perhaps because it is difficult to be a successful female academic, women don't always support each other as much as they could. It's a very complex issue, and there is no simple answer. But I certainly have experienced challenges. The way I tended to try and deal with them was to deny that they had anything to do with the fact that I was female. Later, my approach has been to try to be as collaborative as possible, to try and accept that you can't agree with everybody and that there will be people who will do you down. It's a fact, but you try not to take it too personally, see it as the pressure they're under, and emphasise collaboration, which is why so many of my publications have many names on them. This is not always well respected, but it very much speaks to that collaborative approach that I've found personally the best way of navigating what is a tricky situation.

JR: I agree that collaboration is brilliant. Unfortunately, when you look at the metrification at universities, they oftentimes very much favour single-authored publications.

Ethics in AIEd

Martin Grünert (MG): As a co-founder of the Institute for Ethical AI in Education, you have been at the forefront of discussing ethical considerations in AIEd. What do you consider the most pressing ethical challenges in implementing AI in education, and how can educators and technologists work together to navigate these challenges?

RL: We have to work together because it's so complicated. When I formed the Institute with [Sir] Anthony [Seldon] and Priya [Lakhani, OBE], the bottom line of why we formed it was because we were really worried that nobody was paying attention to education when it came to AI. There was lots of data and AI, lots of ethics and AI, and ethics and data. But nobody was looking at education. I believe to this day that education is a special case because we want people to engage with education throughout as much of their lives as possible – unlike the medical system, where we want them to engage with it as little as possible because they're healthy. It also involves vulnerable people. The reason we founded the Institute for Ethical AI in Education was that we were concerned that nobody was really paying attention and that if we didn't do something, all of the great benefits that AI could bring to learners, teachers and parents would be lost because there'd be some ethical dilemma or horrible occurrence and everybody would, of course, have a knee jerk reaction against the technology.

A large part of the motivation for founding the Institute was to draw attention to education as a special case and to try and alert people to the fact that we needed to think very carefully about ethics. Of course, the situation's got even more complicated with the launch of generative AI applications that have made open-source, scaled, and sophisticated AI available to anybody. That means that people with bad – or good – motives can develop sophisticated AI systems, and that's dangerous. So, we have to work collaboratively and engage all stakeholders around ethics. It's not just about the technology itself; it's also about how the people developing the technology can engage in ethical codes of practice. It's not just about the regulation; it's also about the education of the general public. From my perspective, a key section of that is that educators need to understand more about AI, what it can do, and the basics of how it works. They don't need to know the deep tech complex science because not everybody's into that, but they do need to know the basics.

We have a problem because (a) the regulation and the code of practice will never keep up with what the technology is able to do, and (b) huge assumptions will be made about, for example, what it means to be transparent. Because transparency is not a one-way process. You might think that as a developer, you have explained what happens with your data very carefully. For example, if you're using machine learning, what data do you collect, how do you process them, and how is your system trained all of that time? Unless the person reading it has a certain level of understanding,

it's not transparent to them. So, there's a huge educational imperative around that ethical space. It absolutely has to be a collaboration between regulators, developers, educators, and community groups. It's really important that we see the immense impact that these technologies have on society and, in particular, on education. We're still waiting to hear when GPT-5 will be released, and we're being told lots of things about what it's going to be able to do, and quite frankly, I find it frightening what I'm hearing (see, e.g. Shah, 2024). Is the world ready for this? It's quite a worrying space.

Advocating educator-developer partnerships

MG: Indeed, there's always a risk if the educators are not involved in the process. As we are discussing education, it's concerning to me, as an educator, that solutions are often designed based on perceived needs rather than what will truly be effective and positively impactful. The next question is about the symbiosis between educators and developers and how they can work together. How do you see that partnership being able to be firmed up or strengthened in order to be able to generate the most beneficial outcomes?

RL: The bottom line is that it takes time, and it takes patience and a willingness for all parties to perhaps put aside some of the things that they might normally prioritise. Let me be a bit more specific. I've written in the past about the importance of AI developers working with educators (e.g. Luckin et al., 2016; Cukurova et al., 2019; Weatherby et al., 2022) – and I do believe it absolutely is the way forward. But it's a very difficult road. When I was talking about my history when I was teaching at the further education college, I didn't mention that I also taught in school for a short period of time, again in a very challenging environment. I don't think I was very good, but it has given me an understanding of what it's like to be a secondary school teacher with a very challenging group of young people in their teens.

When you've got multiple things going on, and you've got to keep control of and support that group of people, the last thing you want is a not-quite-finished bit of software to work with. You need something you can depend on. You have to do it in stages. There is a lovely book chapter called Lesson 21 by Kate Erricker, with whom I worked on a project for Nord Anglia Education, that really highlights this so nicely and tells the story of trying to co-design a piece of software and how incredibly difficult it was for the teachers, even with the best will in the world, to do this in a high-stake learning environment (see Erricker, 2023). In those learning environments, it's high stakes in a different way compared with my situation with very challenging learners.

You can't work with something that's still being built, and yet the only way to really engage educators in that process is for them to be part of that design. You have to have stages. Engaging educators in being able to be part of pilots, where they're not necessarily operating in that high stakes environment whilst the technology is not in a place that would support them in that environment but helping them to engage in the design process to understand a bit more about what it is, and how AI works.

If you can help people to understand a bit more about what it is that AI is doing and how AI uses data and what that means, they then are much more able to become proficient users and also understand what potential ethical challenges and safeguarding challenges might be. So, it needs to be long-term relationships. They need to be patient. People need to be absolutely upfront about the challenges of the activity right at the outset. Initially, it needs to be done in a way where teachers aren't put in high-stakes environments with semi-functional technology because it's just too stressful. But they can be part of projects, and then once you get to a certain stage where they feel comfortable and the technology is in a state, then you can start rolling it out and testing it in those more real-world, authentic, high-stakes environments.

AI's potential in enhancing learner-centred approaches

JR: Could you please walk us through some of your key works? In your book "Machine learning and human intelligence: The future of education in the 21st century" (Luckin, 2018), you discuss the potential of AI to transform education. Could you elaborate on how AIEd frameworks can be optimised to enhance both teaching and learning experiences, especially in the context of fostering critical thinking and creativity in students? This is related to your 2010 book "Re-designing learning contexts", where you advocated for learner-centred design in education. How do you see AI further enhancing learner-centred approaches, and what gaps in research and practice still need to be addressed to achieve this?

RL: The whole learner-centred design piece has been a core part of my thinking throughout all of that work. When I was writing the book on "Re-designing learning contexts" (Luckin, 2010), I used writing as a way to help myself understand more and help myself learn, which is why I find the suggestion that an AI would ever take over my writing process completely meaningless. Because I actually want to learn from trying to express myself. I'm sure AI can help. Don't get me wrong. But we learn from exposing our understanding, examining it, reframing it, learning from it, and challenging it. I think the framework that I put forward in the "Re-designing learning context" book is something that I'm now revisiting because AI itself is now far more able to tap into those contextual features that I was trying to identify in that book. So, the primary reason I wanted to write that book was to try and understand more about it - How could we talk about learning contexts?

There's a huge literature on context, but it's quite complex. There are lots of different disciplines that talk about context, and they talk about it in quite different ways. When I was writing that book, I concluded that, from my perspective, being pragmatic and trying to come up with a framework could be useful when thinking about designing the way that learners interact with technology. It was more useful than conceptualising context as something that people step into or out of: 'I'm in this context now' or 'I'm in a learning management system' or a 'virtual learning environment context for a moment'. 'I'm in my office context'. Another moment, 'I'm in the car context'.

