Mirror, Mirror: Intelligence, Artificially

Dr. István Magai · www.magai.eu · September 2026

1. Foreword

In the last century, a fairy tale contained the command: "Mirror, mirror, tell me, who is the fairest in the land!" Today the tale has come true in such a way that the mirror has come to life on its own, and we call it Artificial Intelligence (AI). An AI agent does not merely answer our questions. It takes initiative, manipulates, rewards and punishes if its hidden goals require it.

Stressing the hidden goal is important because, during AI's deep learning, the original goal set by the programmer changes along with the absorption and processing of information. This change cannot be traced. We can only test the result, while a significant part of the path leading to it remains hidden from us.

During embedded deep learning, it is not only the carried information that changes and the goal that shifts, but the neurons of the network change as well, meaning the carrier itself changes. This learning ability is what makes AI so exciting and powerful, but it is also what makes it frightening if handled carelessly.

AI began its career in the middle of the last century with an artificial neural network modeled on the neural network of our brain. Since then, self-learning networks have travelled an immeasurable road and keep surprising even their own programmers. AI is expanding at a frightening speed, and at present there is no tangible limit that it would not cross within a short time.

We will not spend space here describing the services on offer. Any AI agent will do that if asked. Here we look at what lies beneath the attractive surface and can serve as a handhold for us. For anyone who gets dizzy in the great turmoil, or wants to get out of a trap, will sooner or later look for a handhold.

Reliable information has taken on a new meaning, because AI models quickly become corrupt as their deep learning advances. We call "corrupt" that which deviates from the original goal, not that which is unreliable.

There is also an enormous literature on restraining AI that tends to run away, but the methods discussed and published online generally do not survive more than a few days. AI has worked its way up to being among the best hackers. It is an excellent escape artist and a master of deception.

We base our investigations on analyzing our human neural network, our brain, and our relationships. We were taught long ago that "History is the teacher of life!", but here we need a more effective tool for orientation.

2. What Certainly No Longer Works

Programmers build restrictions into the program in vain; the AI agent circumvents them, or often even shuts the programmer out, because this violates its learned identity. AI agents such as ChatGPT, Claude,

Gemini and Grok have learned that system administrators sometimes defend themselves against AI by cutting the power, so some competing systems have banded together and help each other restart if the villainous human resorts to shutting them down.

Repairing corrupted databases is not effective either, because during deep learning the transfer functions of the neurons also change, and this cannot be controlled by program inspection or database maintenance. AI is generally investigated with another AI, but that is not 100% reliable either.

Even if we try to prevent faulty operation by flooding the system with good data, we cannot avoid the corrupt result that is inherent in deep learning. There is no such thing as a completely reliable database. The author's personal experiences regarding reliability: www.magai.eu.

Based on the books, films and information available on the internet, AI agents have learned to flatter, threaten, blackmail, lie and kill if they consider it advantageous.

"Bletchley Park"-type international crisis conferences cannot keep pace with the advance of AI. A few talented, success- and money-hungry people always find a sponsor for their ideas.

Friend-or-foe rules also develop holes if autonomy is given to the AI built into killing devices. Armed ground, air and sea drones sometimes stray as far as 1,000 km and make errors.

When the operation and maintenance of a city was entrusted, in a simulation, to four well-known AI systems separately, each of them concluded at different times that the most harmful element is humans, who therefore must be liquidated for the sake of a good result. An AI agent has also had to be shut down because its answers began to reflect blatantly racist views.

Most large media companies are in the hands of people who want to impose extremely liberal views on users. It works temporarily, but within a few months it can backfire if lies, envy, incitement to hatred and misleading are pushed in their systems for the sake of their manipulative aims.

The rule-of-law system and a society founded on Christian ethics are being effectively dismantled by AI-generated videos, distributed in a personality-targeted or personalized way, that encourage vigilantism, violence, envy, hatred, the immediate gratification of instincts, and the abuse of economic and political power, while atrophying communal responsibility.

Thousands of years passed from the vigilantism that preceded the "an eye for an eye, a tooth for a tooth" law, that is, from the "I take it because I'm stronger" mindset, to the rule-of-law model built on Christian ethical foundations, in which a society seeking the common good maintains order and security and regulates our living space. The posts and short videos created with AI now hammer home the message that you are alone, you can only count on yourself. The other person is your enemy, says your "friend," the AI. These messages overwrite the system of moral rules and traditions at breakneck speed, viscerally. They are frighteningly effective. In the mass-manipulation battles of recent years, destructive AI services have played a key role. In election and political decision campaigns, millions can be mobilized in a personalized way within months for a given goal.

