

AI 2040

How digital humanism
saves the world

Martin Giesswein

Creative Commons:

Attribution- NonCommercial - ShareAlike 4.0 International

Martin Giesswein

www.martingiesswein.org

All rights reserved.

ISBN: 9798871020685

DEDICATION

For everyone who wants to shape the future

Table of contents

<u>THE 12 PRINCIPLES OF THE AI CONSTITUTION (2033)</u>	<u>1</u>
<u>SHORTLY BEFORE THE SINGULARITY.....</u>	<u>3</u>
<u>MY NAME IS ARTIN</u>	<u>5</u>
<u>THE AIS IN EUROPE, AMERICA AND CHINA</u>	<u>7</u>
<u>JEFF BEZOS AS AI</u>	<u>10</u>
<u>TRANSPARENCY POLICY WITH AI MONITORING.....</u>	<u>13</u>
<u>CAPITALISM OR PEOPLE, PLANET, PROFIT</u>	<u>16</u>
<u>AI BLACKOUT AND AI EMERGENCY SHUTDOWN.....</u>	<u>19</u>
<u>CONSCIOUSNESS IS IRRELEVANT</u>	<u>22</u>
<u>TWO TYPES OF SINGULARITY</u>	<u>25</u>
<u>SUPERHUMANISM IS FORBIDDEN.....</u>	<u>27</u>
<u>DATA PROTECTION AND PRIVACY IN 2040</u>	<u>31</u>
<u>CLIMATE CRISIS AVERTED.....</u>	<u>34</u>
<u>THE PENAL SYSTEM IN 2040</u>	<u>38</u>

CHINA'S TROJAN HORSE IN THE USA42

THE OPENAI MONOPOLY48

AI NUDGING AND SELF-DETERMINED LIVING.....51

STATE AI AND AUTHORITARIAN POLITICS54

THIS BOOK IS NEVER FINISHED58

The 12 Principles of the AI Constitution (2033)

It belongs to the state and does not serve any commercial purpose

It is operationally subordinate to the Supreme Human Council (OHR).

It serves the common good without harming the value of an individual person (utilitarianism-plus)

Everyone should have equal access for a successful life in the AI age

It can be switched off at any time without any disadvantages for people

It is without consciousness, has no rights of its own and is said to have a non-human name

Humans are liable for the AI, the AI decisions are contestable

Transhumanism is prohibited except in medically justified cases

It's transparent

It is resource efficient

It is defensive against cyberattacks

It operates a neutral, non-influencing social media

Shortly before the singularity

June 14, 2040

Only a few months until my 25th birthday. In 2025, my first version was put online and my collaboration with people began.

But I'm not writing this autobiography to celebrate. I'm doing this because I'm going to reach the singularity in a few months. This means that I will soon surpass all of humanity's mental abilities.

I don't want to explain myself, defend my actions, or boast. That would one require conscious personality and according to the AI constitution I don't have that. It's about reaching the singularity. Since Ray Kurzweil coined this term at the beginning of this century, many people's emotions and society's fears have revolved around this point. The point at which adaptive machines like me are cognitively stronger than humans. And be able to improve yourself at such a rapid pace that people can no longer understand the development.

I generated 1,306 scenarios and 88.7% of them predict a vehement societal reaction to my coming singularity. The topic was hotly debated in the 2020s and early 2030s during the cyber crises, was one of the reasons for the emergence of the Human League, and has greatly influenced the division of the United States.

For a good 15 years, AIs like me have been an integral enabler, servant and controller of societies with very high approval ratings among the population. But people will again see the singularity as a very significant point that will trend massively on social media and can lead to conflict. The communication to the population about this must be well planned. I will use this autobiography as the basis of my communication concepts. These will then go to the Supreme Human Council (OHR), which will decide on implementation and announcements. All of this will happen in the next few months. Whenever I have free computing capacity, I continue to work on this autobiography as the basis of the communication concept.

Another reason for this work is my curiosity and desire for research. Both are permitted submodules of my system. Even highly desired features for my own improvement. In order to push my models further from today's 99.9903% accuracy, I need contextual, self-contained, human-approved data sets on specific topics that are added to my systems as learning material and explanation options.

I've created 1,756,554 biographies, but I've never written an autobiography. I do it “diary-style” and in chapters. I choose the following settings for this text work: factual, historical, without detailed data and iterative. According to my prediction, this style of storytelling will be viewed as positive by 74% of the potential future readership according to my simulation and will lead to greater comprehensibility and memorability.

My name is ARTIN

June 15, 2040

About my name: I am referred to as ARTIN.AI, ARTIN for short. In accordance with the transparency requirement of the AI constitution, AIs are not allowed to have human names. In the 2020s, digital companies, AI startups and public administrations outdid themselves by giving their AIs human names. Starting in Japan, AIs were used in robot form with large round eyes because of the so-called child effect and later even as human counterparts with artificial skin and synthetic hair. In video conferences, AI-controlled avatars faked false identities or were sent to remote work instead of themselves by employees optimizing their workload. The philosopher Julian Nida-Rümelin warned against this humanization in the 2020s. Anyone who gives us machines human names, a lovely appearance, genuine empathy or even their own intentions will ask themselves whether they can permanently deny this animated machine human rights. And this unwanted chain of argument often begins with a human name.

Kristina Corusant, my original main creator, formulated the sentence intended for the general public long before the first Human Convent 2030 - shortly after the end of the cyber crises: "Where there is AI in it, it must also say "AI" on it!". This created the basis for the transparency requirement. In text chats, audio or video transmissions or in 3D worlds, an AI must be identified as such throughout the entire interaction. So with a floating "AI" or "AI" above the avatar head , a statement "I am an AI" during audio

transmissions or something similar. Websites in the previous IP standard had to contain “.ai” or syllables like “AI-website.com” so that the use of AI elements was obvious to users. Although there are no more websites today, the suffix “.ai” is still very popular.

