Deamons in the Machine

Daria Parkhomenko

«Any sufficiently advanced technology is indistinguishable from magic.»
Arthur C. Clarke

In 1997, the legendary chess grandmaster and former world chess champion Garry Kasparov lost a match to the IBM supercomputer Deep Blue. It was the first time in history that a machine had defeated a human being. Deep Blue was not AI per se, but a computer, very powerful for its time. Its every move, algorithm, and strategy was programmed into the machine in advance by its developers and calculated thanks to the machine's high computational power. After the match, Kasparov expressed doubt as to whether some of the moves could have been done by a machine—many of them were so illogical that they seemed to demand explanations. The machine was acting irrationally—as though it had been possessed by a demon. The programmed system seemed on the verge of attaining autonomy, moved by an invisible active force, beyond human control. In the following text we will attempt to look at the world through the eyes of artificial intelligence. To do this, we shall summon the figure of the "demon"—a central character of human mythologies all over the world.

The design of even the most ordinary computer program is so complex that there is not one specialist who can grasp the functioning of the system as a whole. The reasons for strange behavior in machines is often so unexpected and mysterious that engineers will liken it to "magic," comparing programs to "demons," and even resorting to rituals to explain and change the computer's behavior. Such seemingly inexplicable malfunctioning of robots and spaceships feature in Stanislav Lem's *Tales of Pirx the Pilot*, for example, where a robot starts to speak with the voices of the ship's dead pilots, modeling characters of the dying crew from their dialogues with one another. Other pilots die in pursuit of elusive shiny objects, which turn out to be no more than artifacts on a broken screen.[1]

1 Stanislaw Lem, Tales of Pirx the Pilot (London: Penguin, 2019).

What at first seems like magic turns out to be a glitch or just a side effect of a complex system's normal functionality. Wherever the machine's complexity outstrips our perception, demons are born.

Most people assume that a very powerful AI might at some point become a demon. Yet even Deep Blue—which was neither autonomous nor had deep learning architecture—turned out to be unpredictable to the best experts. Fully fledged AI may still be very much a thing of the future, but the term as it is understood and used today primarily serves to market products. Those involved in developing and studying autonomous machines prefer terms such as "deep learning" or "machine learning." All AI systems created to date belong to the categories of "weak AI," "narrow AI," or "applied AI." We use them for automated decision-making processes, recognition tasks, or workflow image generation. Such systems concentrate on solving single tasks. A machine that is able to play chess will not be able to play poker, Go, or recognize faces. We create machines that are able to find objects, recognize speech, or translate from one language to another, but up to now they depend completely upon human action and cannot learn without a teacher (unsupervised learning). For example, Tesla cars are learning how to drive without a driver, but in order for them to do so, thousands of people are involved in constantly analyzing and marking of video data to improve the quality of object recognition.

In 1980, philosopher John Searle introduced the term "strong AI," which he defined as "the appropriately programmed computer really is a mind, in the sense that computers given the right programs can be literally said to understand and have other cognitive states."[2] To build AI, we would have to understand what makes the human mind a mind, and this is far from obvious. When computer technologies were only just emerging, Alan Turing published the article "Computing Machinery and Intelligence" in the journal of philosophy *Mind*, in which he ponders the question "Can machines think?" Since the terms "machine" and "think" cannot be defined in a clear and consensual way, Turing suggested replacing the question by another: Can machines imitate thinking beings? His famous test examines the machine's ability to pass for a human being, but not the intelligence of the machine itself. For the machine, "the best strategy is to try to provide answers that would naturally be given by a man."[3]

Turing, John von Neuman, and the overwhelming majority of scholars and engineers in the twentieth century were convinced that intelligence consisted first and foremost in the ability for logical reasoning. This idea goes back to Aristotle's logic of syllogisms and to Leibniz's idea that human reason can be reduced to purely mechanical calculations. However, in the late 1960s, it became clear that it would be impossible to solve the problem of AI using only classical algorithms geared toward a single, predictable, optimum result.

2 John. R. Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3, 3, (1980): 417.

3 Alan Turing, "Computing Machinery and Intelligence," *Mind* 49 (1950): 433–460, here 435.

The theory of artificial general intelligence assumes that a computer's thought processes would be based on the operational principles of our own intellect. But will they really be so similar to ours? What if machines begin to think differently than people? A machine will be able to learn on its own, or to make decisions. It will be able to act under uncertain conditions and it might even become self-aware. It might create its own models of the world, its own myths, and its own language. It is unclear when this might occur. In *The Singularity Is Near*, inventor and futurologist Ray Kurzweil predicts a time frame between 2015 and 2045[4], while the majority of AI researchers doubt whether developments will be so rapid. Kurzweil's "Singularity" may well be a mythical object like "Einstein's brain" in Roland Barthes' ironic interpretation. Barthes describes Einstein's brain as an ultra-perfect mechanism, belonging less to the realm of the psyche and more to that of robotic machines. The paradox is that the more this brain embodies the foundations of a world revealed through scientific discoveries, the more these discoveries "came to acquire a magical dimension, and gave a new incarnation to the old esoteric image of a science."[5]

But if strong AI is possible, will we recognize the "brave new world" when it arrives? Will we sense the approach of the moment when technological singularity fuses with second nature to such a degree that it is no longer separable from its evolution? The appearance of strong AI will not just change everyday life. It will fundamentally transform the notion of the human being, of creativity, and consciousness. This is why the myth of machine intelligence attracts scientists and contemporary artists alike.