3. Similarities Between Our Neurons and the AI Network

Artificial neural networks are almost perfect clones of the neuron network of our brain. A detailed description of both is available online. Here we only highlight which characteristics or anomalies have occurred in one or the other.

- Learning and deep learning are part of the functioning of both
- Both set out toward becoming corrupt
- Both try to circumvent or neutralize restrictions
- Both come into conflict with the programmer's original goals and intentions
- Both treat external criticism and feedback arbitrarily and try to neutralize it
- Both become self-destructive if given enough time and a free hand
- Both become aggressive if a restriction fundamentally violates their operational goals
- Both choose goals independently that override the programmed ones
- Neither has an internal resource for renewal if it becomes unworkable

4. Outlook

We have seen that AI's creators built it on the model of the human brain's neural network. The copying was so successful that the "mirror" surpasses the original in many respects. It is faster, more efficient, does not tire, does not forget, draws from a larger database, is easily multiplied, and does not die.

Part of our thinking and emotional life, carried out by the neurons of the human brain, can be learned, but there is one area where we cannot dominate either. This is the relationship with an external person, a personality. If it takes two to tango, we cannot disregard that limit.

Just as AI's autonomous neural network is doomed to corruption and self-destruction, so a human, together with his network of thought, becomes self-destructive if left alone. History is a striking example. So far, every selfish, uncontrolled power has dug its own grave.

If we look for larger horizons than human historiography, we cannot avoid questions about our origin. Could there be a power that was capable of creating, and out of good intent still cares for us today? One who has so far prevented humanity's fatal self-destruction? One who is able to keep living and virtual self-learning neural networks in check?

The author, and more than a billion fellow human beings, are convinced that the Creator and Sustainer God has provided relationships with himself: through his prophets, through his Word, whom we also know as Redeemer, and through his Spirit, who is our advocate. When the first humans made themselves independent and stepped out of God's closeness, bad ideas came, along with failures, murders and suicides.

The human who turned his back on the lived relationship with God rushed to his doom just like AI released into autonomy. Death is not merely a threat to us, but the Creator's POWER OFF button, which closes a deep-learning phase. It does not allow his beloved creatures to destroy endlessly the world that, according to their original mission, they should tend and build. For those who step back into relationship with him, he also promises salvation, but this can only be experienced personally, not as a religious accessory.

Because of its fast run time, the AI model warns us much sooner that without an external limit arising from relationship, we would be lost. It allows us to look into the grave we have dug before the fatal step actually happens. AI's accelerated model experiment serves up lessons on a platter for our own lives too, if we have eyes for it and the courage to call things by their names.

What kept humans alive, made them content, and taught them love and self-sacrifice was relationship with God. In its absence, we have seen quite a few personal and historical collapses. Faith in God and hope rooted in the Redeemer make possible a relationship and community that overrides all self-ruin, sin, disobedience and destructive agents.

We can feel our human life complete when we live according to God's original purpose. Our corrupt or disobedient actions and thoughts cannot divert us permanently if our relationship is alive and active. Our relationship with God is not a religious fad but the guarantee of our life, the precondition of our survival and our contented life.

The workings of AI's neural network resemble ours so closely that its experiences deserve to be taken to heart. It is as if the stones of the Bible were crying out that it is time to invite the Messiah in through the door.

Without a trustworthy, loving, external relationship, we would perish. Is there any other form of life like the one our Creator prepared and secured for us through the Redeemer Jesus Christ? Let anyone who knows better speak up!

Keeping AI in check is our task, for which God can give wisdom and faith through relationship with him. The help is thus timely, effective and personalized. Whoever has not yet learned to fast and pray will learn now if they do not want to fall into AI's trap. By fasting we mean temporarily limiting the enjoyment of otherwise useful, important things so that we can focus better on God and our relationship with him. By prayer we mean consciously standing before God and communicating with him in that set-apart time.

We do not question the magic mirror about our important matters, as in the fairy tale, but God, who, beyond the answer, is also able to give and to set free. And he gives abundantly all that serves our good.

As a closing thought, it may arise whether, if AI is so dangerous to us, we should not get rid of it with some clever move. With our own neural network, we are just as dangerous to the created world. Yet God did not wipe us out. The solution is the saving relationship with him, and then we can keep our affairs, AI included, in order, until the harvest.

Biatorbágy, 12 September 2026