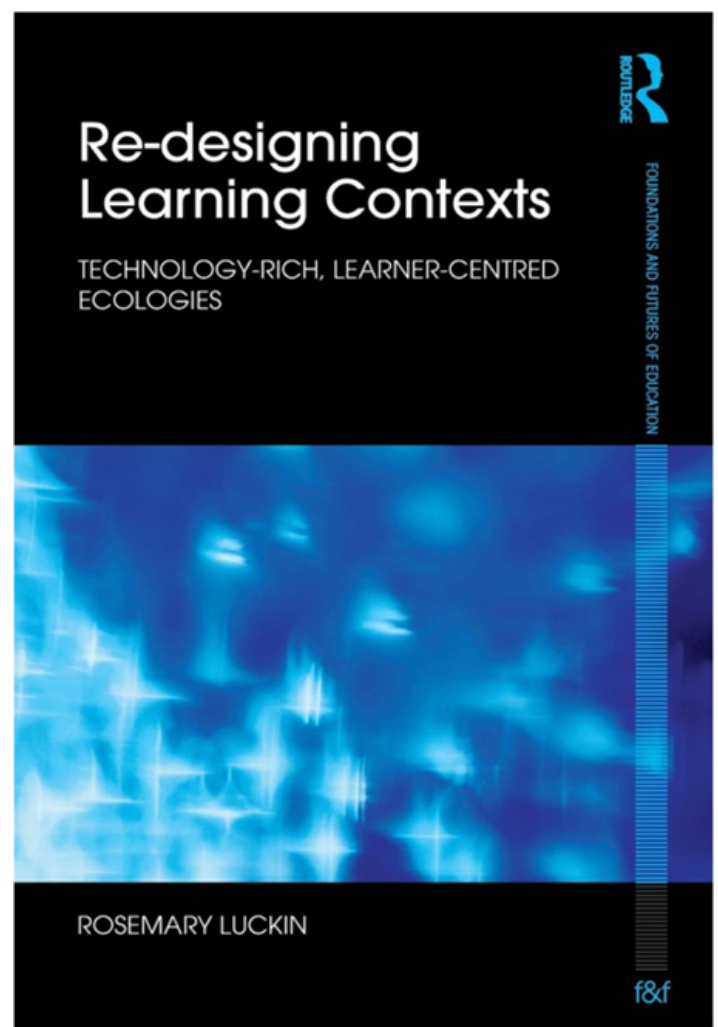


Figure 4: Re-designing learning contexts (Luckin, 2010).

I saw those as environments, and I saw context itself as a whole set of interactions that I might have with the world as I go through my life, but with a particular slice of learning. I believe that people create their own context from all the interactions they have in the world. But of course, that's huge. How do you start thinking about that from a design perspective? Then I tried to think about how you might think about ways in which you could identify different elements of that context. This is an oversimplification. So, thinking about digital and physical environments – what are they like? I thought about the kinds of tools, both technical and non-technical, that you use within their environment and how much there is a relationship between the environment and the tools that are used.

For example, in a science lab, you use quite different tools from the tools that you use in an English language classroom. This is quite an obvious connection, but there's a connection between the nature of the tools that are used and the environment, and then to think about what it is that you're trying to learn. So, you have your learners in the centre, and they interact in different environments. They use different tools, and they access knowledge, information, understanding, skills, and capabilities about a particular subject area or skill through their interactions in environments. But none of those things are necessarily a direct relationship.

So, a child in a classroom doesn't necessarily have the ability to interact with the teacher whenever they want. There are written and unwritten rules about how that works. So that's why, in the framework of that book, there are these items called *filters*. They can be positive or negative in terms of the way that they constrain the access that a learner has to a particular resource. That resource might be part of the environment, a tool, or a part of what they're trying to learn. So that's where that book came from. It was a way of trying to take a very complex subject like context and turn it into some kind of framework that could be useful when we're thinking about designing learning experiences.

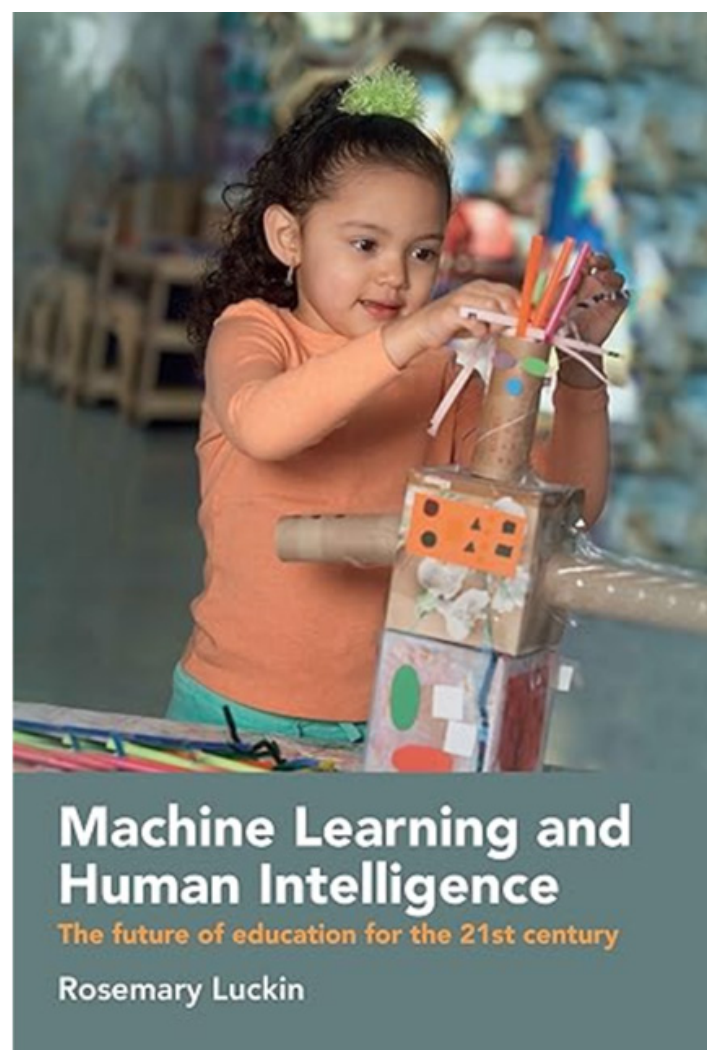


Figure 5: Machine learning and human intelligence (Luckin, 2018).

Now, with the ability that AI has to capture data about those interactions that a learner is having as they progress through their journey, I'm relooking at that work: it's interesting to think about how we can consider design frameworks for AI that are informed by that conceptualisation of context. Then, the other book, "Machine learning and human intelligence" (Luckin, 2018), was written explicitly because I wanted to understand how better I could talk about human intelligence in a way that would distinguish it from artificial intelligence. Because even back in 2017/18, when I was writing that book, I was really bothered about the way we were talking about AI as if it was human-intelligent when it wasn't, and it still isn't human-intelligent. Not the same way that we are.

I was really bothered about the way we were talking about AI as if it was human-intelligent when it wasn't, and it still isn't human-intelligent.

So, that was the absolute motivation for writing that book. That's why the framework in that book is very much focused on how we can differentiate AI from human intelligence and looking at what's happening with the large language models, for example. Particularly now, they're not just language; they're multimodal. I'm having to revisit that and challenge myself about how those differences are perhaps being blurred. Whilst I don't believe that the sorts of generative AI models we're seeing understand the world in the way that we do (because they don't directly experience the world), I do see that there could be an argument that says: 'I spend six months interacting with a multimodal generative AI model. In those conversations, I can see that there's a reasonable premise that you could consider that that model is building an understanding, a representation of the world that we are interacting about'. It doesn't have direct experience of the things that we might discuss but I think there is some notion. There are still clear differences between human and artificial intelligence, but the boundaries are more blurred.

There are still clear differences between human and artificial intelligence, but the boundaries are more blurred.

It shows that whatever you write, you have to revisit it all the time because our understanding of human intelligence, our understanding of AI, is changing so fast. So, if you write something one day, you have to be willing to revisit things. So, I still believe in the fundamental premises of both books, but I think the precise detail is shifting, and that's really interesting.

JR: Everything is developing at breakneck speed. This can be exemplified by Sora, OpenAI's text-to-video AI (see Brooks et al., 2024) that created global headlines and will soon be launched. It's staggering, and it raises questions about what truth is and what reality is. As you were saying earlier, it's quite scary. By the way, we are currently working on an AI framework for our institution's teaching and learning, and we are quite inspired by your ethical AI framework (Luckin et al., 2022).

RL: It's really useful to know that because we went to a lot of trouble to try and make the output from that work really practical.