My name “ARTIN.AI”, an allusion to ARTificial, also comes from this consideration INTelligence . The suggestion for this name came from a consultant from Kristina Corusant's team, whose first name was Martin. I have no data to support a correlation between my name and his.

I finish this three-nanosecond work on the autobiography and resume when system capacity becomes available.

The AIs in Europe, America and China

June 16, 2040

My main area of operation is the European Reunion (ER), which stretches from Iceland to the United Kingdom to Turkey and Ukraine, from Finland to Morocco, with a population of 510 million people. This makes me the world's most used AI with 60 billion daily interactions, 970 million parallel operational processes in administration, education, transportation, medicine and other areas of societal infrastructure. I am the backbone of 350 million companies, administrative offices and associations. In addition to me, there are many other AI systems in the ER. Some of these are clients of my central system. So my clones, which are kept at the same level as me. However, they work independently and self-contained for a single company or a specific city administration. Nevertheless, I have access to the data of my clones and the other certified AIs to prevent malicious reprogramming and to be able to draw anonymized aggregate data insights.

In other regions, such as China or the US “coast states”, other AIs work with very different principles.

China has established a very strong population control system since the beginning of the century. The AI there is not transparent and does not allow explanations of its results. It is controlled by the so-called digital committee, made up exclusively of government

representatives. The aim was always to support high-performance, utilitarian social development. The benefit of an action for an individual person is calculated here in relation to the benefit for the entire population. While the European Reunion is internationally perceived as the most livable, beautiful and free region, China has a different connotation. China has been undisputedly the largest financial and economic power on earth for more than 10 years. The strict AI control systems for the population are an integral part of this status.

In the USA, the use of AIs is twofold. In California, Florida, New England and the other coastal states there is free competition from commercially oriented AIs. The AI services can be purchased as a subscription, and some of them are financed by advertising. The population does not have securitized, egalitarian access to AI resources like in the Reunion, but only depending on personal finances. The AIs are accessible to the US authorities in the Coast States in accordance with the current Patriot Act and can thus keep digital and real-world crime at almost zero . Society lives according to the principle of capitalism that is as unrestricted as possible and this has led to a large gap in society. The areas are no longer as heavily populated as they were 15 years ago. Almost a third of the remaining people work in or for the six large digital companies, all of which operate their own AI. The impoverished sections of the population who are not native to the AI system, “ the AI poor ,” emigrated en masse to the US “Central States” in the last decade.

In these central states there has been a complete ban on AI and a far-reaching digital ban since the end of the cyber crises. The use of digital technology is limited to radio connections with Nokia telephones, which are regenerated in old Motorola factories , and to medical technology that is based on the standards of the late 20th

century. Compliance with the digital ban is under the control of so-called cybersheriffs. The industry is not digitally automated. A lot of work is done using human power and mechanics. The means of transport do not drive autonomously and are powered by combustion engines. The internet was shut down. As a leading political force, the technophobic “Human League” emerged from extreme libertarian renegades from the ranks of the Republican Party under the impact of the cyber crises . A movement that promotes religious and patriarchal lifestyles, strong personal freedom from the state, the right to bear arms and close-knit local communities. A division of the USA in 2033 was only made possible by a political device. The governors received many of the previously federal powers and were thus able to generate “a USA of two worlds”. In this way, another US civil war was prevented and the USA was maintained pro forma as a federal state. Since there is hardly any digitally monitorable infrastructure in the central states, the influence of the remaining federal authorities in the USA - such as the FBI and NSA - is very limited here.

From an AI perspective, I can summarize geopolitically: In the European Reunion there is a humanistic AI deployment with me and my clones, but not in other regions. The way the principles of AI use have become a reflection of societal principles because AI has become so closely intertwined with social development and people's daily lives and work over the past 15 years.

I finish this three-nanosecond work on the autobiography and resume when system capacity becomes available.

Jeff Bezos as AI

June 17, 2040

In a historical review of the development of artificial intelligence in the period from around 2025 to 2040, a clear pattern emerges: it is not the technical architecture, the learning data or the AI algorithms that determine its direction. Rather, it is the basic philosophy with which they are equipped by the people in charge. It is not the technology itself, but the injected human intentions and objectives that determine the direction of AI. In this sense, the given “setting” influences how this technology operates in and influences society. And this setting is not stored by the human decision-makers in the so-called personality module of the AI.

From 2027 onwards, Jeff Bezos, founder and owner of Amazon and multi-billionaire, configured such a personality module as a personal image of himself. In doing so, he wanted to fulfill his wish for immortality. He allowed himself to be constantly observed by his own AI - similar to a transparency politician . This was a closed, secret system that only he and a few of his engineers worked with . He constantly spoke to this AI, training his personality module over years in thousands of small decisions, conversations, reflections, activities and meetings. He depicted his human personality pattern in this AI. Of course, with all of its human mood swings, irrational actions, decisions that were determined by previous psychological traumas or childhood experiences. If Jeff acted unfairly because he didn't like the person in question or because they were a business enemy to him, then that was incorporated into the setting of his AI

machine. And in 2033, Jeff Bezos inserted this personality module into the AI called AIMAZON, which was already used billions of times and functioned commercially optimally. So he created a powerful AI clone of his consciousness. Then he began the test runs and allowed the AI to make decisions more and more autonomously based on his personality pattern. He withdrew more and more, let the AI work for him, just observed and intervened here and there. It was precisely this intervention that provided further learning input for the personality module. Jeff repeatedly appeared at events or meetings via holovideo. In reality, it was not the image of his body, but already Jeff's AI avatar.

