STEER

Signposting for steering minds

Draft framework for rating AI-based educational technologies

In 2011 the UK education company STEER began a programme of research around cognitive processes which could be described as 'machine-resistant'. The programme was led by DR Simon Walker, an applied cognitive biologist

STEER welcomed the potential benefits of machine learning. However, STEER identified that describing human cognitive functions that were unlikely to be replicated by machines was an educationally, economically and existentially important question. STEER speculated that such a study would inform sustainable and strategic education directions; safeguard uniquely human skills; qualify the remit of AI in education; inform moral development of AI and clarify themes of human identity.

To achieve this, STEER drew on specialisms from philosophy, psychology, cognitive and computer science, education and pedagogy. The STEER programme centred around a series of large quantitative school studies testing a theoretical data model which, previous research had indicated, might describe cognitive functions likely to be uniquely human. Whilst advances in machine learning were accelerating rapidly, the basic architecture and coding principles of machine learning had been established back in the 1980. Now applied at scale, and across networked machines, harvesting a wider variety of data, these processes were achieving rapid gains in replicating and beating some instances of human cognitive skill which had previously been considered uniquely human. Therefore, cognitive functions which did conform to this computational direction were considered machine-resistant.

The programme methodologies, outcomes and conclusions were published in a series of studies between 2012-15. The term 'steering cognition' or 'steering' was used to refer to the cognitive function described. STEER identified that steering contributed to both student mental health and learning-to-learn skills.

From 2013- STEER initiated a concurrent development programme to engineer a set of useable technologies which schools could embed to measure, track and improve steering in their students. Since then, more than 100 UK schools have implemented the technologies to track and support 30,000 children and 7,000 teachers over a multi-year programme. The programme is funded by schools who want to achieve better outcomes in student mental health, learning-to-learn skills and quality of teaching, by becoming a 'steering school'.

About the lead author: Simon Walker

Dr Simon Walker is co-founder and Director of Research at STEER and Honorary Fellow at Bristol Graduate School of Education. He holds Masters degrees in Biology and Applied Psychology from Oxford University, where he taught a graduate programme for seven years. His research demonstrating the role of steering cognition in adolescent mental health and learning, has led to the development of pioneering tools reducing mental health risks, and improving learning-to-learn skills, for more than 30,000 children across the leading schools in the UK.

The draft framework

In 2018 STEER began work to develop a framework for rating AI-based educational technologies. This framework seeks to integrate the full range of perspectives considered relevant to educational decision-making: moral framework, philosophical considerations, practical limitations, commercial realities, political risks, cognitive foundations, pedagogical processes and sound educational outcomes.

The development of AI-based educational technology should be understood within a longer tradition of methods used to improve education. Lessons can be learned from the success and failure of previous examples. In particular, education technologies have frequently been mistaken for educational benefits in themselves, rather than being seen as simply potential channels which may/may not enhance educational outcomes. Technology developers have sometimes been guilty of a lack of understanding of pedagogy, adolescent psychology and school realities, leading them to claim benefits which then prove to be unachievable.

At the same time, teachers have sometimes been seduced by the promises of new hard and software to dramatically improve or alter the fundamental basis of teaching and learning. Busy teachers sometimes de-value their own expertise believing too easily that technology developers possess some 'magic' to solve hard problems faced in the classroom. Numerous stories indicate that this faith is sometimes misplaced.

The promises of AI are, in the words of a machine-learning expert at Oxford University, currently experiencing 'market bubble characteristics'. A healthy degree of scepticism is required to see beyond the hype and evaluate both risks/ downsides as well as possible gains.

The draft STEER framework is intended to contribute an explicit and reasonable set of questions which can be used to scrutinise new AI educational advances. The framework is comprehensive but not exhaustive. STEER has not yet developed, or tested, the most appropriate method to rate the 14 items but envisages a 1-5 or 1-10 scale.

Future development of the framework

STEER recognises the need for agencies engaged in this area to collaborate. To that end, STEER will make the framework available to not-for-profit bodies who are committed to the responsible development of AI-educational technologies, subject to international copyright laws,

STEER envisages the development of the framework to improve its scope and application, prior to publication and distribution.