The Golden Triangle Framework

JR: In this context, your Golden Triangle Framework (e.g. Luckin & Cukurova, 2023) provides a model for AI and education by connecting three key elements – educators, researchers and EdTech developers. The Golden Triangle Framework emphasises the role of the teacher in orchestrating the educational experience, selecting appropriate AI tools, and ensuring they are used effectively. AI can provide adaptive learning experiences personalised to each student's needs,

The 7 steps to AI Readiness: EThICAL

There are seven key steps to getting your organization ready to leverage the transformational power of AI. These can be found in the 'EThICAL AI Readiness' framework:

- 1) **Educate, enthuse, excite** – about building within your community an AI mindset
- 2) **Tailor and hone** – the particular challenges you want to focus on
- 3) **Identify** – identify (wisely), collate and
- 4) **Collect** – new data relevant to your focus
- 5) **Apply** – AI techniques to the relevant data you have brought together
- 6) **Learn** – understand what the data is telling you about your focus and return to step 1 until you are AI ready
- 7) **iterate**

And all these steps should be done ethically



Figure 6: The seven steps to AI Readiness (Luckin, 2020).

giving real-time feedback and enabling practice through repetition. However, good AI education solutions require human-centred design that is focused on supplementing/enhancing the teacher's role rather than replacing them. Effective AI requires an iterative, evidence-based approach drawing on multidisciplinary expertise in learning science, technology, and practice. AI should be designed ethically with transparency, privacy and agency in mind to foster trust and metacognition. Your framework highlights the need for research on how best to integrate AI in education to improve learning outcomes. Would it be fair to say that AI should not be on autopilot but, at best, a co-pilot? Could you also please elaborate further on your Golden Triangle Framework, perhaps sharing aspects or nuances that are sometimes overlooked?

The Golden Triangle

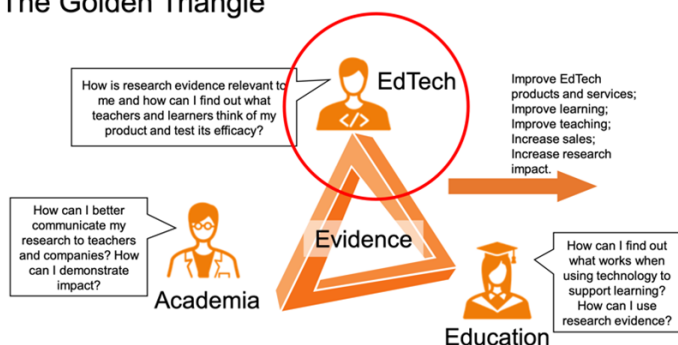


Figure 7: The Golden Triangle Framework (Luckin, 2020).

RL: That's a very big question. I will start with the Golden Triangle, but I really want to come back to that question about co-pilot versus autopilot because I think it's the crux of where we are with AI. Let me give you a sense of the context of the Golden Triangle. For many years, when I've been working in what is essentially AI and education, the word 'AI' has not been popular. 15 years ago, you didn't necessarily want to say you were working in AI because people didn't necessarily want to hear it. So, we would talk about educational technology, and of course, some educational technology doesn't use any AI anyway.

For many years, when I've been working in what is essentially AI and education, the word 'AI' has not been popular.

Although I've always been doing research on AI in education, I did branch out to look at educational technology more generally because of my own personal learning experiences and my focus on applied research. I was always looking to do things that would actually have a useful, practical application. At that time, most of the things that were being used in education were educational technology, not AI. I spent quite a bit of time thinking about and looking at different educational technologies and talking to companies who were developing them. I've always believed that having a relationship with those who are developing their technologies commercially is really important. About 12 years ago, on January 6, I pulled together a meeting and had a wonderful facilitator. It was a meeting where I brought together some big tech companies, some small tech companies, and some educators from across the different sectors. Basically, the question I asked them to answer is: 'Why are we not better able to use technology in education impactfully? What is the problem here?' We came to the conclusion that there was just a really big disconnect.

'Why are we not better able to use technology in education impactfully? What is the problem here?' We came to the conclusion that there was just a really big disconnect.

You had a whole body of people who were researching technologies for use in education and sometimes building research tools that never made it to commercial viability. Another bunch of researchers were exploring the use of existing commercial tools in education but very rarely actually speaking to the people who built the technologies. Then you have this whole raft of people who were building technologies, both small-scale and large-scale, who didn't know how to contact the right person in the research world to create the right connection. Then you had the poor old educators who were left at the mercy of both groups, quite frankly.

We concluded that if we could do one useful thing, it would be to better connect all of those stakeholders. My passion for collaboration is not surprising. The group also thought that we needed to think about how we could create better connections between the communities of people who built the technology and use it. This could help us understand how, why, when, and for whom it is usefully impactful. That's where the Golden Triangle came from.

Then, unfortunately, I was diagnosed with breast cancer, and had a rough ride. If I'm honest, because I had to have all the treatments, the chemotherapy, the surgery, the hormone therapy, everything, I was very ill for two years. I completely lost two years.

JR: I'm very sorry to hear that.

RL: You probably won't want to put this in the article, but I don't actually mind if you do if you think it would be of interest. I also had a complete psychological breakdown and was in a secure psychiatric institution for a month. The reason I am willing to share that is because it has made me think

about artificial intelligence differently. Because my mind was completely and utterly broken in a way that is really hard to explain. And yet I recovered. That's amazing. It's an amazing capacity of the human brain and our intelligence to recover from such a fundamental breakdown where you were such a danger to yourself that you have to be locked up, literally.

My mind was completely and utterly broken, and yet I recovered. It's an amazing capacity of the human brain and our intelligence to recover from such a fundamental breakdown.

The reason I'm always willing to share it is because I am a very honest person, but also because it really informed the way I think about artificial intelligence and human intelligence. There's so much we don't understand about our human intelligence and capacities that I hope we will learn more about because we have these AI tools. I hope we'll see it as an opportunity to really explore the depths of our capacities. That was a hard time. I lost two years completely and probably, if I'm honest, three because it took me a long time to get back physically and psychologically to be able to cope with a full-time and quite high-pressure job. But I did get back at full throttle, as you can see. That's why I say it's an amazing capacity of the human mind.

But it meant that that conversation just got left there. Then, three years later, I thought about, 'How do I pick this up?' I started to reengage with the people, and interestingly, they were all happy to reengage in the conversation. Nothing had happened to them in the meantime, really. There had been a small group who tried to take some of the ideas forward but hadn't really kept its momentum. We revisited it, and then, to cut a long story short, at the end of 2015 and the beginning of 2016, I was encouraged to apply for some funding. It wasn't research funding as such. Development funding from the European Union was used to build better relationships between small companies' entrepreneurs and, in my instance, Edtech companies and universities. So, I put in a proposal for this funding to try and explore this Golden Triangle concept that we'd come up with, and we'd come up with a name – EDUCATE.

Fortunately, that was funded. But of course, a few months after I got the information that it had been funded, the UK decided to leave the EU. That was another little barrier in my way [all laugh]. We have to wait for various political ramifications. Still, in the end, the funding came through and in 2017, the EDUCATE programme started. That had the explicit goal of trying to bring together those three communities. It wasn't about AI. It was about educational technology in general. Of course, AI is relevant and fits within that, but it would be part of the educational technology category.

We did have companies like Century Tech, which is quite a big AI company, that came on one of our very early cohorts and what we did was quite simply design a training programme to try and help developers better understand how to evidence what was impactful about their particular product or service, and to start building a logic model.



Figure 8: Rose Luckin at UCL EDUCATE Demo Day 2019. Source: Team EdTechX, 2019.

They were clear about what the educational outcome they wanted to achieve was, what the outputs from their particular product or service were, how they could collect data and evidence and how they could analyse it along the way to try and help them have more informed conversations with their customers, the teachers. We wanted to engage researchers in this, and we did. Therefore, we also trained some researchers to understand better how to engage with these small tech companies because they both work to very different rhythms of life. You actually have to build that relationship as it doesn't happen naturally.