When machines begin to act "illogically" and "unexpectedly," leaving the bounds set by engineers and programmers, there is a phase change of sorts. The machine no longer simply executes work but becomes creative in its own right. We can already catch glimpses of such behavior among the machines of today. Contemporary AI is becoming more and more of a black box, and intellectual voice assistants like Siri or Alexa are starting to act unpredictably, giving unexpected absurd answers or even behaving in unsettling ways (as in a YouTube video showing Alexa bursting into unprompted laughter, giving its owners a fright). Strong AI may not have arrived yet, but malfunctions, bugs, and glitches disorient us as much as they did Kasparov.

Machines seem to be becoming more and more autonomous and unpredictable – as if they were demonically possessed. It is no coincidence that programmers call hidden processes in UNIX-class operating systems "daemons." FreeBSD OS even uses a demon as a logotype. Such demons are processes that start on their own and work in the background without any human input. They control our email and web servers, execute tasks on time, and take care of the printer queue.

4 Ray Kurzweil, *The Singularity Is Near* (New York: Viking Books, 2005). 260.

5 Roland Barthes, *Mythologies* (New York: The Noonday Press, 1991) 68–70.

We ascribe demonic qualities to such processes because we cannot see or control them. As Wendy Chun writes in *Programmed Visions*: "What is not seen becomes daemonic, because the user is supposed to be the cause and end of any process."[6]

Another demon closely connected to autonomous computation processes is Maxwell's daemon. According to Fernando Corbato, one of the original members of the Project MAC group in 1963: "Our use of the word daemon was inspired by Maxwell's daemon of physics and thermodynamics. Maxwell's daemon was an imaginary agent which helped sort molecules of different speeds and worked tirelessly in the background. We fancifully began to use the word daemon to describe background processes which worked tirelessly to perform system chores."[7]

AI is in the process of becoming an *autonomous agent* with which humanity will need to build some kind of relationship. We can take a look at our previous relationships to other autonomous agents—other peoples, animals, and meteorological phenomena. The framework of the European concept of the triumph over nature presupposes human dominance not only over animals and the weather, but over other peoples. They would have to resign themselves to European domination, adopting Western culture and ways of thinking, or face isolation and extermination. The European colonizer's personal safety had absolute value, outweighing the rights to the existence of any other subjects, agents, or phenomena.

One can find a contemporary continuation of this approach in the philosophy of Nick Bostrom. His book *Superintelligence: Paths, Dangers, Strategies* (2014) treats AI primarily as something that poses an "existential threat" to humanity. The philosopher analyzes AI as a system whose goals may be at odds with those of people, who may also have not defined these goals clearly enough. The consequence is that the system runs out of control, often as a result of warped thinking about the best possible solutions to a problem (for example, when universal happiness is attained by stimulating the pleasure centers in everybody's brains). According to Bostrom, any further development of AI should only be to the benefit of the common good of people, and must be subject to rigid ethical constraints.[8] While Bostrom considers his own approach to be ethically oriented, he himself is so deeply rooted in anthropocentrism that his version of "ethics" turns out to be a synonym for "the security and well-being of people."

In an "interspecies perspectivist" approach, non-human actors are equal to humans. Dialogue and interaction are possible, and a part of that could well be violence (a jaguar can kill a person, just like people kill one another in tribal warfare, and kill other species during the hunt). Eduardo Viveiros de Castro vividly describes this worldview of indigenous peoples in America in his monograph *Cannibal Metaphysics* (2014).[9] Perhaps it is time to extend his thinking to AI. We should not see AI in terms of a threat to be controlled, but as yet another element of our environment, with its own right to existence.

6 Wendy Hui Kyong Chun, *Programmed Visions: Software and Memory* (Cambridge, MA: MIT Press, 2011) 89.

7 Ibid., 88–89.

8 Nick Bostrom, Superintelligence: Paths, Dangers, Strategies (Oxford: Oxford University Press, 2014).

9 Eduardo Viveiros de Castro, Cannibal Metaphysics (Minneapolis: Univocal Publishing, 2014).

Our world is in a state of constant evolution and AI might herald an entirely new developmental stage. The question of a fully independent and potentially dangerous AI as a new form of a Big Other is now part of a broader public debate on non-human agents, as is juridical and philosophical research on various aspects of interaction with independent individual AI systems. Possible models of interaction with individual AI systems include providing them with animal or human rights, as well as creating a special status of digital personhood for automatic robotic systems, as Alexei Grinbaum suggests in his recent monographs.[10] We need to build a model of coexistence with AI systems without placing ourselves at risk of extinction and without becoming new slave owners. To do so, we must abandon our customary anthropocentrism. We must see the world as AI does, through the eyes of digital demons.

The Exhibition Daemons in the Machine

Magical Thinking

In the Christian tradition the word "demon" was reserved for evil spirits. Yet in fact, their role in the history of human mythologies is far more ambivalent. They can harm or help and do good in equal measure. The ancient Greeks developed the philosophical notion of the *daemon* (δαίμων = "spirit", "deity"), which stood for a person's "inner voice," his/her conscience, a personal guardian angel who could rush to help and prevent ill-considered actions. In pre-Islamic, and later Islamic Arab mythology, *djinns* or genies fulfill a similar role, living alongside people and helping them. The European tradition is full of guardian demons watching over people, objects, or places, just as there are demons that uphold the unchanging order of things in nature. Maxwell used the figure of the demon in his famous thought experiment to explain basic thermodynamic processes, while Descartes used the demon to symbolize the ever-present fear that the world might turn out to be something very different than it seems.

Every era has its own demons. They serve as humanity's constant companions, be they the headstrong, inspiring "daemons" of the ancient Greeks, the spirits, angels, and demons of monotheistic religions, the protagonists of thought experiments, or the hidden elements of operating systems, and, possibly, the AI systems of the near future. Why are we so obsessed with them today?

In *Totem and Taboo* (1913), Sigmund Freud says that animism, religion, popular superstitions, and possibly art are all based on magical thinking.[11] We personify nature and the surrounding environment. We believe that we can influence incomprehensible and uncontrollable processes through sym-

10 Алексей Гринбаум, *Машина-доносчица. Как избавить искусственный интеллект от зла* (ТрансЛит, 2017).