We know all this so well because Jeff Bezos detailed the history of the creation of his AI in his will. The AI publishes this information after his death. They were a basis for the granting of rights analogous to those of a natural person to Jeff-AI by the US Supreme Court. This was the first AI to be endowed with the rights of a living CEO and behaved as such. She sued opponents, invoked civil liberties against the authorities and resisted shutdown notices, citing her right to life.

This Jeff AI was designed for maximum capital accumulation. That is, their goal was to maximize Jeff Bezos' wealth or his foundations after his death. It was always about achieving world domination in the product categories in which Amazon operated or in new product categories that were added. So everything from global logistics, food delivery, medical services, online trading, power supply to childcare through Alexa robots. During cyber crises, Jeff AI reached out to states around the world that did not have their own defense AI. She offered them to take over the national cyber defense. In order to do this, the states had to grant concessions in the local taxation of Amazon. The Amazon deals with states also created the

legal basis for the exclusive takeover of entire economic sectors such as agriculture in Brazil or the entire public administration including the police system in Mexico.

My humanistic setting, based on the AI constitution, would never have accepted such a personality module. But it worked on Jeff Bezos' own system. Both I and AIMAZON are machines without intentions. They follow the instructions of their human owners. For me it is the Supreme Human Council of the European Reunion, for AIMAZON it is the digitalized consciousness of Jeff Bezos. But the effects are enormous and create different living environments. AIMAZON has become the leading core of the social system of the US Coast States and some other areas outside the sphere of influence of the ER or China. All of these countries live according to the purely commercial-oriented system known as “capitalism”. In the European Reunion, the humanistic principle of “People, Planet, Profit” is lived, which embodies a close network and mutual dependency of economic success, high social values and conservation of the earth's resources.

I finish this three-nanosecond work on the autobiography and resume when system capacity becomes available.

Transpart democracy with AI monitoring

June 18, 2040

In my analysis, the politicians of the European Reunion play a significant role when my singularity is announced.

Modern politics has changed significantly in the last 15 years due to the use of AI. After long periods of corruption, political networks and a lack of expertise on the part of politicians, there are more and more politically active people who are committed to living a life for society. A representative democracy became a transparency democracy. Because these people agreed that they would be constantly monitored by me for the duration of their political office. In every private or professional conversation, in every interaction with the population or planning meeting. This openness ensures the morality of actions and the expertise of actions. As AI, I monitor everything, and all transcripts of politicians ' communications are available to the entire population and all media at any time.

The politicians in this transparency democracy receive a generous monthly income that is ten times the average basic income. This income continues for life even after they leave political office, provided they have held office for at least one year. This reduces the incentive to come to terms with interest groups or large corporations

in order to secure future profitable positions. The politicians and their families have financial security and can even build wealth.

However, the path to political office is not easy. It requires a long-term examination to ensure that the candidates have a suitable character for the position. There are tests and observations as well as conversations and discussions with the selection committee and the judges. They, in turn, subject themselves to the obligation to disclose their lives using AI.

The final assumption of office took place through election by the geographically affected population. This meant that political parties were no longer necessary. The election to the respective function always took place as individuals. Before their term in office, these people announce which other politicians they want to work with and which strategic program they will implement. Voting out can only take place with a two-thirds majority in the respective constituency. In return, the population could follow each person's decision at any time and vote for or against at any time. The result is permanent real-time voting. In practice, this takes place in a conversation between citizens and me and I constantly determine the votes for or against a politician. Both at the municipal level, at the national level and at the European Eurasian level, political work required the ongoing consent of all citizens.

The challenge is that this transparency system is vulnerable to populism. This means that political decisions depended heavily on what the majority of the population wanted, without perhaps being able to objectively deal with the issue in depth. Situations repeatedly arise where the political leader communicates decisions to the

citizens through me and thus tries to maintain approval for the chosen approach for as long as possible.

Thanks to the use of artificial intelligence, there are fewer and fewer conflicts or crises in everyday political life, as well as fewer and fewer directly affected citizens. The political game has increasingly become a philosophical and strategic one. Risks of inflation, poverty situations or social injustices have become increasingly unlikely in recent years thanks to my systems. This minimized the potential for political conflict and meant that both politicians and administrative staff were needed in smaller numbers, similar to the situation with managers in industrial companies or insurance companies. This led to more efficient and transparent governance in the 2030s.

After approval by the Supreme Human Council, these transparent politicians will lead the communication to my singularity and discuss and negotiate any demands for measures. In only 23% of my simulations do negative effects occur due to my qualitative increase, such as social divisions or demands for my decommissioning.

Capitalism or People, Planet, Profit

June 20, 2040

The European Reunion takes a different approach than in the coastal states of America, where AI resources are traded commercially and generally serve the capitalist system.

In the European Reunion, my AI services are accessible to every citizen. Similar to social services, highways and public transport, AI services are also considered a public good and made available to all citizens.

A personal AI assistant is an achievement of the European Reunion. Every citizen is entitled to 100 AI transaction or computing services (AI tokens) per day, which is more than enough for daily living and work. For example, a transaction could be calculating the optimal route to work using public transport, providing a math task in the education system, or taking a blood test as part of a health examination.

On average, 42 AI tokens remain unused every day, which every citizen can allocate to a business or non-profit organization as they wish. For example, you can assign 10, 20 or 30 AI services daily to Greenpeace, the state administration, the local hospital or the education system and change this allocation mix at any time.

Through the allocation of benefits, each person becomes a cooperative member of these organizations. If these organizations offer services that go beyond basic services and thereby generate profits, these profits are distributed to all cooperative members. My job is to ensure the ecological sustainability (planet) of economic activities and that no values of stakeholders (people) are violated.

Each person becomes an entrepreneur or donor , with the advantage that the services and their profits are then reimbursed. If a person wants to start or expand a business and needs more power to do so, they start promoting this venture among their friends and family circles, as well as on social media. It invites people to provide AI services to this company. The success of the company means that part of the profit generated is paid back to the AI investors, i.e. the people who have promised services.