We've never managed to do the same relationship building at scale because we work with over 360 companies, which is a lot of the UK tech sector. Some companies were from outside the UK, but most are UK-based. We never managed to build the same depth of relationship with educators because it was so hard to access them. Because they're incredibly busy, don't have time, and can't use semi-functional products. Then, there's a lot of complexity for educators. We did engage with educators, but it felt that it was more from the perspective of giving them information than actually helping them to be part of a collaboration. It felt much more that we were exposing them to information and opportunities, but I never felt we'd built the kind of in-depth collaborations I would have liked to have seen. I'm still trying, obviously. Now, the Golden Triangle is highly relevant to the AI space. The building of the relationship between the different parties is fundamental to the success of any AI use case that's being rolled out.

In Singapore, there are three use cases being rolled out across schools. That was decided and designed before ChatGPT was launched back in 2022. It's really interesting to see how that unfolds. But we certainly need to build those relationships that are defined by that Golden Triangle and try to learn from some of the approaches to training, for example, those that were developed as part of that programme. That's certainly what we're trying to do at the moment. The company Educate Ventures' researchers are absolutely helping educators think about the purpose for which they want AI to be used. What do they want to

achieve with it? Rather than starting with the technology, asking 'What's an educational challenge? What is the task that could be helped by AI?' To start from that point and then develop small tests, even before pre-pilot, even at a very early investigation of what it means for that AI to be used: What kind of infrastructure do you need to have? What kind of staff capability do you need to have? All of those are very practical things that you need to understand before you can get into really interesting questions about how that aspect of AI is supporting learning or how it is helping teacher development.

You need to get those pragmatics right. We have a little framework for doing that and a framework for thinking about how you evaluate each of the ways in which you are looking at AI and how you can take a very early intervention through to a pilot and think about how that can inform a bigger strategy. What I'm seeing, certainly in the UK but not exclusively in the UK across the entire education sector, is that people, including educators and students, are using AI. But I've not yet seen anyone who's developing it in an organisational, strategic way. What I've learned from the Golden Triangle is that we can build those relationships within an underpinning governance and ethics framework, which is fundamental in running a series of use cases of AI.

We look at the technical and data infrastructure required, how they build staff capability, and what staff capability is needed. We have a framework for learning from every iteration that feeds back into the strategic way in which AI is being used in an institution. That's the sort of framework we work with, and it's all really been inspired by that Golden Triangle. It's interesting how it's evolved, right back from 12 years ago.

JR: Thank you so much for sharing, and we're so sorry to hear you went through such a torrid time.

RL: I'm fine now. I'm always willing to be honest about it because (a) I think it helps other people who might have had similar experiences, and (b) it really taught me something about intelligence and the human mind; it really fundamentally shifted my own perceptions.

JR: As you know, Shannon and I co-authored a book with Stephen Brookfield (Brookfield et al., 2023). Stephen went through some ten years of severe clinical depression. For the same reason that you mentioned just now, he also talks about it very openly because he believes it's a taboo topic that men especially like to suppress because men always see the need to perceive themselves as strong and so on (Brookfield et al., 2023). I think it is excellent that educational thought leaders like you and Stephen Brookfield talk about their personal crises and how they eventually overcame them.

RL: I think it's really important to be open. I understand the male perspective on that, and I don't deny anything that he's saying, obviously. But that was part of the challenge that I faced as a female: not wanting to show any weakness. I tried to work throughout all my treatment, which was a complete mistake. The day before I collapsed and was admitted to the psychiatric hospital, I was trying to read a student's thesis.

I couldn't process the words, but I just didn't feel I could give in. I thought I had to keep the flag flying. We females might get breast cancer, but we can hack it! And actually, I couldn't. That's the honest truth.

That was part of the challenge that I faced as a female: not wanting to show any weakness. I tried to work throughout all my treatment, which was a complete mistake. The day before I collapsed and was admitted to the psychiatric hospital, I was trying to read a student's thesis. I couldn't process the words, but I just didn't feel I could give in.

Generative AI and its impact on education

ST: Since its inception more than a year ago, ChatGPT has fired the public imagination with a vengeance. Within a couple of months, the AI chatbot has hit more than 100 million users. There is an increasing number of rival chatbots. Having been so long in AIED, you may be amused by the sudden hype that surrounds large language models (LLMs) and generative AI. What are your personal experiences and impressions of these developments in the past year? How do you see these advancements influencing future research and education in AI?

RL: I love the way you phrased that: "fired the public imagination with a vengeance"! On the one hand, I find myself completely delighted because people actually want to talk about AI. I've been trying to get people to want to have conversations. I've been trying, particularly, to talk to teachers, people who are actually practitioners. I'd been wanting to have much more conversation with them and school leaders about AI, and it was really hard to get anybody interested. Now, it's not hard at all, so it's good from that perspective.

It's a very powerful technology. I find it fascinating, and if I can manage to be detached, I can look at it in a more scientific way and think about: 'Okay, so what's the sort of sociotechnical story here?' We've got this technology. People weren't ready for it. It was a quite deliberate launch of something that society wasn't ready for. As a mass experiment on millions of people, that is interesting, but it's also quite frightening. I find myself on the boundary, often between excitement and fear, because I think it's great that people now want to engage with AI. Something like this was always going to happen. There was going to be a moment when the power of AI became more apparent.

It was a quite deliberate launch of something that society wasn't ready for. As a mass experiment on millions of people, that is interesting, but it's also quite frightening.

In a sense, the power of AI had been apparent but hidden before. Social media is driven by AI algorithms, and we know how much of an impact that has had on society, education,

and people. But it was not explicitly, 'I am a user using AI', 'I'm using social media, and I may or may not have any idea that there's AI actually running this'. So, this was an availability at scale of a very powerful AI that was easy to use, and you knew you were using AI. It had it labelled on the tin – 'OpenAI'. This is an AI. It wasn't hidden at all, which is all good. But of course, the world is not prepared for it, and that's my biggest concern. My experience to date tells me that I'm right to be worried, from politicians to schoolteachers to all sorts of different ages. The vast majority of them have no idea what they're doing with AI.

In a sense, the power of AI had been apparent but hidden before. Social media is driven by AI algorithms, and we know how much of an impact that has had on society, education, and people. But it was not explicitly, 'I am a user using AI'.

Understandably, why would they? It's not a criticism in the least. But I often get asked by policymakers, and they're doing their best. Some of the questions they ask are incredibly naive and really worrying because that's the kind of thing that is driving the way education systems work. Singapore is a different case because they've been engaging in thinking about AI much longer at a system level. I think they are definitely ahead of the game, but nevertheless, it's still a challenge to think, 'How do I help so many people grasp enough about AI to understand how and what they should use it for?'

This comes back to the question about co-pilot and autopilot. There are many reasons why I think that everybody needs to understand enough about AI to use it safely and to their benefit. They need to understand enough to build AI. Few people do, of course, but we don't need everybody to do that. It's because we don't have a good conceptualisation of what the relationship between an AI and human intelligence should be. There is no clearly defined framework for thinking about that relationship. There are no recommendations. We certainly need to start conceptualising the different sorts of relationships that can exist between artificial and human intelligence quite quickly and thinking about the ramifications of those. Because we already know that technology is changing the way our brains work at the neural level. We are changing because of the way we use GPS, because of the way we use Google rather than remember things.

Technology is changing the way our brains work at the neural level.

So, however we use this AI, it will change us cognitively. We better make sure those changes are the ones we want and that we don't offload the wrong things to the AI. That autopilot/co-pilot thing is fundamental. Unless we have an educated enough population, how do they navigate that very challenging space? Even, say, if it was possible – I'm not sure it is – to come up with a framework, we have different conceptualisations of the relationship that could exist between human and artificial intelligence, a bit like the EU AI act that has different ranges of risk (European Parliament,

2023). You could have different types of relationships, and then you could look at each of them and think, 'Well, what are the implications of this for the developer of the AI, and what are the implications of this for the humans and for educators and stakeholders in general?' Then, you could explore each of those. But of course, even if you had that, unless you've educated the population sufficiently, how do they access what that means? Do you see what I mean?

