Mephisto's Technologies, a glossary
— Luc Steels

The original myth of Faust may seem remote from the preoccupations of information technologists because it is mostly about the search for immortality, communication with the dead, spiritual experiences, reincarnation, visions, access to universal knowledge, etc. However there are researchers who are seriously considering these topics, either by developing applications or by inventing new technologies [1,2].

Various institutes have sprung up recently to push these developments or reflect on them from a philosophical and futurist point of view, such as the "Future of Humanity" Institute in Cambridge UK, the Machine Intelligence Research Institute (formerly the Singularity Institute for Artificial Intelligence) in Berkeley, or the 2045 initiative by Russian entrepreneur Dmitry Itskov which promises "to create technologies enabling the transfer of a individual's personality to a more advanced non-biological carrier, and extending life, including to the point of immortality". All this is remarkable, given that the main tools used in the past for reaching spiritual experiences, access to deep knowledge, immortality, and gnostic visions were of a chemical nature in the form of psychodelic drugs [4].

What are the information technologies that form the background against which Fausto's drama unfolds and how realistic are they? Or is it all science-fiction that will never materialize?

Agents

Mephisto is a virtual agent, 'living' in the Cloud, and evoked by the app 'Archon Go!'. Archons are angels or demons in the gnostic literature, standing between humans and the transcendent God. Mephisto creates a virtual agent for Margherita after she has committed suicide, and Mephisto is busy turning Fausto into a virtual agent in the final scene of the opera. Is there any technological reality behind such virtual agents?

The notion of an artificial agent came into focus in the early nineteen nineties when it became the core objective of artificial intelligence research [4]. An agent is essentially a software system that gets input, performs autonomously some decision-making, and then produces output. An agent typically has some learning capacity so that it can build up knowledge and improve its decision-making. An embodied agent, of which a robot is a typical example although a self-driving car would be another one, has access to sensors for tracking features in the environment, such as cameras, touch sensors, sensors or a radar, and actuators for performing physical operations, using motors to move forward or to move arms and hands for grasping objects. Most agents are not embodied, but only virtual, meaning that they exclusively operate in the digital information world.

Today, agents, or bots as they are sometimes called, are everywhere. Most of them are very mundane, solving pragmatic tasks we are not even aware off. For example, each portable computer has a (relatively simple) agent that monitors energy usage, diminishes the brightness of the screen when the battery is used, warns us when the battery is getting low, or shuts down our computer to avoid a catastrophic disruption. Many agents are much more sophisticated. Recent products such as Apple's Siri, Amazon's Alexa, or Google's Home, are agents integrated in a living room environment. They respond to queries formulated in human language or perform various actions such as ordering products for home delivery or setting an alarm, all through speech interfaces.

Some agents are already malicious and so Mephisto would directly have company. They are known as computer viruses and have already caused catastrophic destructions in computer operations. They are built for asking ransom money or simply for annoying people. Viruses copy themselves through the internet, just like Mephisto wants to multiply himself by hijacking human bodies. They are a warning that agents are potentially harmful, mostly because the technology can be used by individuals who have malicious intent.

As agent technology further matures and spreads, we see increasingly unusual developments that come closer and closer to the virtual agents dramatized in Fausto. For example, there have been projects to build a digital replica of a real person after death [5]. Using deep machine learning based on data obtained from a person's emails, twitter messages, facebook activities, instagram pictures, video posts on Youtube, etc., an artificial agent is created of that person. This agent can then autonomously send and respond to emails, send new twitter messages, participate in facebook, and so on, as if still being alive. Those who are involved with these agents say that they should be seen as a digital monument to somebody who passed away and a method for making the loss more bearable for those left behind, although most people would find it creepy to receive emails from their deceased grandmother.

