

AI and the future of children's social care: why waiting is not an option

By Professor Sir Anthony Finkelstein CBE FREng,
President of City St George's, University of London

A personal perspective

I recognise that you will think at the conclusion of this that I am an overenthusiastic tech evangelist, but I'm not. I'm an engineer with, in general, a rather cautious view of the benefits of the adoption of new technologies. I've been through the cycle enough times to have reached that point.

I believe that enthusiasm for new tools must be tempered with caution about their consequences. I understand that the adoption of technology is socially mediated and that organisational institutional structures are complex and resistant to change. Equally, however, as an engineer, I'm unafraid to acknowledge the possibilities of technical, technologically induced disruption and large shifts in the way things are done and the way society is organised.

The evidence for this is clear. I first encountered artificial intelligence in the late 1970s to early 1980s, at the very end of what was called the AI Winter, signalled in the UK by a report by James Lighthill in 1973. At my first sight, I was interested, but not more than that. I could see the relevance of what were then called knowledge-based approaches and so-called expert systems, and these gave rise, in my case, to a long-standing concern with the use of logic and symbolic reasoning that actually has shaped a good part of my later work.

The history of AI

Machine learning, specifically neural nets, were familiar to me. They reflected the so-called biologically inspired strand of AI. Interesting conceptually and important in applications such as vision and pattern recognition. But otherwise, I thought, a side-line. Besides, they appear to have substantial practical limitations that I assumed would constrain their further development.

Roll forward. Computer speeds increase, growing exponentially, as data volumes and accessibility algorithms are refined leveraging these capabilities. And it seemed reasonable to assume in light of this that machine learning would benefit. I anticipated, correctly, that some of the familiar applications would take advantage of this better image processing or perhaps better speech recognition. I was aware that people were building ever larger language models, and I expected that maybe natural language and machine translation would benefit from these.

I think the overwhelming majority of computer scientists anticipated what would happen with large language models. I recall my first use of ChatGPT, and remember I was aware of what was going on, but I was profoundly shocked at the

behaviour of the system, at what it could do. Indeed, despite the fact that I knew how it worked, actually in some reasonable technical detail, I could not comprehend it. I simply didn't understand what massive scale amplified, admittedly, by some neat engineering would yield.

I had to repeat to myself that I was not seeing search, but rather the results of a statistical process giving rise to predictions at the level of words, tokens and text fragments. To this day, whenever I use these tools, I still struggle to really intuit what those models are doing, and it's really important we hold on to the shock, because we've crossed a frontier.

The implications, the impact of technology and the ways in which innovations are applied have been much studied. For advanced technologies, the translation from lab to broader uptake generally takes an extended period. There's a gap between big gaps, in fact, between early adopters, early majority, late majority and laggards.

The current state of AI

But this is not what's happening with AI: take-up is progressing extraordinarily rapidly. Productivity opportunities and applications are proliferating. The leverage that can be secured from integration with existing platforms and data resources are evidently very large and are emerging on an almost daily basis. This is so much the case that in business, impatient with even the accelerating pace of deployment, individual use for routine tasks has become commonplace. It is impossible to say with any precision where this might lead, not least in the context where the models themselves continue to develop at extraordinary speed.

Whilst AI has limitations, some of which I will discuss, many of these are intrinsic to its nature. Part of what makes it both powerful and fascinating. With reference to hallucinations, where the tendencies towards approximation, creativity and reinterpretation, these are not flaws. They're features that open up new ways of thinking and working. Understanding these tendencies is crucial to harnessing AI effectively, rather than risking dismissing it prematurely.

"To dismiss AI for its imperfections is to overlook the very essence of what makes it valuable: generative, associative, adaptive intelligence that complements rather than replaces human insight."

It's a system that generates, suggests, and adapts, with all that implies for its use. It is, in my view, a strange behaviour



to ask an AI tool to do things it's not designed to do and then present the inevitable shortcomings as evidence of its failure. It's very strange behaviour to apply AI, where its known limitations might yield adverse effects without applying obvious and readily available mitigations. And it is, in my judgement, the strangest behaviour to see an amazingly powerful technology and focus on its envisaged failings and not its immediate potential. **Scepticism is valuable. Dismissiveness is counterproductive.**

We should also keep in mind that **the advent of AI does not make all of the rest of computing suddenly go away.** We still have search. We still have information retrieval. We still have computational models and so on. **AI is an addition to the digital ecosystem, not a replacement for everything that came before.** Indeed, perhaps one of the most important digital developments is the convergence of AI with a range of other technologies. For example, virtual and augmented reality, amongst others. These convergences will multiply AI's impact, requiring us to think holistically about technological transformation.

An imaginary scenario

So, what does AI mean for children's social care? Now, I want you to recognise that I am an occasional tourist in your professional domain and not an expert. But let me set out an imaginary scenario and you can ignore the details or adapt accordingly.

So, a child is recounting their family and the relationships in the households with which they interact. This is as described, transcribed in real time. The voice, key relationships and names are extracted and correlated with other records, and inconsistencies are fed back in real time as prompts for the social worker. Queries are composed and can be sent to other external data sources: health, education, policing, justice. They can be made and catalogued. They can be reconstructed in privacy preserving ways. The responses can be risk assessed with the social worker reviewing those assessments.