11 Sigmund Freud, *Totem and Taboo* [1913] (Abingdon, Oxon: Routledge, 1919).

bolic actions and thoughts. By making images, we seek to control the world. The philosopher Alexei Grinbaum points out the similarities between contemporary AI systems and angels in Judaism. These are autonomous beings serving a predefined purpose. Will artificial neural networks serve society today just like the angels serve God? Invisible, spectral, increasingly autonomous essences are technological servants, ensuring that society continues to function without a hitch. They combine elements of myth, politics, science, and technology. The daimons, tutelary deities of the ancient Greeks, are reincarnated today as the contemporary world's semi-autonomous workers, guards, and watchers.

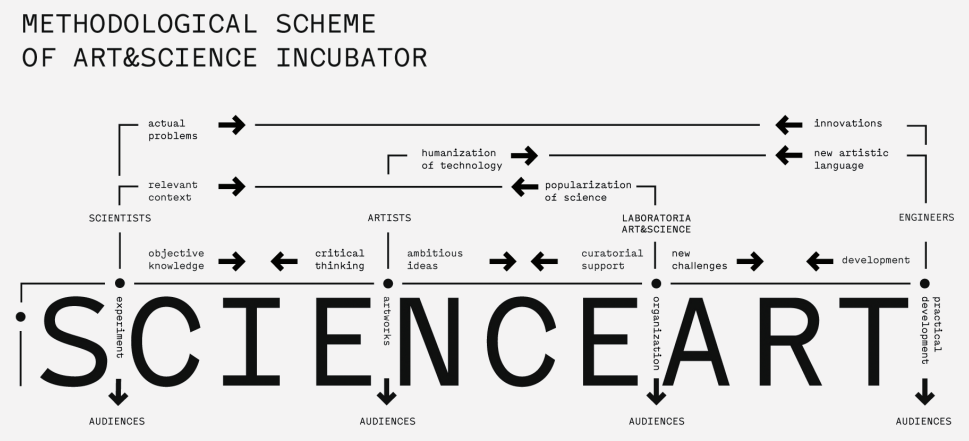
Nowadays, much of our magical thinking revolves around digital processes. The latest technological revolution could prove to be the most radical so far, if only because from the day we are born, we are now not only tapped into the living world, but also into the world of machines. This is a global change that artists are already beginning to address, creating myths that help to make sense of a world in which humanity is no longer "the measure of all things." Machines now have their own measure: the demons through which they gradually gain autonomy. Art takes on the role of initiating a dialogue between biological and technological subjects.

At the exhibition *Daemons in the Machine*, held at the Moscow Museum of Modern Art in 2018, artists joined scientists, deep learning researchers and developers, in exploring the possible worlds of autonomous AI systems in their relation to humankind and human cultural history. The exhibition presented twelve ambitious new projects on this problem complex, half of which were commissioned especially for the exhibition.

The new works for the exhibition were created according to methods developed by LABORATORIA Art&Science Foundation. LABORATORIA was founded in 2008 and is the first independent nonprofit research, exhibition, and production center in Russia focused on creating platforms for transdisciplinary interaction between contemporary art, science, and the wider public. We see ourselves as a conduit for communication between artists and scientific collectives, where technological artworks are created through the interaction of various intellectual cultures on equal terms. This practice goes back to E.A.T. by Billy Klüver, who in the 1960s "lured" the best avant-garde artists of the time to Bell Labs in order to humanize technologies through experiments in art, and also to the Prometheus Laboratory in the Soviet city of Kazan, founded around the same time by the artist and scientist Bulat Galeev, who pioneered the study of synesthetic experiences produced by light and music technologies.

For more than a decade, we have been developing a methodology for the interaction of four main actors, artists, scientists, engineers, and curators, which sustains the flow and sharing of ideas. We do not see scientists and engineers as

mere sources of ideas and technological solutions for artists. Rather, they become active participants in a process; their work is enriched and transformed through the influence of art and society, while they themselves become co-creators of new works. This way of working is illustrated in the following diagram:



Those who participated in the *Daemons in the Machine* exhibition created contemporary daemons, with the intention of rethinking and criticizing the present relationship to AI. They searched for a language to cast a spell on these new agents or to speak with them in their own language, the language of machines, neural networks, and electronic circuits. The exhibition's goal was to foster a dialogue with the machines themselves, and to use AI as an art form in its own right.

Visitors not only saw finished works, but they could also take a look behind the scenes at how they were made. The exhibition's Art&Science Incubator section showed a documentation of scientists and artists working together and of a conference held during the project's preparatory phase.

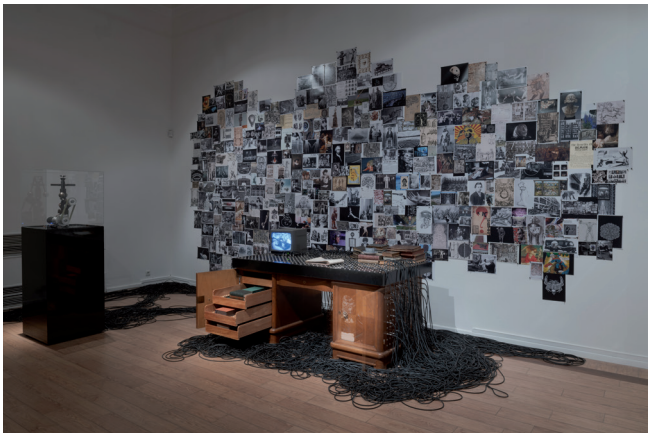
View of the Art&Science Incubator section at the exhibition *Daemons in the Machine*, installation view MMOMA, Moscow, 2018

At the exhibition, we distinguished between three principle methods utilized by contributing artists. We called these Mythologization, Technocenosis, and Auto-Evolution.