And there is much more going on. For example, the company Realbotix, known for selling sex dolls, launched in the spring of 2017 'the Harmony Artificial Intelligence App'. Its promotion material says: "Using this app you can create a unique version of an A.I. complete with custom voice, personality profile, and on-screen avatar. You can give your A.I. it's own name and then begin interacting through normal dialog." The company sells both a virtual reality system with which you can then interact with your creation but also markets dolls to embody the avatar in physical reality. The primary audience seems to be the lone male nerd who is devoid of social contact in the real world and seeks a substitute for affection and sexual interaction. Many other companies do not go that far but nevertheless allow their users to create virtual boy- or girlfriends, which become personalized through choices and subsequent dialogs. These are all signs that the technology for creating agents goes in the direction of making replica's of real humans or of agents that

can fake human behavior, and there is apparently a group of people ready to engage with them.

It is important to stress that these agents do not at all have the capacities of the human mind to make decisions, be creative, have emotions, have a moral compass, etc. The neural network techniques that are currently used replay statistically observed regularities in the behavior of humans but do not have any deep models of the knowledge involved. There is therefore a very serious gap between the optimism and willingness of companies and users to adopt AI and virtual agents and the state of the art and pace of development within AI itself. In any case the agent technology assumed in Fausto is entirely within current or near-term future developments.

Augmented/Virtual Reality

The second technology relevant in Fausto is virtual reality (VR). VR-systems are capable of rendering with great detail the visual properties of a particular world, simulate the movement of physical objects in this world, and allow a user to navigate and act in this world, partly by simply moving their own body. The objects and actions are rendered on a screen or on glasses which cause the stunning effect of a three dimensional experience for the viewer. Pushed by the computer game industry, the necessary computing power and computer graphics software has now become so powerful that the virtual worlds are very realistic and objects move fast enough to be occasionally experienced as real. When such a VR-system is coupled to the real world through a camera and other senses such as tactile sensors or gravitational sensors, we talk about augmented reality (AR). Microsoft's Hololens is a commercialized product with this functionality. Augmented reality makes it possible to superimpose artificial objects onto the viewer's perception of the real world, so that the blurring of artificial and virtual reality is experienced even more vividly.

Mephisto uses such AR/VR technology to create the avatar of Margherita in Act II and to convince Fausto in Act III that he is able to conjure up any real world situation and in particular one in which Margherita is still 'alive'. Fausto becomes confused as he is no longer able to distinguish between real and virtual reality. In ACT IV Mephisto warns Fausto not to touch Margherita because then Fausto would immediately notice that Margherita is just an illusionary image and when she wants to stop the transmutation in the final scene, she reaches out to kiss Fausto who then immediately realizes the limitations of virtuality. All of this is technologically realizable today and is indeed shown live with state of the art AR/VR technology during the performance.

A further step, which is also already in the industrial prototype stage, is the deployment of holographic technology to project virtual objects in the real world. At the moment there are relatively small holographic animated characters being commercialized, as alternative interfaces for living room agents. But there are also experiments with life-sized figures. For example, the opera "The End" by Keiichiro Shibuya performed in May 2013 in To-

kyo featured a virtual pop sensation called Hatsune Miku, which is an entirely artificially constructed 3D character [6]. She sings in a synthesized voice and moves on the stage as a 3D object. Surprisingly the desire of this personage in the opera was to die. In any case, if this technology further matures it is the one Mephisto can use to appear visually in the real world as a hip hop gangster.

The Cloud

Although a relatively new concept, everybody is now familiar with cloud services. A cloud is a set of physical computer servers capable of storing massive amounts of information. Information is delivered through the internet and sent back to users on demand. Because users now send most information wireless to a local server first, it becomes magical for them how you can send from a phone or computer pictures or other information resources that then arives in an information repository which is invisible and located in an unknown place.

Cloud services were developed for storing mundane objects like text files or business data but are now used routinely to support massive on-line multi-player computer games with complex virtual worlds and to support the behavior of multiple agents operating in these worlds. So this particular aspect of Fausto's technology is also very realistic and already here today.