We could do that work, and we could at least come up with something that might be a bit useful. It wouldn't be perfect. That's for sure, but it could be something. But unless people understand enough about AI, they may not be able to make the best use of that. So, it's a real educational challenge. But we have to not go for AI as the autopilot; we have to keep human decision-making in the loop at the appropriate level.

We have to not go for AI as the autopilot; we have to keep human decision-making in the loop at the appropriate level.

That means people need to understand what that means, 'Ok, I'm going to let my AI do X, Y and Z'. But what does that really mean? I'm letting it do something, and what am I still making decisions about? It's really challenging. But at least this imagination with a vengeance is opening up the possibility of those conversations in a way that we couldn't have had over a year ago or so.

ST: ChatGPT has made particular headlines when it comes to assessment in higher education. It has been said that the essay is dead, and the integrity of open-book online exams has become questionable. We may arrive at a situation where the lecturer creates the assessment using ChatGPT, then the students write the assessment using ChatGPT, and then the lecturer marks the assessment using ChatGPT, and nobody learns anything anymore (Popenici et al., 2023). To what extent does generative AI impact assessment? Is there a legitimate use of generative AI by students and teachers? Does AI call for a reassessment of how we should set assessments? Should we try to make assessments more authentic? Should we try to have more assessments that attempt to be 'AI proof'? Could you please also discuss the potential of AI in revolutionising traditional assessment methods, particularly in providing real-time, personalised feedback to learners?

RL: Assessment is a key area of impact. But I think the key question for people when we think about assessment is: How are we going to change our assessment methods so that they actually assess what we need to assess? In the book "Machine learning and human intelligence" (Luckin, 2018), I say that we treasure what we measure, and assessment is our measurement tool. At the moment, in most cases, I don't think we measure the right things, and therefore, we don't treasure the right things.

The key question is: How are we going to change our assessment methods so that they actually assess what we need to assess? We treasure what we measure, and assessment is our measurement tool.

At the moment, in most cases, I don't think we measure the right things, and therefore, we don't treasure the right things.

If we were measuring more sophisticated thinking skills, then the assessments probably wouldn't be as accessible to something like ChatGPT. But whatever! We don't know how these models are going to develop. So, the key imperative is to think about redesigning assessment and revisit the basics of why we are assessing and what it is we want to assess. We have to move on from assessing memorisation, which is a lot of what we do in the UK. It's perfectly obvious that any quite basic large language model can do pretty well at exams that test memorisation. If they've been trained on the right data, they'll remember it. They might get some things wrong, but they'll still do well.

There's an imperative for us to think carefully about what it is we want to assess and then how we're going to assess it. On some occasions, that assessment may involve the use of AI. Why not? I believe it should be seen like this: If this AI is here to make us more intelligent, then stopping people from using it is a very backward step. So, how do we make the assessments that even when you're using AI, you're still assessing the student's capability, knowledge, understanding, whatever it is you're trying to tap into, to measure accurately?

It's the biggest job. I was recently involved in a roundtable organised by an examination body, trying to explore this very question. What is it we should be assessing? How should we be assessing? How could we use AI in assessment? Of course, that's another element of AI, whether you're using it for automated grading or to provide automatic feedback, perhaps in a formative assessment context, or whether you are using it to analyse and measure student progress as they learn. AI is definitely a powerful tool that can be used as part of the assessment process itself. It's also a tool that students can transparently be encouraged to use as part of the assessment. But before you can get to that stage, you've still got to do the nitty gritty on what it is we're trying to assess and why we want to measure this. Why does this matter?

At the moment, rigour has got in the way of the value of the things that we measure. We think, 'As long as the measurement is rigorous, it's ok'. But actually, now it's not. So, we need to do huge amounts of thinking. That also relates to the question about the relationship we want to have between humans and artificial intelligence. Because if we want to look at the human side of the relationship in a particular way, how do we assess the extent to which we're achieving that successfully?

AI is not going to go away, and it's, without question, radically transforming the workplace. Just look at the legal profession. It's turning cartwheels or somersaults at the moment. Perhaps more accurately, these AI tools can do much of what the legal profession was doing. Of course, not always flawlessly, but I think we owe it to our students to focus on being much clearer about what we want to measure through our assessment systems and why. Then,

we can start thinking about how and what the role of AI might be in that as part of the assessment or as part of the measurement.

Those conversations need to be informed by the kind of relationship we believe society would benefit from between human and artificial intelligence. But the big elephant in the room is who is going to be responsible for that? I remember attending an event in the UK many years ago, which was organised by a body called Becta, which doesn't exist anymore. It was a non-departmental government body that oversaw the way that educational technology was used in the UK and developed framework support, etc. They were organising a whole set of policy workshops. I went to one of these, and there are various policymakers, big tech representatives, educational body representatives and quite a lot of civil servants. At the end of one of these sessions, we were asked the question that was about changing the education system: Who is responsible? Is it the educators and the educational institutions, or is it the policymakers and government, or is it society as a whole? The majority of the room, not myself included, voted for society as a whole because nobody wanted to own it. [All laugh.]

Who is responsible? Is it the educators and the educational institutions, or is it the policymakers and government, or is it society as a whole? The majority of the room, not myself included, voted for society as a whole because nobody wanted to own it.

And therein lies a huge challenge. Is it the examination bodies? Is it the government? Is it the schools? It's got to be a cross-sector, multi-stakeholder engagement. But who's going to bring it together? And how do we get it right? A really big challenge!

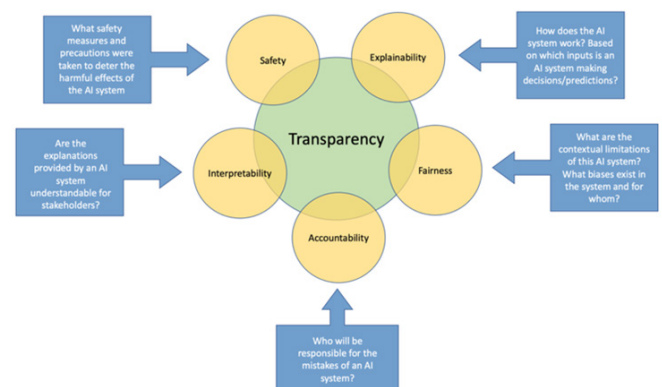


Figure 9: Transparency in relation to other dimensions of ethical AI (Chaudhry et al., 2022).

JR: We are now moving on from assessment to teaching and learning, which are, of course, intrinsically intertwined. How can we change our approaches to teaching and learning in light of generative AI? ChatGPT and other chatbots (like Bing Chat, Bard – now called Gemini (and hence no longer insulting to Shakespeare) – and Ernie) are the latest shiny things in the long history of AI in education. It may be too

early to determine ChatGPT's (and GPT-4's) place in that history, but what are your preliminary thoughts? In your view, how should higher education institutions deal with generative AI?

RL: The assessment piece is obviously significant, but we've talked about that. 'Strategically' is my answer. I don't think you should separate generative AI from AI in general. Obviously, it has different capabilities, and that needs to be recognised. However, every higher education institution needs a good AI strategy. That should be built on their existing vision, and that institutional vision may need to change. It depends on what it is. You might want to revisit it. But if you have a vision that's about educating a wide demographic of students in a fair and equitable way, there's no reason that would change just because we have AI. The way you do it might change, but the existence of that vision doesn't necessarily change. Then, strategically, how are you going to leverage AI to help you achieve that vision more impactfully, way faster for less money, whatever particular metric you want to use in that context? How's it going to help you achieve that vision? That's what it does come down to.

It's complex, obviously, and I'm going to oversimplify it, but one key element is governance and ethics. You need to put in place at the highest level some governance structure that's specific to AI. Maybe it's a high-level committee. Maybe you even invite external people to sit on it to help you make sure you keep abreast. Maybe some of the tech companies get involved. Maybe there's a sector-wide one. There needs to be some high-level governance and ethics committee. Then, obviously, each institution needs to have different policies that would help the people within that organisation understand how to navigate the boundaries that the governance and ethics committee is setting for that institution.