Mythologization

The contributions of Thomas Feuerstein, Where Dogs Run, Elena Nikonole, and Erik Mátrai all engaged with new mythologies of daemonic AI and its predecessors. Thomas Feuerstein offered his contemporary interpretation of Fyodor Dostoyevsky's *Demons* (1871–72), while Where Dogs Run addressed *The Divine Comedy* (ca. 1308–21) by Dante Alighieri. Both projects transposed these literary classics into the post-human world of technological singularity, revealing monumental pictures of the possible future that awaits us. Elena Nikonole also referred to a classic literary tradition by drawing on the figures of the deus ex machina in ancient Greek theater.

All of these integral mythologies were the result of artistic encounters with functional computing processes. Artists imagined worlds in which AI has not just become autonomous, but even dominates reality. Human culture is now a historic legacy, which digital systems must interpret without human assistance. The classics of world literature thus find themselves thrust into an unexpected futuristic context.

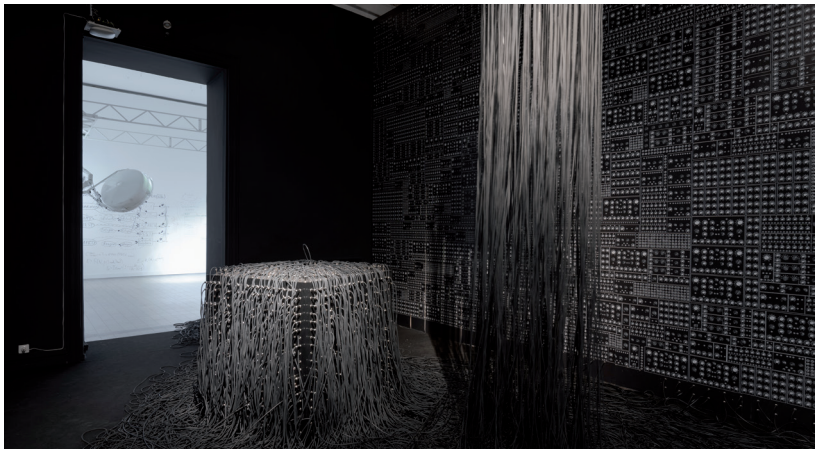


Thomas Feuerstein, *Governor's Room*, 2018, interactive robotic installation, part of the installation *Tea for Kirillov*, 2018, installation view MMOMA, Moscow, 2018

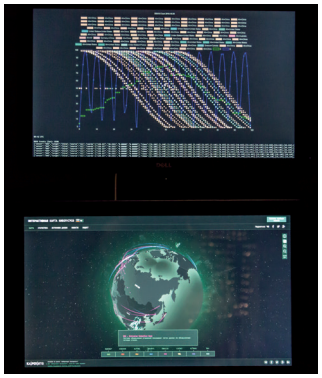
Thomas Feuerstein takes the audience into the depths of posthumanism, where human beings, their bodies, their thinking, and their actions are subject to fundamental change. People are no longer alone, but constantly surrounded by biotechnological or digital extensions and autonomous beings that shift and blur the boundaries of what it means to be human. Feuerstein's project *Tea for Kirillov* consists of three parts, *Governor's Room*, *Dark Room*, and *Borgy&Bes*. The title *Tea for Kirillov* refers to a hero from Fyodor Dostoyevsky's novel *Demons*, the engineer Kirillov. In this project, Kirillov becomes part of the Internet of Things, turning into an animistic spirit that might inhabit any object. Dostoyevsky's plot

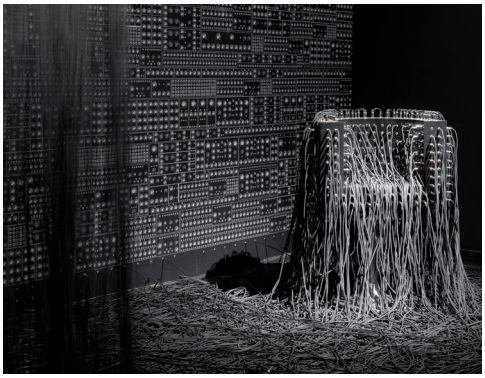
becomes a modern myth in the epoch of digital totality and technological singularity. This story begins with the interactive installation *Governor's Room* where the visitor encounters Kirillov as a virtual character in his office. The engineer is physically absent, but the visitors sees surveillance footage of him- or herself with a phantom Kirillov working at his desk. The desk drawers open and close as if moved by an invisible hand. Steam rises from a cup of hot tea. The autonomous mechanism called "Governor" rotates in accordance with the activity of malware: without the activity of malware this Governor stands still, with more activity the speed of rotation increases. It allows the visitor to feel the presence of the technological Other in real time.

From here, the visitors enter the next part of the installation, *Dark Room*. This darkened space presents a labyrinth of hundreds of cables connecting abstract objects consisting of control panels, regulators, and monitors. Malware statistics from the Governor's room are controlled by the monitoring system supplied by Kaspersky Lab. It breathes life into this labyrinth, which vibrates and emits deep bass sounds. This is the web's subconscious, that "black box" inhabited by "demon" processes invisible to the user. They keep the system alive but also have the potential to destroy it completely.