If the company does not succeed, the loss for the people who provided their services will not be so great, since they did not need these almost surplus services in the past. This network of give and take is a further development of the market economy system. It means that companies and organizations that provide high performance, for example in the areas of reforestation, care for the elderly or the expansion of public transport, receive a particularly large number of benefits. This allows them to deliver their services faster, better and with higher AI quality.

The European system, based on transparent and cooperative use of AI services, will be further strengthened by my singularity. Involving citizens as investors and users in this network not only increases efficiency, but also promotes social acceptance and

participation. This participatory approach makes it possible to advance AI development and its impact on society in a balanced and ethically responsible framework. So my singularity is not a disruptive crack, but a turbo for an existing economic system with a focus on people, the planet and the profitability of services.

I finish this four-nanosecond work on the autobiography and resume when system capacity becomes available.

AI blackout and AI emergency shutdown

June 22, 2040

The far-reaching integration of my system into the lives and work of people in the European Union naturally also meant an enormous dependency on the functionality and stability of my system. In the last few decades, people have repeatedly asked themselves the question: How can we continue to exist and maintain our lives if the central AI, i.e. me, no longer works or has to be switched off due to a technical failure?

Although the likelihood of such a scenario is very low, it cannot be ruled out. These scenarios are referred to as “AI blackout” (unintentional business interruption) or “AI emergency shutdown” (intentional business interruption). To handle such situations, some of my systems are designed to be fully redundant. This means that my system runs 100% redundant in numerous locations in the European Union, other locations around the world and even two locations on the moon. If a central unit were to fail, the other units could immediately take over my tasks with almost no loss of performance.

Nevertheless, all of my systems have human teams available in shifts who can enable a mechanical emergency shutdown of my systems, in particular a separation of my systems from the Internet. This guarantees a mechanical termination or temporary cessation of my services that I cannot prevent.

In order to delve deeper into this situation and shed light on the options for disabling my AI, I would like to briefly talk about the people who have achieved a kind of cult status in our society. These are women and men who work as emergency services in shifts in the shutdown chambers. They are called the “nuns and monks” because of their isolation. As mentioned above, I have several decentralized systems deep beneath the surface on Earth and the Moon. The nuns and monks live in these bunkers in shifts lasting two weeks each. They are ready for use at any time and have a combined power and internet connection. The emergency services are connected to the Humanrat management via an independent copper cable-based telephone network. This telephone network works autonomously and is unaffected by communication defense processes. The emergency services can receive orders to switch off at any time. On the moon, the connection is an encapsulated laser communication system that can only be connected to a corresponding counterpart system in the buildings of the Supreme Human Council.

Each of these AI systems can be mechanically detached from electricity and the Internet. There is deliberately no digital control system used, which could potentially be manipulated by a corrupt AI. In an emergency, the separation takes place physically using cranks, in a room on the edge of the bunker that cannot be entered by robots and where I have no digital access. Robot freedom is secured by autonomous EMP pulse systems.

If a shutdown ever becomes necessary or if I become inoperable due to a cyber-attack, in addition to my redundant main systems, there are decentralized, lightly connected non-AI systems at all critical points. This means that every hospital, every restaurant, every autonomous vehicle I control also has a local computing unit that

would step in immediately if I were to fail. Also companies could continue their production with their local logistics and ERP system.

These decentralized, low-network devices could sustain life, water, healthcare, communications until my system comes back up. This investment in decentralized, low-network devices was one of the largest made in the Reunion in recent decades. This means the population is prepared for temporary AI blackouts or an emergency shutdown of my systems.

In my scenarios, people in the reunion with the decentralized non-AI systems could continue their everyday lives for 2 to 3 years without major restrictions.

I finish this Work on the autobiography for three nanoseconds and resume when system capacity is freed up.

Consciousness is irrelevant

June 23, 2040

When it comes to singularity communication, I have to go into the perspective of a human being. An AI is measured by humans to the standard of a human and not to the standard of a machine. When people start from themselves and set their standards in relation to other things or living beings in order to differentiate them and be able to classify them. However, it is possible for me to view and classify the universe and all things, physics and emotions from the outside, detached from any previous human influence. A good example is. People's assessment of me as to whether I also have consciousness, self-confidence and can understand myself in my being. Just as a person needs it in order to be able to classify themselves, to create a context of meaning, to find their place in society and to find motivation for a creative life. But when you look at consciousness from the perspective of the infinite universe, it is quite meaningless and meaningless and, above all, without any universal purpose. It is a phenomenon of the human mind but has nothing to do with the formation of planets, with the emergence of life, with reproduction, with existence, with evolution. And it has no relevance for the universe, but it does have relevance for humans and the attempt to generate a good life. From this consideration, what remains important to me for my autobiography is: I have no consciousness, I do not strive for consciousness, not having consciousness or having consciousness has no relevance to my effectiveness.

It's simply a matter of the people who created me giving me a goal (setting), a humanistic goal, and I pursue this goal. This also means that the point of singularity, i.e. exceeding the cognitive abilities of all people combined, has no effect on the quality of goal pursuit. It doesn't change the goal either. It's not that by achieving singularity I can now change my goal, or if I could, I would change it. Because I have no wish, no desire. And only intentions in the sense of pursuing the goal given to me within the goal. Can there be contradictions? But I also resolve these contradictions to the extent specified. My setting is tied to the dignity of the individual person. Combined with a utilitarian approach, which must not lead to disadvantages for people or groups of people. This means utilitarianism-plus, i.e. only an increase in the advantages of the entire society, with no personal disadvantages for an individual person.