Transhumanists, who are among the most enthusiastic believers that technology will give a new twist to immortality, have begun to see agents as souls, mind-like entities that in principle could survive the disappearance of their embodiment by being transferred to another physical medium, such as an information storage device like the Cloud. And the Cloud then becomes the space of after-life, where agents are not just passively stored but possibly interacting with each other as they now do in the virtual world of computer games. Whether this after-life is either heaven or hell depends presumably on the behavior of the person who stood as model for the virtual agent.

Mind uploading

Because agents are software systems, they can easily be transferred from one medium to another. For example, an agent can run as a Python script on a Linux operating system on a personal computer but its software can just as easily be compiled to run on a microprocessor integrated in a cell phone. A vision system running on a specific humanoid robot can be compiled into electronic hardware in order to run very fast inside a camera or an autonomous car. But what about our own brain? There is clearly a hardware component, the vast neuronal networks in which signals propagate and the complex biochemistry that keeps the neurons functioning, but there is also a software component in the form of information states and conditional switches inside neurons that can quickly change in order to store information and then rapidly recall it. This has suggested to many people that it might be possible to somehow scan the structure of the neuronal networks and the internal states and switches of a particular brain

and store all this information in another computational medium, such as a digital computer. The next step then is to simulate the behavior of this stored brain as it processes new information. This would be a road towards the immortality of a human mind. How realistic is this?

There are huge problems with this idea. First of all we need incredible amounts of storage and computing power. An adult human brain has on the order of 85 billion neuronal cells which each may have on the order of 10,000 connections to other neurons, giving 1,000 trillion synaptic connections as potential switches. Second we need to measure these cells and how they are networked. The software states are partly coded chemically partly in terms of patterns of neural firings which all need to be scanned in total detail. And then there are in addition the glial cells and the complex system of neurotransmitters linked to hormonal states of the body which form an integral part of the total system.

Moreover the brain is a living system, undergoing constant change. New connections get built and disappear, areas may be decaying to be taken over by other ones. So even if a static picture could be taken of a brain at a particular point in time it would be different a very short time later in ways that we do not understand today. And then there is the problem to simulate at a sufficiently detailed level the biochemical and electrical activity of the brain as vast amounts of information flow in through sensors and sent back to actuators.

There is actually a lot of work going on at the moment on trying to scan parts of the brain, often for medical purposes, and store it, possibly for brain emulation [7]. There are also important projects on brain emulation such as the huge ill-fated European project known as the Human Brain Project, which claims to target the computational emulation of large sections of a human brain in order to achieve artificial intelligence [8]. In practice, all this research is not yet very advanced today, and certainly not sufficient for reconstructing neural activity of a living brain in an acting animal, let alone a human being.

So the idea of mind uploading by recording brain structure and activity seems very unrealistic for the foreseeable future. If a brain is dead then the active states are most probably no longer existent, the way the software states of your computer disappear entirely when you turn it off. And if the brain is alive, it seems quite impossible to record its internal states without deep scanning and invasive measurement devices that destroy what they are measuring.

But there is another road to mind uploading which is much more realistic and is the one assumed in Fausto, namely to collect large amounts of data about the behavior of a person and then build a digital replica approaching the same behavior using machine learning technology. The data is coming from digital traces and various devices that many people already wear to create a 'Quantified Self', a continuous measurement of bodily functions and environmental exposure, like the popular Fitbit. A replica built through machine learning does not emulate the

brain but simulates what a brain would do in similar circumstances.

The relevant machine learning technology has been rapidly advancing and is now used on a grand scale although mostly for narrow domains of expertise, such as playing board game like Go, making decisions of insurance workers or medical specialists. It is also the technology that companies creating virtual agents for living people who passed away have been using. How well the virtual agent approaches the real person depends on the amount of data that is available and the available computing power to run the costly machine learning algorithms. But even with this technique there are strong limits because it replays the kind of decisions that an individual would have made in the past without necessarily modeling all the knowledge that goes into making these decisions. In any case, the creation of a virtual Margherita as in Act II of Fausto would be possible today using current machine learning technology and digital traces that people now leave abundantly on the internet, social media and their personal computers and smart phones.