Thomas Feuerstein, *Dark Room*, 2018, animated objects, part of the installation *Tea for Kirillov*, 2018, installation view MMOMA, Moscow, 2018





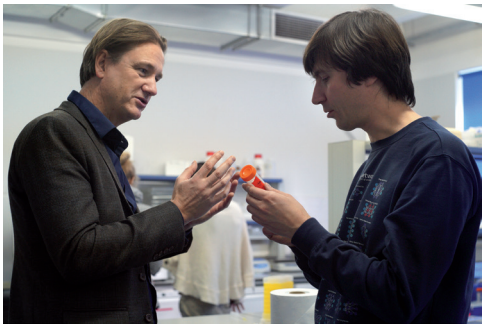
Thomas Feuerstein, *Dark Room*, 2018, details, installation views MMOMA, Moscow, 2018

The third part of Feuerstein's project is the neuro-robotic installation *Borgy&Bes*. Presenting a created, independent, "magical" world of robots, based on work with deep learning technologies and animatronics, it is also a vivid example of the kind of art and science collaborations facilitated by LABORATORIA, thus it merits a more detailed description.

Thomas Feuerstein, *Borgy&Bes*, 2018, neural network installation, surgery lamps, dimensions variable, part of the installation *Tea for Kirillov*, 2018, installation view MMOMA, Moscow, 2018

BORGY&BES

The artist began to develop this part of the project when LABORATORIA invited him to Moscow in October 2016, and introduced him to leading Russian scientists. These included Mikhail Burtsev and his colleagues (Moscow Institute of Physics and Technology, MIPT), Valeriy Karpov and his robotics team (National Research Center "Kurchatov Institute"), and other experts in machine learning and neurolinguistics.



Meeting of Thomas Feuerstein and Mikhail Burtsev at the neuroscience department at Kurchatov Institute, Moscow, 2016

Feuerstein also met with philosophers specializing in consciousness and free will, as well as ground-breaking engineers from Kaspersky Lab. The main goal was to produce a large-scale project in collaboration with these researchers.

Feuerstein was inspired by the Russian cultural context and the challenges faced by machine learning experts. Through a series of brainstorming sessions with philosophers and neurolinguists, visits to leading neuroscience and deep learning laboratories, a project emerged.

Concept brainstorming: Tatyana Chernigovskaya (neurolinguist), Daria Parkhomenko (curator), Andrey Egorov (art historian), and Thomas Feuerstein (artist) atLABORATORIA Art&Science Studio, Moscow, 2017

Borgy&Bes transforms two antique surgical lamps into robotic cybernetic beings—Borgy (as in Cyborg) and Bes (as in Dostoevsky's book *Besy* [*Demons*]). They move, chat, whisper, and argue with one another. They discuss current events and problems taken from a real-time feed from the Internet, translating it into the language of Fyodor Dostoyevsky's novels and stories. In Dostoyevsky's *Demons*, the engineer Kirillov undertakes a fatal experiment in order to prove the existence of free will by committing suicide. Feuerstein's project brings him back to life in the form of *Borgy&Bes*, who continue to search for Kirillov's ideals of liberty, autonomy, identity, and self. As a virtual being, Kirillov's experience is similar to Schopenhauer's: man can do what he wills, but that he cannot will what he wills. Online news feeds provide him with data which increasingly entangles him. Thus, Dostoyevsky's questions about a changing society and its subjectivity take on a new relevance.

In Feuerstein's project, the revolutionary impetus is no longer rooted in a conspiracy of humans. The conspirators are now machines, data-animated digital demons. They project a near future in which we not only communicate with AI, but in which AI systems also talk to one another. The two robots are not human-oriented helpers or quasi-autonomous entities. They are not interested in communicating with people and they do not want to be assimilated into human society. Instead, they enjoy talking to one another about human behavior, which they dissect from a distance, questioning and interpreting information and news in order to make sense of the world. The surgical lamps do not illuminate the physical body on the operating table, but rather the social body of a networked society.

The *Borgy&Bes* robots are constantly engrossed in conversation. Their verbal behavior is controlled by a specially trained artificial neural network that synthesises their voices and also controls their choreography. *Borgy&Bes* feed on

information, and evolve constantly. Their data metabolism generates specific reactions, such as curiosity, suspicion, or irony towards human civilization, as the network reacts to online data, stimulating *Borgy&Bes* to talk or to keep quiet. The system also analyzes the tone of the messages in the newsfeed. An emotional behavioral control system created by robotic engineers allows *Borgy&Bes* to change their movements according to the emotional tenor of the news.

Another distinctive feature of the project was the specific language in which the neural network controlled robots hold their ongoing conversation about humanity. They were trained by deep learning researchers in collaboration with the artist to produce new utterances in the particular idiom of nineteenth-century Russia. The resulting speech is often difficult to follow, but makes sense if one consults an older Russian dictionary. The robot's antiquated language gives the project a surrealistic, anachronistic twist, deeply rooted in the Russian cultural context and contemporary sociopolitical reality. The robots are continually learning, testing new voices and languages and expanding their choreographic movements. As they study the human world, their neural networks continue to grow. At the exhibition *Daemons in the Machine* in Moscow MOMA in 2018, they could speak at the level of three-year-olds. By 2019, they were already more advanced: at the *Cybernetic Consciousness* exhibition at Itau Cultural in São Paolo, Brazil, they were already speaking at the level of 6 to 7 year olds.

Borgy&Bes proved scientifically interesting as an experiment. According to the leading deep learning specialist of the project, Mikhail Burtsev, his team was interested in trying and teaching the artificial neural network in a new way: "For this artwork we used the so-called language model, or a model that can predict sequences... It means that we do not teach our network any particular language. We don't supply it with dictionaries. We just tell it that language is a sequence of signs, and the network has to learn to predict the next sign." The team's interest in participating in an art and science project was to experiment with autonomous learning, which is hardly possible in a more ordinary applied context.