Draft framework for rating AI-based educational applications				
	Area	**Intended outcome**	**Sample question**	**Scale**
1.	**Machine vs human definition**	To clearly distinguish uniquely human cognition from machine cognition. E.g. types of datasets, enquiry and task which machines can / cannot contribute to educational improvement	To what extent does a new pedagogical technology enhance a uniquely human cognitive function?	
2.	**Moral / ethical development**	To safeguard the moral underpinning of AI coding from school upwards-through education, selection supervision provision	What provisions are in place to supervise the moral, ethical, political biases of those developing this AI technology?	
3.	**Psychological implications**	To recognise the varied ways in which AI may alter the social contexts of education and to evaluate its psychological consequences (e.g. VR, student facial monitoring etc)	What evaluation has been made of the wider psychological consequences for users of this technology?	
4.	**Educational management**	To evaluate impact of outsourcing cognitive 'engine load' from education systems for both: - an individual - an educational system e.g. the benefits and risk of AI-enhanced teaching for students but also education system management.	What research has been conducted to understand any unintended losses, or regressions, that result from users relying on this new technology?	
5.		To prioritise AI- applications which enhance human learning rather than unintentionally diminish human cognitive work	What provisions have been made to position this technology appropriately such that users use it to enhance and not diminish human learning?	
6.	**Cognitive objectives**	To focus on education which develops student metacognitive skills and expertise because of the 'governor' role such skills provide in supervising AI-applications	What provisions are in place to train student / teacher metacognitive skills in schools where this AI-technology will be deployed?	
7.	**Pedagogical scope**	To prioritise pedagogical processes which foster empathy, exploration, judgement, discrimination, rather than diminish them	By what method does this AI technology seek to improve these human skills?	
8.	**Real-world application**	To develop AI which has real world applications rather than simply class-room based application. E.g. to address current risks facing human decision making in an internet age, such as amplified group-think, fake news etc…	What efforts have been made by the developers to apply their technology to real-world problems and contexts? Have they been successful in better interpreting those contexts as a result of using this technology?	

9.	**Knowledge hierarchies**	To teach epistemology: to enable students to distinguish between data, information, knowledge, abstraction, felt experience such that they can discriminate between lower-level machine analytics and higher-level data interpretation.	How developed is the target educational system in positioning the new contribution of any AI technological data within a robust hierarchy of epistemological validity?	
10	**Practical constraints**	To evaluate the practicability of any new AI- applications in terms of the evidence that it can be feasibly deployed within the confines of the school environment, adolescent behaviour, limitations of available time, hard/software and teacher skills	How feasible is it for teachers to practically use this technology in-class/ school? What constraints of time, skill or costs exist? What intangible obstacles have been considered? E.g. reluctance of students to use 'wearable technologies' in the social setting of the classroom?	
11	**Developmental methodologies**	To evaluate the extent to which researchers have developed any proposed AI-applications through in-school action-learning rather than merely in a lab?	Was this technology developed in dialogue with school users?	
12	**Commercial and political perspectives**	To educate users that, like any other field, AI is being developed in a market in which both moral and amoral private and stare actors are acting to exert control for various ambiguous ends through data	What safeguards are in place to prevent abuse of this data for commercial or political ends? To what extend does this AI-technology improve the ability of students to interrogate motives and intents of real-world applications of AI?	
13	**Multi-disciplinary supervision**	To provide healthy critique involving a range of expertise and perspectives, not just the technological: philosophy, teaching, computer science, cognitive science. Recognising that no single discipline should be preferenced in legitimising the development and use of AI in education	What range of perspectives have been provided by the AI developers on its basis and implications?	
14	**Technology audit**	To maintain an up-to-date evaluation of AI- applications as exemplars, examples of emerging trends and in order to provide sector guidance	Is this technology necessary? Is it unique? Has it been tested already elsewhere? What is its significance and reliability?	