A simple exchange format can be provided, augmented with information from the social work record that can be used in an automatically constructed workflow to construct a genogram, a standard pictorial representation for these relationships, and you can go on. A short video can be synthesised with an avatar of the social worker explaining the genus gram in terms the child themselves have used perhaps using any one of the languages that family might be using, perhaps supported by chat interaction, and a redacted and or synonymised version can be produced that can be shared with other experts.



Now, that's just one extended use case, but it's not a fantasy. It actually relies on technologies that are readily available now achievable using existing apps and services. And this scenario excludes all the obvious routine tasks: compiling reports, sending requests to other organisations, filling in forms, maintaining key information items in the record, scheduling, collating, searching and summarising. It excludes too training and testing.

What does AI really mean for children’s social care?

AI offers the potential for highly personalised and interactive support, adaptive in real time, generating new pathways and processes, evaluating progress dynamically and adjusting to different styles. **But let’s be clear: AI in children’s social care is not about replacing human interaction.** Face to face engagement is not going away. On the contrary, **AI has the potential to supplement personal engagement, making it deeper, more meaningful by automating routine tasks and freeing up social workers for what matters most:** listening, co-creating solutions and enabling access to critical resources. The role of the social worker will evolve, but it will remain indispensable.

Beyond children’s social care, the welfare, care, health, education and justice system are equally subject to change. **AI will not just reshape how we deliver social care; it will transform everything around it.** It will reconfigure the infrastructure. It will influence the policies, the funding priorities and the strategies. And this is not some distant pipe dream. The consequences are unfolding now and we will need to evolve structurally and technologically.

The pace of change is accelerating, though I recognise that there may be technological discontinuities that are difficult to assess at this moment. But adaptation is not optional. It is imperative. Ultimately, we will see children and adults who are different because of these changes. **AI will not just change the ways in which we provide care, it will change the service users themselves, their expectations, their cognitive processes, and the ways in which they engage. It’s going to change society. So, the real question is, of course, what can and should we be doing right now?** We must lean into change. We must start the dialogue. AI is here. We must engage with it rather than resisting. Waiting is not a strategy; active participation is. We must listen to both social workers and children and families.

They’re already using AI in ways that we may not anticipate. Ignoring this reality would mean failing to understand the people we need to support. I think there’s a question about who we’re watching and who we’re learning from. I wonder whether this may be a broader point about public service, about looking at others in public service any longer forms the right benchmarks for what we need to be doing about technology. So, are we benchmarking against organisations that also adapting or against outdated models of delivery?

Recommendations for change

- We must build the right foundations
- We must invest in core infrastructure and platforms. Without the right infrastructure, even the best AI strategies will fail.
- We must treat data as an asset because ultimately, our opportunity to make the most of AI will depend on our unique access to data. High quality data sets will define success. If our data is flawed or inadequate, so too will be our AI driven insights.
- We must experiment, iterate, and refine. Waiting for perfection is not an option. **Failure to experiment is a greater risk than experimenting and failing.**

A strong skill base will be a prerequisite. Some of this, but I suspect not all, can be achieved through training and development from within the existing workforce, but new expertise will be needed, requiring fresh approaches to recruitment and professional development. There is a need for the right partnerships, the existing partnerships networks. Dare I say, suppliers may not be positioned to deliver what we need. New collaborations will be essential and there will be a need to develop the capacity to invest in a plan. Organisations often fail to make it through technology transformations because at the point when changes required, they do not have the ability to make it happen because all their resources are locked into their existing capabilities.

Strategic foresight is critical and conscious of the time. So, we must address the critical risk issues head on. AI brings enormous opportunities, but real risks. Ignoring them does not mean make them go away; addressing them head on is the path to responsible innovation.

So, I think we need to evaluate what’s *not* done as well as what *is* done. Decisions by AI must be clearly understandable and accountable, with designated responsibilities for oversight and compliance. Critical decision making must follow a human in the loop approach, ensuring AI generated recommendations are reviewed by suitably qualified individuals.



Systems with high stakes applications cannot be fully autonomous, and final decisions must always be made by humans who bear accountability for these decisions. Ethical governance cannot be an afterthought. It's a necessity. AI systems can reinforce existing inequalities. Bias in AI must be identified and minimised, particularly in training data and algorithmic design. Bias audits and fairness testing are required to identify and mitigate discriminatory impacts before deployment without intervention. AI can amplify systemic biases rather than alleviate them.

Machine learning requires large scale compute with associated demands for energy, and thus consequences for sustainability that must be assessed and mitigated. AI's environmental impact must not be overlooked in the rush towards adoption. No technology is neutral. AI can be used to empower or to exploit, to mislead as well as to inform. This is a broader societal challenge. A powerful technology in the hands of an unethical organisation or individual poses a risk. Ethical AI is not an abstract discussion. It's a responsibility.

“Ultimately, AI in social care is not just about efficiency. It’s about possibility. But to realise that potential, it is necessary for professionals to be proactive, intentional and bold. Passivity will mean missing the moment. Collectively, we can’t wait for change to happen. We need to lead it. If we don’t, we’ll leave it to Silicon Valley to define our future for us.”