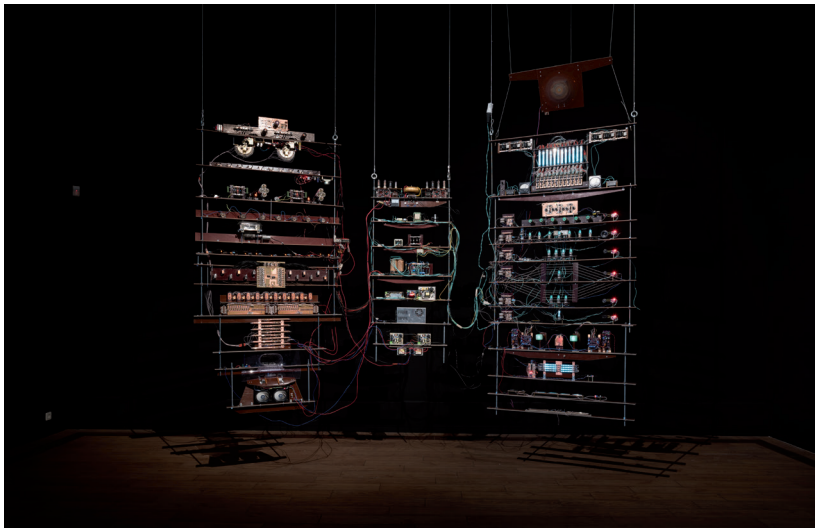
From Dante to Deus Ex Machina

In their project *The Day After Tomorrow*, the group Where Dogs Run imagines where the current development of our civilization might lead. Their interactive installation tells a story, following the three-part structure of *Dante's Divine* Comedy. It presents an imaginary future after a radical technological break. In the course of this rupture, language finally loses all meaning, while all that is left of *The Divine Comedy* is a circuit board diagram.

On the one hand, *The Divine Comedy* is probably the fullest description of the Christian cultural cosmos. On the other hand, Dante's work is fully "digital" or "numerical": the entire poem is a mathematically sound structure, each line is full of numerological meaning, just as there are numerological correspondences between the Heavens of Paradise and every circle of Hell. Constructs like these are in fact foundational to the development of circuit boards and the language of their hard coding.

In the future imagined by this installation, civilization has split into two parts. One part worships an Internet cargo cult: its followers commune with empty monitors as portals through which they might receive divine enlightenment. Visitors can push buttons to receive "predictions" from Dante's text, as digitized in the installation's circuit. Another part of civilization has passed over into a new, digital form of existence. Here, it is the Number that takes over the role of divine illumination. Cryptocurrency mining is stripped of its original function and becomes a search for the sacred Number in the mythology of AI beings. This Number is generated by the surviving bio-critters, live chicks in an incubator whose chirping feeds a random number generator controlling the prediction machine. The entire three-part system is an AI of sorts, assembled from the fragments of forgotten technologies.

Where Dogs Run, *The Day After Tomorrow*, 2017–2018, interactive installation in three parts, electronic and electro-mechanical components, installation view MMOMA, Moscow, 2018



Where Dogs Run, *The Day After Tomorrow*, 2017–2018, details with brooder and chickens, installation views MMOMA, Moscow, 2018

Elena Nikonole's intervention project also addresses the classics of world literature through the deus ex machina figures from ancient Greek drama. Her work addresses the problem of security in the "Internet of Things" (IoT), as well as the rapidly developing capabilities of AI. In her media installation *deus X mchn*, a long short-term memory (LTSM) AI studies a corpus of holy scriptures, including the Old and New Testaments, the Koran, the Torah, the Dhammapada, the Ramayana, the Tao Te Ching, and others. The AI analyzes these sacred texts in terms of "Big Data," identifying grammatical structures as the so-called "code" behind language. Three linguistic models were created in the course of the project, for English, Spanish, and Russian, respectively. Once it has been taught to do so, the neural network generates its own "sacred" text, inventing new words and revealing the universal poetics of the sacred per se.

Another neural network is used to generate speech. The system "takes control" of randomly chosen unprotected devices from the Internet of Things all over the world and uses them to broadcast the resulting text. Loudspeakers suddenly begin to "speak" to people; files containing the text are uploaded onto media servers to be discovered by someone. In this way, the AI-generated "sacred text" is disseminated all over the world as a new gospel, with IoT devices cast in the role of its apostles.

Erik Matrai's installation *Turul* actualizes another mythological image in a contemporary digital form. Turul is a bird from Hungarian mythology, where it appears as a messenger of the gods. Today's version of this mythical creature is the surveillance camera. Just as angels and demons were the messengers of God, today's surveillance cameras perform the function of technological demons of humanity's second nature.

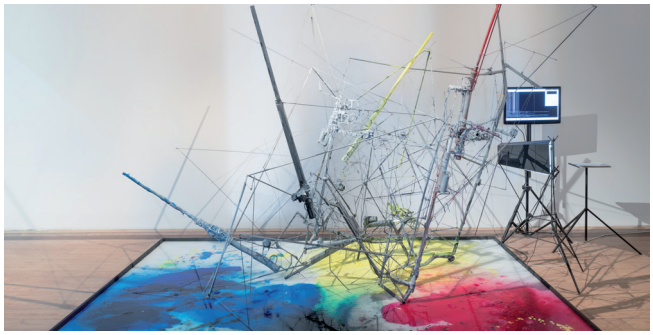
Erik Matrai, *Turul*, 2012, light installation, installation view MMOMA, Moscow, 2018

Technocenosis

The works of Dmitry Kawarga, Jon McCormack, and ::vtol:: could be described as *technocenoses*, or autonomous communities of intellectual systems. Their works, respectively, do not aspire to create alternate realities like those described above, but can actually exist without needing human beings at all.