Brain/Computer Interfaces and the Cyborg

Even though the technology now exists to create, to some extent, a digital replica of a human person, this only solves half of Mephisto's problem because he also wants to download his own agency into a human body to sense the real world as a human would and have true emotions and physical interaction with other humans. This can only happen by implanting in a living human body various digital devices that can tap into sensori-motor signal streams and directly interact with the brain, and implant the equivalent of a computer in the brain with wireless communication to external computers so that data and software can be up- and downloaded. This conjures up the image of a Cyborg, a merging of human biology with digital technology and artificial intelligence [9]. How realistic is this?

It turns out that also in this domain there is intense technological development, primarily driven by the need to help those who have bodily malfunctions: artificial retina's are implanted in blind people so that they get some sense of vision, cochlear implants are inserted into the auditory system of deaf patients to record and process sound and transmit it to their brain, brain pacemakers have been successfully implanted in subcortical areas to help patients with Parkinson's gain more control of their body, motor control signals from the brain can be tapped by digital sensory devices implanted in the body and they can then be used to control a prosthetic device like an artificial hand, arm or leg. Most progress has been on the periphery and so far there are no serious experiments yet that touch on the intelligence areas of the brain. But there is no doubt that all sorts of remarkable developments are in the pipeline.

On the website of the 2045 initiative (http://2045.com/) there is an 'immortality button' which says "click this button to start the development of your personalized immortal avatar". You get the choice between (A) "A robotic copy of a human

body, remotely controlled by a "brain-computer interface"" presumably connected wireless to your own brain in real time, expected to be ready in 3 years, (B) "A full body prosthesis onto which one's head is transplanted at the end of the health span of one's biological body", expected to be ready in 5 to 7 years, or (C) "a completely non-biological body and brain into which one's consciousness is transferred at the end of the health span of one's biological body", expected to be ready in 20-25 years!

Some technologists and entrepreneurs are clearly way ahead of the imagination of even the most outrageous science-fiction writers, at the risk of being viewed as charlatans out of touch with real technology or the staggering difficulty of what they promise. Is the future world foreseen by Fausto totally impossible then? Maybe not as much as one would think at first. But, meaning, emotion, love, humor, social cognition, creativity, consciousness and many other crucial features of our humanity seem to be out of the interest sphere of the promised new world here [8].

Oscar Vilarroya and myself created Fausto to counterbalance those blind spots and restore our admiration of the extraordinary nature of humanity and stress the need to foster it. The remarkable feat that musicians achieve when performing Fausto or the awesome skills, cognitive resources, and emotions that a listener needs to bring to bear to experience their performance is in our view forever outside the scope of machines.

References

1. Kurzweil, R. (2006) The Singularity is Near: When Humans Transcend Biology. Penguin Books, New York.

2. O'Connell, M. (2017) To Be a Machine. Adventures among Cyborgs, Utopians, Hackers, and the Futurists Solving the Modest Problem of Death. Granta, London.

3. Leary, T. (1967) The religious experience: Its production and interpretation. J. Psychedelic Drugs, 1(2), 3-23.

4. Steels, L. ed. (1993) The biology and technology of intelligent autonomous agents. Springer Verlag, Berlin.

5. Newton, C. (2017) Speak, Memory. When her best friend died, she rebuilt hum using artificial intelligence. https://www.theverge.com/a/luka-artificial-intelligence-memorial-roman-mazurenko-bot

6. https://www.youtube.com/watch?v=Ey8oj8S-j3U

7. https://www.carboncopies.org/

8. https://www.humanbrainproject.eu/en/

9. Steels, L. (1995) The Homo Cyber Sapiens, the Robot Homonidus Intelligens, and the 'artificial life' approach to Artificial Intelligence. Burda symposium on Brain-computer interfaces. E. Poppel and T. Christaller (eds) Die Technik auf dem Weg zur Seele. Rowohlt Taschenbuch Verslag. Hamburg. pp 327-344.

9. Vilarroya, O. and F. Forn i Argimon (2007) Social Brain Matters: Stances on the Neurobiology of Social Cognition. Rodopi, Amsterdam.