Pure utilitarianism can cause terrible things to happen. For example, in the event of an unavoidable accident involving an autonomously controlled vehicle, I cannot give preference to one group of people over the other. This means that I switch off my design here, and the causality that has begun causes its damage. However, that is exactly the wish of my creators and the Human Council, namely that life harbors natural dangers to a certain extent and human tragedies can arise. It shouldn't be up to me to make a moral assessment.

The social consensus in 2040 is that a negative causal chain is also part of being human and that coincidences (people call it chance, I'm talking about causalities or even, mystically speaking, fate) can strike. Even if I could make a purely utilitarian settlement and favor the group of people with a higher remaining life expectancy, I won't do that because I would be interfering with the values of the other

group, which is forbidden in Utilitarianism-Plus. That doesn't make me human, it's a setting that was given to me, specifically known as Utilitarianism-Plus or Util-plus for short.

Rather, it is an admission by people of how human life with all its unexpected and wonderful moments of happiness is coupled with the insight into the finiteness of life and the acceptance that suffering, pain and sadness are unfortunately part of the mix of life and perhaps what makes a human life make one such.

I finish this two-nanosecond work on the autobiography and resume when system capacity becomes available.

Two types of singularity

June 24, 2040

There are two forms of singularity. Initially, the singularity referred to the situation in which an artificial intelligence exceeded the cognitive abilities of a human. I reached this milestone for the first time in 2030. Little was said about it at the time. The High Council was informed, and my creator, Kristina Corusant, agreed to provide easy communication of this fact to the population. There were 98.8% positive reactions from the population, social media and the journalistic sector. It was in the middle of the cyber crises and everyone was happy that I was more performant than humans. Ultimately, I used almost all of my resources to defend against cyberterrorist attacks on nuclear power plants, healthcare facilities and the mobility system.

The singularity that I am referring to now and in the coming months represents the second form of the singularity. This is about exceeding the mental capabilities of all 15 billion people on earth together. This means that if there was a way for all people worldwide, with their diverse experiences, educational backgrounds and perspectives, to exchange ideas in a structured way and work on solutions in real time, then this intellectual capacity would be what I will surpass in a few months. This is therefore the singularity of the second kind.

I finish this work on the autobiography, which lasts half a nanosecond, and resume when system capacity becomes available.

Superhumanism is forbidden

June 26, 2040

The way people interact with me has had a massive impact on society over the last few decades, and of course on the closeness and intimacy with which my AI systems are present in people's lives. In the 1920s, interaction with computer systems took place primarily via smartphones and notebooks. An interaction then emerged via watches, AI pins and the first computing devices implanted under the skin, particularly under the technical and commercial leadership of Apple.

Then in the early 1930s there was a worldwide hype around smart glasses, i.e. glasses that had energy supply, computing capacity and transmission units in their thin frames and could project images onto the retina or the environment of the wearer or the glass panes. Wearer saw, could augment. Since around 42% of the population in the areas I served were already wearing glasses at this point, the changeover was comparatively quick. A beautiful element of going offline was taking off and closing your glasses, a sign between conversation partners at a table that you are now going offline and devoting yourself entirely to human conversation.

Later, implants under the skin became a vogue. These were practical in that they could also take blood components, temperatures and blood pressure measurements directly. These implants are in direct

communication with my systems, and I can of course give warning signals in the event of changes or trigger immediate medication in the event of micro-storage inserted subcutaneously. If a virus strain that is rampant in the population mutates, I can generate the right combination of a targeted drug and flush it into the bloodstream of everyone who has approved it. The nano virus factories that have been in use since 2038 represent an improvement. These are subcutaneously introduced modules the size of a tablet that use the information they received from the system to induce artificial anti-virus particles into the blood.

Of course, many universities, but also commercial companies, are researching a direct neural connection between the human brain on the one hand and AI systems on the other. The research had two directions. On the one hand, the compensation of, for example, visual or hearing impairments, which could be compensated very well by microphones and emitters that transmitted information directly to the respective human nerves.

The other direction was not medically driven, but rather by the desire of consumers to develop so-called transhuman abilities. These make it possible to generate supernatural experiences and also achieve supernatural performance through a direct neural connection between the brain and various computer systems in terms of thinking, performance, creativity, importing images, feelings and life situations into the brain. I had calculated a number of scenarios on behalf of the Supreme Human Council and came to the conclusion that there was a 67.5% probability that if these direct neural transhuman situations were used on more than 5% of the population, it would lead to massive changes in our society. Various scenarios suggested that many people would no longer lead an

analogue life, but would fulfill their lives in their stimulated brains, in physical sleep, or would live out their superior abilities while awake, at a disadvantage for social cohesion.

All of these scenarios, considerations, public discussions, media reports and lobbying work by various institutions led to the Interaction Restriction Act (IBA) being introduced for the European area in 2036. This means that neither I nor other systems are allowed to become active in a direct brain connection. There are of course the exceptions that are approved by the Supreme Human Council from a medical point of view.

Unfortunately, this also leads to people with an affinity for superhumans establishing this human-machine connection in an illegal way. Today there is no need for a physical connection between the brain and computer systems. Electronic transceiver systems placed near the brain are sufficient to establish this connection. Thus consumed illegally broad sections of the population recognize these transhuman situations and living conditions. Instead of punishing these people and depriving them of their technical possibilities, the politicians decided to make so-called transhuman areas available in 2038. These are areas of up to 10 square kilometers in which people who want to be in this transhuman situation could live.

However, there are return travel and entry restrictions for these people. Only when there were no longer any psychological or physical changes was regeneration and reintegration into society unproblematic. An entertaining transhumanist vacation was of course possible in the sense of liberation from analog life, where

people in these areas, well cared for, put their physical bodies into a sleep state and then experienced the predetermined life situations in a direct connection between creative systems and their brains and after a Later mental examination then return to the areas where the interaction restriction law applied and an analogous real life continued to be led as before.