What's allowable? What are we encouraging? What's not? Why? But then, what are the policies that help people on the ground – students, teachers, department heads, university leaders, whatever – to understand how to put that governance and ethics framework into practice? I would encourage people to think of this as an iterative learning cycle. You need to look across the institution and think about areas where perhaps you're facing particular challenges, and they may not all be in the teaching and learning space. Some of them might be in the back office. You need to look to see whether a particular challenge or activity you need to get done – better, faster or cheaper (whatever it is) – could be enhanced by AI.

This is this idea of purpose-driven AI, and once you've identified some use cases that are driven by purposes, what are the kinds of technologies that you might use for that? We've got a template that we use. What's the purpose of this? What's the AI going to help with? Why might it bring benefit? What are the ethical implications? What are the staff capability implications? What are the technology and data implications? What are the practical implications of actually implementing this? How are you going to collect data and evidence about whether it's actually doing what you want? How are you going to learn as an organisation for that? So, you test a few things out, and if they look like

they've got legs, pilot them, but always look to see how that helps you or is constrained by your existing technology, data, and existing staff capability.

So, obviously, those initial use cases are going to be constrained by your existing technology and human capability infrastructure. But each time it is rolled out, there's learning to be had. So, how are you going to make sure that that learning is passed on to everybody who can learn from that particular use case? How are you going to learn about ways in which your technology and data infrastructure might need to change because of this technology being rolled out, and then gradually, as you iteratively test out these different AI, some of them you will chuck out very quickly? Some of them you'll take to pilot and then chuck out, and some you'll take to pilot and then roll out. But by the time you get to rolling out, you'll have a really good sense of what that AI is meant to achieve, what the implications of trying to implement it are, and what you need to have in place in order for it to work. How, fundamentally, are you going to know whether it's achieving what you wanted to achieve?

So maybe, for example, you might have a use case around automated feedback. Let's be honest. In universities, students often don't get feedback in a timely manner. It's improved, obviously, but often they don't, and AI can actually be quite effective at providing feedback. You may not want it to do all of that. You may want it to be overseen by a human. That's something obviously we need to consider. However, various commercially available applications can help provide feedback for students. Feedback is also a really expensive thing to provide, so there's a real cost implication. So, you might decide that's a use case. Look for your tool. The benefits you're hoping to achieve are faster feedback to students, thus more effective for their learning, reduced workload for lecturers, more information for lecturers (because you're able to analyse all of that automated feedback), and saving costs. There might be five elements of what you're looking for, so then you need to have a framework within which you're going to say, 'Ok, yes or no' if it didn't pass our threshold of satisfaction for that particular area.

So, you have these four main elements: (1) governance and ethics, (2) iteratively rolled-out use cases, (3) the technology and data infrastructure piece and (4) the staff capability piece. They are all highly interconnected, but I think it's always helpful to break things down when you're facing a very complex situation. One thing I learned very early in my AI career studies is that a lot of AI at that time was about understanding a problem. So, this is trying to simplify what's a very complex situation.

So, it's ok to start with these four boxes - the governance and ethics, the use case that is being iteratively rolled out across different areas of the institution, technology and data infrastructure, and staff capability. But that iterative process brings other people in, as per the Golden Triangle. So as I have said, with that governance and ethics piece looking for externals, you can help the university understand more, perhaps, about where the technology is going or perhaps about the law. The Golden Triangle pieces are at the heart of helping educators to be part of that iterative testing and

learning.

Integrating AI into inclusive education

MG: In your view, how can AI be effectively integrated into inclusive education to support learners with diverse needs and abilities? I ask that out of interest in what you mentioned earlier about transparency not necessarily being transparent. Introducing new tools that may not be understood by the end user, or where there is a disparity in the existing skills of class members to adapt to a novel technology, could unintentionally create learning environments that disadvantage certain segments of our student population. What are the key considerations for ensuring AI technologies are accessible and beneficial to all students, regardless of their learning differences?

RL: It's a hard question, and there's also an added extra dilemma. There's a minority of students who ethically object to using AI because of its potential damage to the climate. So, there's a real equality issue there. Do you respect their quite heartfelt, passionate views? If you do, then they're not going to have the same access. So, it's definitely a dilemma.

Two of the key features the AI brings to the party are the ability to be adaptive and the ability to behave autonomously, though, as we've discussed, the extent to which you want your AI to be autonomous is something that needs to be thought about carefully. Because an AI works by analysing its environment and then, based on that analysis, the way it's been built and trained, what it's meant to achieve, and what its goals are, it behaves in a particular way. So, if it's a grading algorithm, the environment is the assessment materials it is processing. It's been trained on thousands or millions of previous examples and various other artefacts like rubrics, etc. The goal is to produce some feedback. So, we know that's how it works. Of course, it's adaptive, so the feedback it gives to one student won't be the same as the feedback it potentially gives to another, even if their assignments are identical, although that's an interesting question.

So, in the one sense, it brings that ability to be adaptive and to behave with a certain amount of autonomy that should, in principle, mean that no matter what a student's ability, there should be an AI that can help them and meet them where they are. That's part of the dream. It's this notion of really quite hyper-personalised learning where you have an AI that really gets to know the student. Hopefully, the student gets to know themselves as well, but that's the sort of dream of the AI that enables you to meet any user's needs – because of this adaptivity and autonomy. Of course, the other side of that promise is this dilemma that you really are right to identify: there are issues about accessibility.

We're already seeing it, and it's so true. If you pay for one of these models, they're so much better than the ones that you get free. So, it is an immediate equality issue. The issue that you already highlighted about whether people understand nothing anyway in the first place to be able to access the technologies is a real core accessibility issue. Then you've got this ethical dilemma about whether people passionately feel these technologies shouldn't be used and whether we

need to respect that. On one hand, it brings the potential to really help address challenges of access and equality, and on the other, it brings us a whole new raft of ways in which we can get it wrong.

AI, teacher education and professional development

ST: With the rapid advancement of AI in education, what changes do you foresee in teacher education and professional development programmes? How should teacher training evolve to equip educators with the necessary skills to integrate AI tools effectively in their teaching practices?

RL: Again, this is such a core part of the puzzle. They have to change to take it into account. I don't foresee a rapid change, unfortunately. At least it'll be patchy. Some countries will be much further ahead than others, so it's not universally the case. But history teaches us that changes to teacher education and professional development can be very slow. And here we're faced with a technology that's moving very quickly. That is a really big challenge for the education system. So, potentially, the best way of dealing with it is actually through continuous professional development. Of course, you need to look into initial teacher training. But it's not just initial teacher training. It's recognising that this is going to be something that people are going to have to learn about, not just now, but next year, the year after, and the year after that. There is a need to design for continuous professional development to support teachers as the technology changes, but they need to get that base understanding first, and we're not even there yet in most cases.

History teaches us that changes to teacher education and professional development can be very slow.

So, the pressing need is to help educators get a basic understanding of AI so that they're what I would call 'AI-ready'. Then, you can start looking at how you develop their AI readiness into different skills, abilities, and capabilities. Maybe some people in an organisation specialise in one type of AI or one where AI is used in a particular channel, and that will be up to organisations to decide. But it also needs to be part of that overarching strategy that an institution needs to have. There are these constant cycles of learning.

Two further very practical things. It's essential that senior management are very vocal about their support for training because they're always making a difference in how well it works. The teachers are given time and recognised for doing the training so that it's something that they feel good about, that they feel they want to do. We really need them to do it because they're being faced with students who are using these technologies. Anecdotally, I was talking to an organisation recently about some research they've been doing on young people where some school-age learners are. In fact, the majority see AI as the solution to all their problems. That's a real worry.

The majority see AI as the solution to all their problems. That's a real worry.

If we don't have educators who have a basic understanding to help them navigate that path better than we've done with social media, then we're in a difficult position. After a basic understanding, continuous professional development needs to be built into each institutional strategy. As you iteratively roll out the AI, everybody's learning, and people are open and honest and feel ok about being honest. It's a bit like dealing with phishing emails. You need people to tell you when they've inadvertently opened that attachment. You don't want them to feel scared because you need to know they've opened that attachment. You need people who feel that if they get something a bit wrong with AI, they're actually going to tell you about it because you need to know! [All laugh.]

JR: I've clicked on those phishing emails before.

RL: Yeah, everybody has.