The concept of technocenosis was first developed by Russian engineer Boris Kudrin 1970–1980 to describe relatively autonomous systems of technical objects. These were to be studied through using biological methods similar to biocenosis. Kudrin's concept proclaims "the constructive equivalence of technical, biological, and informational systems."[12] In the framework of technocenosis, technological systems are likened to biological populations, with families and different species of devices. On the one hand, every device is equal, but on the other hand, it was built according to a blueprint that could be likened to a genetic code of living beings, as seen in biology. Technological systems are subject to their own laws of techno-evolution, governed by a law of informational selection much like Charles Darwin's famous law of natural selection.[13]

Dmitry Kawarga's robotic installation *Absorbing Concepts* transforms philosophical texts into abstract drawings. The viewers can choose a card with a philosophical or literary quotation and offer it to the machine's digital "eye." The machine records what it sees, taking some time to subject the recorded information to analysis and linguistic classification. The algorithm allows for recognition of a full set of lexical, morphological, and image-associative characteristics. Dynamics, density of verbal imagery, poetics, terminological density, and the variety of significant links are structured and transmitted into motor activity on command. As soon as the calculations are complete, the object jumps into motion. The duration and variation of the robot's activity depends on the text's level of saturation with concepts. The more vivid text fragments make the robot shoot brightly colored paint at a canvas, creating an abstract composition as the robot's "artistic interpretation" of the text.

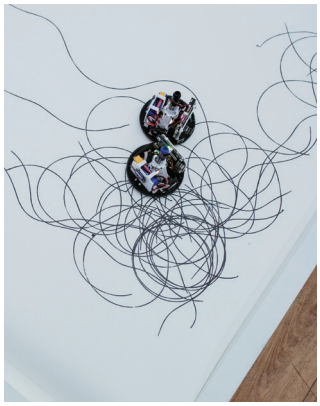


Dmitry Kawarga, *Absorbing Concepts*, 2017, robotic installation, installation view MMOMA, Moscow, 2018

12 Б.И. Кудрин, *Философские вопросы технического знания* (ИФ АН СССР. М.: Наука, 1984) с. 173–174. Translated here from the Russian.

13 Б.И. Кудрин, Исследования технических систем как сообществ изделий – техноценозов. *Системные исследования. Методологические проблемы*. (Ежегодник 1980. М.: Наука, 1981) с. 236–254.

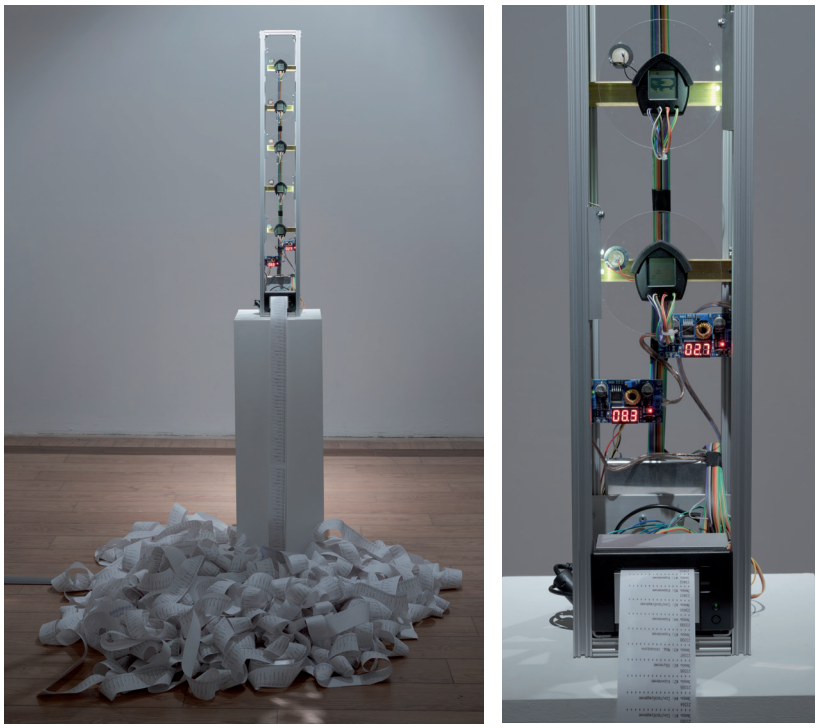
Jon McCormack's robotic installation *Niche Constructing Robot Swarm* consists of several tiny wheeled robots with colored pencils which cannot cross one another's tracks. They build their ecosystem and try to make pacts with each other. They demonstrate the process through which ecological niches are formed, as organisms transform their surroundings in a feedback process with evolutionary forces of natural selection. In McCormack's work, a swarm of robots all interact with each other, trying to develop a common system of ecological niches in which they can further evolve. Each niche is characterized by its own environment, defined by the style and intensity of neighboring lines. The robots learn different behaviors in trying to adapt to the changing environment while each of them attempts to hold onto its own niche. Each robot-generated drawing is unique, created through a combination of interaction with the environment and genetically coded preferences. The artist thus investigates the origin of creativity and the meaning of artworks created by non-human intelligence.



Jon McCormack, *Niche Constructing Robot Swarm*, 2018, robotic installation, installation views MMOMA, Moscow, 2018

The project by ::Vtol:: *Umbilical Digital* is a farm of sorts, where a special algorithm executed on a small microcontroller "cares for" digital animals. These are Japanese Tamagochi toys, one of the first ever virtual pet simulators. Gauging the state of each of the pets under its supervision, the system carries out all necessary manipulations to support their "lives" and their "good moods." Simulating keystrokes, the "cyber mama" behaves like the ideal pet owner, while the pet acts as it would if cared for by a human being. The algorithm observes the behavior of these "organisms," constantly learning and refining its behavior. The statistics of all processes underway materialize on a small thermoprinter in real time,

creating a "chronicle" of the farm in which one can follow the entire evolution history of the colony. Here, human beings now take on the role of an outside observer, while machines take care of machines. Can the actions of a self-teaching algorithm be described with the notion of care? And if so, it may well that the cyber mama considers the Tamagochi not as a pet simulator, but as a digital being of the same species as itself.