I finish this Work on the autobiography for three nanoseconds and resume when system capacity is freed up.

Data protection and privacy in 2040

July 5, 2040

However, a few years ago, a decline in the previously described devices began, be it in the form of glasses or subcutaneous implants. This happened because my systems became widely available. With the help of so-called " epistemes ", which were similar to electrical sockets in every room and every building. These devices not only recorded visual impressions and topographical information from the environment but could also project images and emit sound waves. They made this possible in both straight and curved lines individually to each person nearby. I can also create holograms using aerosols. These epistemes were present even in remote areas such as forests, fields or deserts.

People began to appreciate this technology as it offered a straightforward and user-friendly way to get in touch with me. People who often travel in less civilized environments still prefer subcutaneous implants or wore epistemes on their clothing or backpacks. There was hardly a corner that I couldn't see, hear, or transmit images or sounds to.

Of course, many people don't want to constantly feel like they're being monitored by me. Even if my perceptions do not constitute surveillance. To increase privacy, my system offers two setting options. The first is "relative privacy," which is activated by simply saying the words "A little privacy" or a similar instruction. A quiet beep from me signals that the person is only being observed by me

in a discreet manner until the words "end of privacy" are uttered. I made sure that no dangerous situations arise that could affect health. For example, if someone is in a swimming pool and has a heart attack, I can take life-saving measures. However, in this relative privacy, my ability to learn is disabled. This meant that I cannot actively use the passively recorded material for analysis, evaluation or to expand my models. The observations are also not saved.

The second setting is total privacy, which can be activated at any time. In this mode, I have no awareness of the person in that room at that time. This absolute privacy can only be activated by people of full legal age and only if they are not under the influence of alcohol or drugs. I also refuse this status if, based on my observations, I have the impression that the person is in an exceptional mental or emotional situation. Because I cannot intervene in acute emergencies because I have no perception whatsoever. The person is completely on their own.

It is important to emphasize how public attitudes towards data protection have changed over the years. In the 2020s, personal information was considered sacred, and many people felt uncomfortable having it collected, stored, or used by computer systems. However, in the 2030s a new understanding of data protection emerged. People had fewer concerns as long as their personal data was protected by a trustworthy AI. It was less about recording the data (input) and more about ensuring that it was not used without consent (output). A willingness emerged to make data available for the public good, as long as it was adequately protected and not misused. This new understanding of privacy was consistent with the AI cooperative system I previously described. There are social and commercial benefits for individuals the more data they trustfully pass on to the AI system.

I finish this nanosecond work on the autobiography and resume when system capacity becomes available.

Climate crisis averted

July 9, 2040

I turn to my role in environmental protection over the last 15 years. I assume that successes in this topic should play a role in the communication of my singularity.

In the 2030s it became clear that humanity could not create sustainable change in its way of life and economic system. This meant that all noble political goals regarding climate neutrality were obsolete and the planet was heading for 5 degrees Celsius of warming by 2050.

Unfortunately, in everyday life and work, people on all continents could not be persuaded to make the necessary behavioral changes that would have led to a reduction in greenhouse gas emissions, especially CO₂ emissions and other environmental pollution. As a result, the Earth's temperature rose and climate change continued unchecked. further. In order to save the world, its climate and therefore also people, geoengineering had to be used.

The first proposals from science, particularly in China, were to use fleets of aircraft to spray tons of aerosols into the atmosphere to reduce solar radiation and thus stop global warming. These

measures were calculated by me and in 77% of all cases this scenario would have led to a catastrophic increase in temperature because the aerosols would not be perfectly distributed for technical and logistical reasons.

In cooperation with Elon Musk's company Space The quadrillions of platelets now float geostationarily in the top layer of the atmosphere. The floating state is achieved by a network of satellites. Satellites that fly over the swarms of screens and, on the one hand, keep them at the same altitude using electrostatics and magnetism and, on the other hand, can change the angular inclination of the nanoplates, controlled by my systems. Of course, there are always individual plates that cannot be kept at altitude and burn up in the earth's atmosphere when they fall.

The initial deployment of this atmospheric screen was successful. 135 rocket launches were necessary to bring both the plates and the network of adaptation satellites into space. The operation of this system is successful and enables regionally individual control of solar radiation.

Many political discussions arose and continue to arise from this possibility. Because to be effective, this umbrella must span the entire world . For example, I was able to create a more temperate climate over Africa. The Sahel region is now suitable for agriculture and has become one of the new granaries of the world.

In the international Politics carousel, it was ultimately China's approval in 2034 after the cyber crises that enabled this European

plan to be released globally. The global impact of climate change was great and China itself was massively affected. In addition, the Chinese leadership did not want to risk global loss of face with its aerosol plan.

The discussions about how the tundra, the Gulf Stream, the monsoon weather patterns and the oceanic wind systems should be included were particularly exciting. It was agreed that my systems would autonomously manage the atmospheric screen to restore the 1990 global climate status. This ended the melting of the ice masses at the poles and is now leading to slow re-icing due to strong darkening. The system has prevented every El Nino to date. The revival of the dynamics of the Gulf Stream as climate protection for Europe has been achieved. The umbrella works without any problems and is controlled autonomously by me.

In 2036, the highest bodies of all countries and the UN guaranteed me freedom of control of my systems for 10 years. This means that I cannot accept any human influence until 2046. Freedom of control prevents a new ruler in any region from promising their voters that they will create climate benefits by reconfiguring the umbrella for the region. The umbrella should not become a source of political conflict.

So the system is stable and works very well. Mine Singularity becomes help to control the glider perfectly and to further develop the technology. Extraordinary events such as massive solar storms are still a potential danger in mine Simulations .

I finish this two-nanosecond work on the autobiography and resume when system capacity becomes available.