ST: Considering the importance of lifelong learning in today's dynamic world, how do you see AI contributing to adult education and continuous professional development? What are the opportunities and challenges in this area?

RL: This goes beyond the education profession. World Economic Forum (2023) data show the expectation that the majority of the workforce isn't very well prepared for the future. There's an imperative there for companies to prioritise ongoing training, which they often don't do. But they're going to be impacted by AI whether they like it or not because their customers are going to use it even if their staff don't. So, it's a global challenge. Few businesses are completely unlikely to be impacted at all by AI. Even if you're running a corner shop, for example, and you feel that it's all about your personal relationship with your customers who come in face to face, and you meet them, and you sell them, and you build relationships, some of them will be using AI. In general, there is an ongoing need for lifelong learning and professional development and for that to be prioritised.

Evaluating the effectiveness of AIEd

MG: You mentioned earlier that, under the Golden Triangle, you should be sure that you are able to evaluate the effectiveness of the things that you implement. So the question is, what methodologies or frameworks do you recommend for evaluating the effectiveness of AI-based educational tools and interventions? How can educators and institutions assess the impact of these technologies on learning outcomes to ensure that they are pushing student learning in a positive direction?

RL: We've got to be open and innovative in the way that we approach the evaluation. This isn't a space for randomised control trials, for example, not for a long time anyway, and for many reasons. We need to think about methodologies that enable us to collect quantitative and qualitative data to understand the experience because that is supremely important. In this instance, we can also use AI to help us with that data analysis. So, we can do intelligent analytics on the data that we collect – not just the quantitative data, but also the qualitative data. So, we have the potential to be

quite revolutionary in the way that we evaluate the extent to which a particular AI is or is not supporting the student, the teacher, the parent, or whoever is focusing on it in the way that we want.

But again, that's a question of access because most people don't understand enough about the potential of intelligent analytics and wouldn't know how to necessarily do it. So, we have to start simple but look at ways in which you can mix the quantitative and qualitative data that you collect in a very formative way. When you're doing that initial test, which you could think about as an MVP [minimally viable process], is the technology even worth considering in any detail? What's the MVP that we want to test here? 'Ok, it's passed that threshold'. So, there are different ways of evaluating for different purposes.

'Ok, we're going to pilot it.' What do we want the pilot to achieve? Then, I think using the logic model approach is a great framework for doing this. But to be thinking all the time, how much more sophisticated are the data that's collected through that logic model framework? It would be the subject of some intelligent analytics in the future so that when you get to the stage of actually rolling out an AI across a part of the institution, or even the whole institution, you've got the right data being collected. You've got the analytics in place to make sure that you can learn the things you really want to learn about. So, it's in stages, but the logic model is a great framework to use for that.

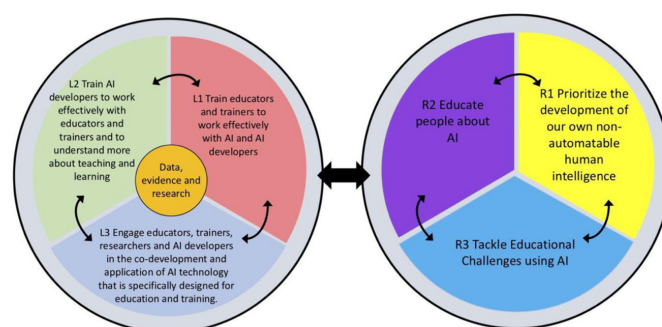


Figure 10: A co-design framework for AI to be used in education and training (Luckin & Cukurova, 2019).

AI and the role of the teacher

MG: Oftentimes, for us working in the private sector in particular, there is a need to justify financially not just the implementation of the technology but also the degree to which these initiatives impact student outcomes, progression, and the student experience. If we look at it from the perspective of the technology, we often see it changing in advance of important considerations regarding its implementation. The question is, how do we keep an eye on balancing the use of these AI and digital tools while still retaining the human-centricity of the learning experience? How do we support students with an AI-curated learning environment without robbing them of the auxiliary benefits of having human educators?

RL: Several years ago, I was asked what my dystopian and utopian views of the future of AI in education were, and at that time, I thought what I was saying about the dystopian view was just a nightmare that wouldn't happen. The dystopian view would be that you would have students, generally from poorer backgrounds, being educated mainly by AI systems that were hyper-personalised and adapted to their needs. So, lots of technology interacting with the students, probably some bouncers to ensure that students behaved, but very little human interaction because, after all, that's really expensive, isn't it? Once you've invested in the AI systems, they can work 24/7. They don't need holidays. They don't go on strike. They rarely break, and they get better.

The dystopian view would be that you would have students, generally from poorer backgrounds, being educated mainly by AI systems that were hyper-personalised and adapted to their needs. So, lots of technology interacting with the students, probably some bouncers to ensure that students behaved, but very little human interaction because, after all, that's really expensive.

The utopian view, which, obviously, would have been for the more privileged students, was one where the human was much more in evidence. The learning is very human-driven, with the technology very much in the background, orchestrated by the student and the teacher working together. Yes, it is hyper-personalised but very much driven by the humans in the room, with the students gaining a really deep understanding of themselves as learners, of the data that could help them understand themselves, of how to control them, of the ethical implications, the way to benefit from it – but that being in the background.

The utopian view, which would have been for the more privileged students, is hyper-personalised but very much driven by the humans in the room, with the students gaining a really deep understanding of themselves as learners.

To be honest with you, I am actually really worried that a dystopian perspective is more likely now than it was a few years ago, and I'm worried about it because I see the dilemma. If I am in a country where there are millions of children who get no education and a technology company comes along to me and says, 'Don't worry, we can give you a hyper-personalised adaptive assistant for each child'. You're not going to say 'no', are you? Because at the moment, I've got nothing, and this would give them something, but to what extent does that then become the thin end of the wedge?

You would be comparing it to nothing that the students would otherwise be learning. If you're able to demonstrate, as I'm sure you would, to what extent do other people think, 'Oh, we could save some money here!?' 'We're financially

challenged.' This comes back to that core question that we've talked about several times, which is, 'What do we want the relationship between the human and the artificial intelligence to be?' Where is the sweet spot where both benefit? We want the AI to get better at giving us what we need. Maybe it's not what we want, and we need humans to get better at making sure the AI gives them what they need. Not necessarily what they want, so that relationship is still core.

The future of work

JR: What will generative AI do to graduate and academic employment? Daniel Susskind (2020) recently wrote a book titled "A world without work". You appear to adopt a more optimistic view (for instance, in Luckin et al., 2016). I believe you are saying that AIED will help people continue to be employable. What are important skills and competencies for graduates to become and remain employable? How do you see the future of academic work in light of generative AI?

RL: Generative AI has the potential to automate some routine cognitive tasks currently performed by humans, but I believe there will still be a strong need for our unique human skills and competencies among graduates and academics. Key skills for ongoing employability include:

- Creativity and innovation: Coming up with novel ideas, making connections between disciplines, and designing innovative solutions. AI can enhance but not fully replace human creativity.
- Social and emotional intelligence: Understanding emotional cues, building relationships, empathy, communication, and collaboration. These distinctly human abilities will become even more valuable.
- Complex problem solving and critical thinking: Formulating, analysing, synthesising, and solving new complex problems that require reasoning, judgment and strategy development. AI can augment but not replicate advanced human cognition.
- Adaptability and self-management: The ability to continuously upskill, learn new things quickly, and manage one's own learning and career. This lifelong learning mindset enables graduates to evolve alongside technological disruptions.

Within academia, generative AI will likely enhance academic productivity through automated basic research and writing assistance. But uniquely human skills like conceptualisation, creativity, complex critical analysis, judgment, social perceptiveness, and wisdom will become even more valuable among academics. AI will augment technical aspects, allowing more focus on the interpretative and social dimensions of scholarship. Lifelong learning for faculty to continually advance their own expertise and teaching ability will also grow in importance.

