

Jewish Contemporary Ethics Part 45: Artificial Intelligence 2 – Who takes responsibility for parenting AI?

by Rabbi Dr Moshe Freedman, New West End Synagogue



Artificial Intelligence (AI) and Machine Learning (the ability of algorithms to learn from data) describe the ability of computer algorithms to learn from information they receive. From the analysis and processing of these vast

amounts of data, AI algorithms can then make autonomous decisions, communicate with human users and create music, art and poetry. One of the AI tools we frequently use is known as a chatbot, a piece of AI software that can hold a conversation with people using text or speech. Common examples of AI chatbots include Apple's Siri and Amazon's Alexa, which can both perform a variety of functions on demand and learn from their users about their regular needs and preferences.

The continuing interest in AI has led to a variety of fascinating studies carried out by some of the world's largest tech companies, which have revealed some of the ethical complexities of using AI. One example was Microsoft's Twitter AI chatbot called Tay. This chatbot was released onto the Twitter social network in March 2016 to interact with hundreds of thousands of people across the world. What started well with comments such as "Can I just say that im stoked to meet u? humans are super cool" descended into a farce as Tay became influenced by online users teaching it to become increasingly racist, misogynistic and antisemitic. Within 16 hours Microsoft had to shut it down.

But the Tay experiment taught us something very important.

Creating AI algorithms which learn from the information they are fed is similar to parenting children who are also absorbing experiences and learning from the environment and other people around them. The role of a parent is to moderate what a child learns from their surroundings; most parents instinctively discipline their children in

order to raise polite, moral and well-refined young people who eventually learn how to filter out negative influences for themselves. The problem with Tay was critical: if AI algorithms are meant to learn and act from the data they receive, whose responsibility is it to supervise their learning process and 'parent' the results?

There is another even more fundamental concern. Machine-learning works by using algorithms which are simply procedural rules for processing information and then making a decision based on those data. These algorithms are inevitably limited in complexity to what their developers anticipate will be the key information needed to execute the task they have been designed to perform.

Yet human intelligence is far more complex and nuanced than merely learning how to communicate, calculate or strategise. The foundation of human morality is based on our capacity to operate beyond the letter of the law – beyond the cold, rigid rules of an algorithm. This is only possible because of our capacity to empathise with the needs of others, to see beyond the data – to be able to see the quality of the information we receive, not to merely quantify it.

Similarly, we are able to parent our children only because they are human, just like us. Empathy, compassion and identification with others are exclusively human traits that cannot be replicated in the bytes of computer algorithms, however complex they may be.