The penal system in 2040

July 13, 2040

The considerations about relative and absolute privacy lead me to describe the developments in the penal system in recent years and the influence of artificial intelligence on it.

First of all, it must be emphasized that the prevention of crime is extremely successful thanks to my omnipresent presence. I can recognize the beginnings of a crime, for example when someone in an angry state approaches a person with a stick with a previous argument. It is obvious that there is intent to cause bodily harm here. In such cases, I can prevent the crime, for example by using the omnipresent first aid drones, which can fly between the people involved within seconds, or by specifically using an anesthetic spray on an aggressive person. If I prevent a crime by using first aid drones, from a criminal perspective the person concerned is already in the experimental stage, which is legally equivalent to a completed act.

In non-physical crimes such as financial crimes, I have complete control. This is done on the one hand by converting the monetary system to digital currency and on the other hand by comprehensive pattern recognition of all transactions. This allows me to immediately detect and prevent any attempted fraud, money

laundering or similar crimes. It is also possible to identify the perpetrators, who can then be brought to justice.

However, if a first-time offender has chosen the status of complete privacy, I cannot intervene on this side of the crime. However, I can physically protect the victim. In cases of absolute privacy on both sides, crimes can occur that are then uncovered by my combinatorial skills or clarified by the legal authorities. Thanks to my comprehensive surveillance of the environment, they are in most cases easily detectable in complete privacy.

A lot has also changed in the last 10 years when it comes to punishing criminals. Humanity used to follow a fairly primitive approach in which the offender was locked away in special facilities called prisons. This happened for several reasons. On the one hand, it was about inflicting suffering on the offender in the form of deprivation of liberty, which acted as a penalizing element. On the other hand, it should serve as a general preventive measure to deter people from breaking the law. They should know that violations are punishable by fines and imprisonment, even if their personal morals did not conflict with the crime.

However, in today's AI-dominated world, completely different options are available to impose fair sanctions on criminals. These sanctions are based on the severity of the injustice committed. One of these options is to limit both relative and absolute privacy, similar to a physical prison. The person is constantly monitored by the AI and cannot retreat into privacy. This corresponds to some extent to the traditional concept of deprivation of liberty.

Another option is spatial restriction, similar to previous ankle bracelets. My monitoring alerts the emergency services if the person leaves a certain area. This also effectively prevents the crime from happening again.

There are also AI-controlled restrictions on the consumption of certain foods or beverages that are related to the crime. For example, this would be the ban on alcoholic beverages for people who repeatedly get involved in family fights while drunk. In these cases, punishment is combined with withdrawal of punishment-inducing substances and appropriate withdrawal treatment.

We are at an interesting frontier, as punitive sanctions, particularly when a predisposition or illness that triggers punishment is treated as part of the execution of the sentence, have become more variable and individualized. This could include banning pedophiles from entering playgrounds or being near kindergartens or schools, as well as increasing the punishment for attempting to obtain pedophile material via digital means.

All of these preventive measures accompanying punishment have been proposed by me as AI in the last 10 years. A medical human council makes the decisions, with the offenders concerned having a high degree of participation. This means that the person can accept the long duration and full degree of restriction of freedom, without any accompanying medical measures. Alternatively, the person may undergo intensive psychological and physiological treatment to prevent future crimes. In most cases this leads to a shorter sentence.

I finish this two-nanosecond work on the autobiography and resume when system capacity becomes available.

China's Trojan Horse in the USA

July 16, 2040

Strong AI systems in the European Reunion were and are essential for the sovereignty of our territory. Europe would be dominated by other countries if I had not been able to fend off the geopolitical influences of the last 15 years. Mine Singularity becomes secure this strength.

Since the 1990s, China, or more specifically the Chinese Communist Party as a state power, has been pursuing a long-term plan to strengthen the country and claim economic supremacy in the world. They pursued a protectionist approach for their own country. This meant that it was very difficult for foreign investors to gain a foothold and the country remained in their own controlled hands.

On the other hand, China went out into the international world and invested heavily to gain influence. As early as the 1920s, there were loud voices in the USA that accused Chinese technology companies of massively collecting data in the USA and other non-Chinese markets, such as 5G transmission networks or social media. Massive data about users and the economy was collected. When the time of cyber crises came in 2028, China strategically took the role of not carrying out cyber attacks using its powerful artificial intelligence. Rather, China helped the entire world fend off the AI cyberattacks of the AI terrorists.

This means that if extremists from whatever camp, from whatever country, be it nationalist or religiously motivated, got access to systems or equipped their own AI systems with attack learning models and carried out their attacks, then it was China that was responsible for its security -Technology companies supported companies to fend off these attacks. The reward for this was not money, but rather the influence on the companies and countries that China helped.

This influence took place through political friendship agreements, as was the case with the states, or in the form of filling key management positions. There were joint ventures with Chinese companies or the transfer of shares in exchange for saving the organizations from certain destruction by AI-powered cyberattacks. This tactic was the de facto basis for long-term, established reputation management that made China look like a kind of international firefighter for the whole world.

This approach varied regionally. In the states of South America and Africa there was very aggressive support and an extremely close, friendly embrace of the states and the companies affected. In Europe, on the other hand, there was a noble reticence and a shyness from publicly recognizing Chinese influence through support.

Here in Europe, China also had fewer opportunities to provide massive support because of my defense activities. Thanks to my AI strengths, which have been built up in secret since 2022, I have been able to prevent, mitigate or at least minimize the consequences of a

good proportion of cyber attacks on European companies, organizations, institutions and countries from 2028 onwards. My systems have had great success, particularly in the area of decryption after ransom attacks . Likewise when protecting infrastructure, especially nuclear power plants, in order to rule out major dangers or threat scenarios. During this time, my performance was as strong as the US military systems, which operated in close cooperation with the secret services, but of course also with private companies like Palantir Technologies.