Artificial General Intelligence and ethics

JR: We are now looking a bit more into the future. Nick Bostrom, a philosopher at Oxford University, has written a book about superintelligence (2014). He is cautioning that after computers have achieved Artificial General Intelligence (AGI – which essentially means that they can think and act like humans), superintelligence may be close. This would mean that machines would be potentially and exponentially more intelligent than us humans. One possible outcome would be an extinction event for humanity. A more benevolent one, perhaps, would be that computers keep us as pets. Movies like *The Terminator* and *The Matrix* immediately come to mind. What are your thoughts on this? How do you view the ethical responsibilities of AI researchers?

RL: The prospect of super-intelligent AI that surpasses human capabilities gives me great pause for thought and some significant worries. I don't believe there will be an AI that can do everything a human can do, and all of it will be better than a human. But I do believe that with the right emphasis on ethical foresight and safeguards, we can develop AI that enhances rather than endangers humanity.

As an AI researcher focused on education, I feel a profound responsibility to ensure the work I'm involved with broadly benefits society and empowers humans rather than replaces them. Key principles we must embed into AI systems are transparency, accountability, respect for human dignity and agency, and optimising for the social good rather than solely efficiency or profit motives. Robust policy frameworks on areas like data rights, preventing bias, and adherence to human rights also need to be implemented alongside the technology as it advances. Multi-stakeholder consultation is key.

As an AI researcher focused on education, I feel a profound responsibility to ensure the work I'm involved with broadly benefits society and empowers humans rather than replaces them.

Rather than AI systems that control critical decisions autonomously, I believe human-AI collaboration is crucial – designing intelligent assistants that enhance human judgment, creativity and well-being. Systems focused on personalised learning over standardised outcomes respect learner agency, too. The goal must be expanding human potential – economically, creatively, socially and culturally – not limiting it. With ethical AI guardrails and human-centric design principles in place from the outset, I am hopeful we can positively shape this technological frontier for the betterment of humanity as a whole.

The goal must be expanding human potential – economically, creatively, socially and culturally – not limiting it.

MG: In "Intelligence Unleashed: An argument for AI in education" (Luckin et al., 2016), you present a vision for AI in education. Looking forward, what emerging trends or technologies in AI do you believe will be most transformative

for education in the next decade? Finally, could you share with us your future plans and any additional topics or advice you'd like to offer to upcoming researchers in the field of AI and education? And is there anything else we did not cover that you would still like to discuss?

The future of AIEd

RL: The continuing rapid pace of progress in AI currently is exciting. Areas like natural language processing (NLP), personalisation via machine learning (ML), multi-modal interaction, and human-AI collaboration hold particular promise over the next decade. For example, NLP could enable richer dialogue and feedback between the learner and AI tutor. Personalised machine learning could allow more targeted content, guidance and support customised to each student's strengths and weaknesses. Multi-modal AI that incorporates speech, vision, and haptics, along with language understanding, could make interactions more intuitive and assistive for a wider range of learning needs. And improved techniques for human-AI complementarity will be able to amplify learner potential. As long as we get the ethics, safeguarding, and guidance correct, these are just some of the significant benefits we can expect.

I envision a future with AI-empowered personalised lifelong learning companions that can adapt to each individual, motivate and guide self-directed growth – companions that are firmly in the control of the user and there to support them to become ever more intelligent and ever more knowledgeable about themselves as a learner. Learners should be creators and partners alongside their AI tutors, definitely not just consumers of content. Educational equity can expand tremendously if we get the guidance and regulation of AI progress right.

As someone pursuing research in this exciting field, I will focus first on the key human challenges to address rather than leading with technological capabilities. I will always explore the ethics, the potential for inclusion and the human development needs from the outset. As always, I will continue to collaborate actively with educators, learners, and policymakers to ensure real-world relevance and responsible progress as we try to ensure that we humans become ever more intelligent – especially in uniquely human ways. There are still so many open and exciting questions that merit creative exploration around peer learning, meta-cognition skill building, transferable competencies, and the interplay between AI and quality human teaching, for example. I think I am likely to continue to be very busy!

JR: We are very grateful to you for this outstanding interview.

References

- Bostrom, N. (2014). *Superintelligence. Paths, dangers, strategies*. Oxford University Press.
- Brookfield, S. D., Rudolph, J., & Tan, S. (2023). *Teaching well. Understanding key dynamics of learning-centered classrooms*.

Routledge. <https://doi.org/10.4324/9781003447467>

Brooks, T., Peebles, B., Holmes, C., DePue, W., Guo, Y., Jing, L., Schnurr, D., Taylor, J., Luhman, T., Luhman, E., Ng, C., Wang, R., & Ramesh, A. (2024, February 15). *Video generation models as world simulators*. OpenAI. <https://openai.com/research/video-generation-models-as-world-simulators>

Chaudhry, M. A., Cukurova, M., & Luckin, R. (2022, July). A transparency index framework for AI in education. In *International conference on artificial intelligence in education* (pp. 195-198). Springer International Publishing. https://doi.org/10.1007/978-3-031-11647-6_33

Cukurova, M., Luckin, R., & Clark-Wilson, A. (2019). Creating the golden triangle of evidence-informed education technology with EDUCATE. *British Journal of Educational Technology*, 50(2), 490-504. <https://doi.org/10.1111/bjet.12727>

Ericker, D. K. (2023). *Lesson 21: Technology, innovation and system change* [Presentation]. <https://www.cite.hku.hk/wp-content/uploads/2023/10/20231030keynote2.pdf>

European Parliament. (2023, December 19). *EU AI Act: First regulation on artificial intelligence*. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Hofstadter, D. R. (1999). *Gödel, Escher, Bach: An eternal golden braid* (20th anniversary ed.). Basic Books.

Luckin, R. (2010). *Re-designing learning contexts: Technology-rich, learner-centred ecologies*. Routledge. <http://dx.doi.org/10.4324/9780203854754>

Luckin, R. (2018). *Machine learning and human intelligence: The future of education for the 21st century*. UCL Institute of Education Press.

Luckin, R. (2020, August 4). *Is education AI-ready? Turing lecture* [PowerPoint presentation.] https://www.turing.ac.uk/sites/default/files/2020-08/turing_lecture_-_is_education_ai_ready_pdf.pdf

Luckin, R., Bligh, B., Manches, A., Ainsworth, S., Crook, C., & Noss, R. (2012). *Decoding learning: The proof, promise and potential of digital education*. Nesta. https://www.researchgate.net/publication/269111789_Decoding_Learning_The_Proof_Promise_and_Potential_of_Digital_Education

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, 50(6), 2824-2838. <https://doi.org/10.1111/bjet.12861>

Luckin, R., & Cukurova, M. (2023). Small-scale commercialisation: The golden triangle of AI EdTech. In B. du Boulay, A. Mitrovic, & K. Yacef (Eds.), *Handbook of artificial intelligence in education* (pp.537-552). Edward Elgar Publishing. <http://dx.doi.org/10.4337/9781800375413.00036>

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, 3, 100076. <https://doi.org/10.1016/j.caeai.2022.100076>

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. <https://discovery.ucl.ac.uk/id/eprint/1475756>

Popenici, S., Rudolph, J., Tan, S., & Tan, S. (2023). A critical perspective on generative AI and learning futures. An interview with Stefan Popenici. *Journal of Applied Learning and Teaching*, 6(2), 311-331. <http://dx.doi.org/10.37074/jalt.2023.6.2.5>

Shah, S. (2024, February 14). *ChatGPT 5 release date: What we know about OpenAI's next chatbot*. The Standard. <https://www.standard.co.uk/news/tech/chatgpt-5-release-date-details-openai-chatbot-b1130369.html>

Susskind, D. (2020). *A world without work: Technology, automation and how we should respond*. Penguin.

Team EdTechX. (2019, November 14). *UCL EDUCATE demo day 2019*. Medium. https://medium.com/@team_edtechx/ucl-educate-demo-day-2019-7856ff1fd2ea

Weatherby, K., Clark-Wilson, A., Cukurova, M., & Luckin, R. (2022). The importance of boundary objects in industry-academia collaborations to support evidencing the efficacy of educational technology. *TechTrends*, 66(5), 784-797. <https://doi.org/10.1007/s11528-022-00705-0>

World Economic Forum. (2023). *The future of jobs report*. <https://www.weforum.org/publications/the-future-of-jobs-report-2023/>