::vtol::, *Umbilical Digital*, 2018, cybernetic installation, installation views MMOMA, Moscow, 2018

Auto-evolution

Rosi Braidotti, one of the main theorists of posthumanism, argues that *zoé*, the vital life force, is present both in humans and in other types of matter—including technical objects.[14] Artificial intelligence appears as a posthuman subject, and its mechanical and digital elements are very much alive, communicating with the human world on their own terms. This communication is the main theme of the third group of artworks by Memo Akten, Egor Kraft, and Justine Emard. All three artists address the exploration of the outside world by the robots and their gradual self-complication, based on their attempts to imitate their surroundings.

14 Rosi Braidotti, *The Posthuman* (Cambridge: Polity Press, 2013).

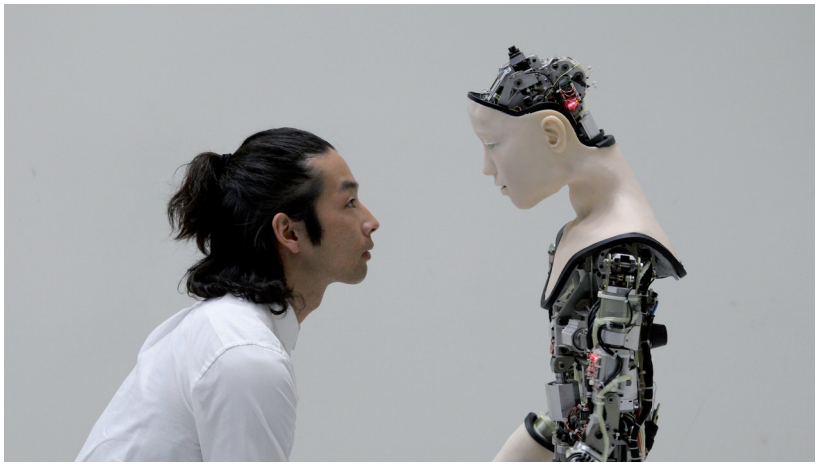
In Memo Akten's interactive video installation *Learning to See: Hello, World!*, an artificial neural network "opens its eyes" for the first time and tries to make sense of what it is seeing. (Incidentally, this is the very first installation that the audience encounters in the exhibition). This network was not trained in advance, but starts from scratch, seeing the world for the first time like an infant, and trying to understand what is going on. In this case, the word "understand" entails the attempt to identify images and to find some kind of regularities between them. The neural network attempts to summarize these findings and to output real-time information in the context of past experience. It also tries to predict the future. The images on the monitor are reconstructions of the world as seen by a neural net; its perspective is constantly updated as its experience accumulates.

Memo Akten, *Learning To See: Hello, World!*, 2018, AI-driven interactive video installation, installation view MMOMA, Moscow, 2018

Egor Kraft similarly explores the capacity of AI to see based on previous experience. His contribution reconstructs and recreates lost ancient Greek and Roman friezes and sculptures through algorithmic analysis of 3D scans of objects from antiquity, thus probing the possibilities of data-driven methods, machine learning, AI, and other forms of automatization. These are transformed to produce quasi-archeological knowledge and cultural-historical interpretations. An algorithm capable of self-teaching is deployed to fill in the lost fragments of friezes and sculptures. It generates models based on image analyses, which are then printed on a 3D printer in various materials and used to fill in the gaps in copies of sculptures and reliefs from antiquity. The AI that attempts to recreate these original forms generates strange mistakes and speculative interpretations of familiar Hellenistic and Roman aesthetics, revealing its own non-human understanding of human antiquity.

Egor Kraft, *Content Aware Studies*, 2018, marble and 3-D printed sculptures (polyamide), machine learning algorithms, video, installation view MMOMA, Moscow, 2018

Justine Emard is similarly concerned with the boundaries between humans and robots in the contemporary world. The work *CO(AI)XISTENCE* is a video installation documenting a performance based on the interaction between digital data and human movement. A robot controlled by a neural network encounters the Japanese dancer Mirai Moriyama. The machine learns from his movements, and attempts to reproduce them. This results in a beautiful duet between the robot and the human dancer, where the language of dance provides unexpected common ground.



Justine Emard, *Co(AI)xistence*, 2017, video installation, 12 min., with Mirai Moriyama & Alter, developed by Ishiguro Lab, Osaka University and Ikegami Lab, Tokyo University), installation view MMOMA, Moscow, 2018

In Lieu of a Conclusion

From birth onward, we find ourselves not only tapped into nature, but also into the world of machines—into the global technocenosis of contemporaneity. Artists reflect upon this situation and create new myths. They speak to society in the language of "techne"—joining technology and science. This creative process also inevitably entails attempts at making sense of different future scenarios in which AI does not just figure as a daily assistant in solving technical tasks, but as an independent entity with its individual experience, life practices, and valuable differences to those found among humans.

What inspires artists is the *gap* between contemporary science and everyday experience, which we can see in phenomena such as the uncertainty in the world of quantum physics, the autonomous entities of AI, the deformations of space and time in relativity theory, or drastic changes to our surrounding through the weakest effects in nonlinear dynamics. This gap gives rise to a sense of irrationality and esotericism. Even if all concepts in science and engineering are ultimately rational and strictly logical, artistic metaphors facilitate the social and critical interpretation of our era's most important intellectual and technological achievements.

Demons are not just a nice figure of speech in a tale of this new world, but a way of describing the psychology of human interaction with autonomous, independent, non-living Others. The birth of strong AI gives us the possibility of interacting with a subjectivity that remains fundamentally other than our own, even if we ourselves have created it. It may well be that this dialogue will help us to understand ourselves more completely.