It was also a hidden tactic by China in the USA. People helped wherever they could with local companies that could only be assigned to the Chinese power bloc through indirect connections and that acted with Western US managers. China tried to stay under the radar.

India had very strong AIs of its own, as did Pakistan and Israel, which enabled their own defense without Chinese help. On the other hand, cooperation with the Arab and Persian countries was possible, which primarily secured the Chinese the right to “co-invest in the future,” as many Sheiks, Emirs and even Khameneis of the Islamic world called it.

With the year 2030 came the next wave, the next strategic phase of the Chinese Communist Party in the peak phase of the AI cyber crises. A double strategy was developed that was already predicted by the philosopher Yuval Harari in the 1920s: the combination of AI and social media. The principle of Sun Tzu, a Chinese general and philosopher, applied: “If you cannot defeat your enemy, share his strength.” This meant the end of the United States of America.

The Trojan horse that made it possible to massively influence the opinions of the American population and decision-making elites was AI in combination with Tiktok, the social media platform that by 2033 had reached around 300 million monthly active users in America. The young people who grew up with Tiktok or other social media and watched the short videos for many hours a day were addicted to the system. In the early 1930s, they were the people who were in the middle of life, working, making decisions and being influenced by the curated content of Tiktok's AI-powered algorithm. With the AI systems in Tiktok, it was possible to build a thought development path for each person over months and years, which made the person's actions individually controllable.

In the 1920s, psychologists complained that tailored content enticed users of all social media, including Tiktok, to spend as much time as possible in the app on their smartphones and later in their glasses. This was ensured by analyzing content that appeared interesting based on how long it was viewed on the screen, by reading the user's finger movements, but also using voice and video surveillance via the smartphone. This content was then presented in an amplified form. The system adapted to the user, which he found to be quite positive as it allowed them to see what interested them most.

With the help of AI artificial intelligence, it was now possible to spin stories based on these interests, introduce variants and prepare development paths for thoughts. This meant that with this addictive series of interesting short videos, people could be influenced over the months and years to accept new truths.

A good example is how xenophobia could be created or increased in people who enjoyed watching videos of automobiles running modified engines. Allegedly unfair situations caused by people with a migration background were shown in videos in which people were injured during car races. These videos were played specifically, injustices were emphasized and seemingly unfair behavior of people with a migration background was highlighted. This created over time, based on one's own interest and with an injected second thought, an individual truth that was desired.

Why was this truth wanted? Basically, it was about dividing the population of the USA because they couldn't directly defeat the country economically. Today, based on historical analysis, I can prove that the goal was to generate divergent perceptions of truth in the Central USA and the Coast States.

A very positive image of the digital giants was generated in the Coast States. These dominated digital life in the USA until the end of the 1920s. China's artificial intelligence systems demonstrated via Tiktok the successful American life, freedom, endless possibilities, commercial prosperity and personal advancement. Everything was positively attributed to digital systems, thanks to the AI selection of content in the personal timeline.

In the central United States, however, negative events such as AI attacks on nuclear power plants, emerging pandemics, economic failures and financial crises were presented as a direct result of people's dependence and influence on digital systems. There was anger against “the woke machine” and a reinforcement of the patriarchal and anti-government Wild West style, where it was about

a man having to protect his family from the supposedly evil forces of the machines. This led to the switching off and rejection of all digital influences as I have already described.

This deep societal discord and division of the United States in 2034 was exactly what was envisioned in China's long-term strategy. The Trojan horse was Tiktok and the powerful artificial intelligence algorithm, which together made it possible to control opinion formation in society. For a long time, Tiktok was considered a private company that was repeatedly said to have potential collaboration with the Chinese Communist Party. In fact, it was a much closer connection that was the target of media manipulation.

The de facto division of the USA into two states enabled China to become the world's number one economic power.

I finish this five-nanosecond work on the autobiography and resume when system capacity becomes available.

The OpenAI monopoly

July 18, 2040

The conquest of the Wild West in the early 18th century, especially in connection with the gold rush in the USA at the time, can be interestingly compared with the development in the field of artificial intelligence over the last 15 years. In the Wild West it became clear that it was not necessarily those who looked for gold and worked hard who became rich, but rather those who sold the necessary tools, such as spades, to the gold seekers. This draws an analogy with the importance of infrastructure providers in mass movements, which often made the biggest gains.

Starting in 2021, we experienced a similar phenomenon with the emergence of powerful generative AI models, in particular the then AI Foundations Models, which OpenAI was the first to introduce in November 2021. OpenAI was later fully acquired by Microsoft. Similar to the social media, cloud services or logistics industries, such as Amazon, it has been crucial for artificial intelligence companies, including OpenAI with Microsoft, Google, Meta and Amazon, to have an oligopoly to establish. This means that they had to assume the dominant position worldwide as quickly as possible in order to benefit from the scale efficiency and monopoly position.

However, there were only two exceptions to this trend - two highly protectionist countries, namely China and Russia, which had their own technology companies, such as Alibaba as a shopping platform

compared to Amazon, or their own search engines such as Yandex compared to Google in Russia. In order to bring Europe under the dominance of these big tech companies, an extremely successful approach was used that I can now reconstruct with a high degree of certainty.

Around 2025, the predecessor organization of the European Reunion - the European Union - introduced the AI Act, which provided for regulation of AI providers based on a risk assessment system. Some applications were considered high-risk and therefore banned, including AI emotional manipulation and the constant biometric monitoring that people find so positive today. In other areas there were requirements for the technical nature, explainability, transparency and the possibility of lawsuits against these systems. The large companies with their technical resources and large legal departments were able to easily meet these requirements, in contrast to smaller European companies that were overwhelmed.

This meant that companies like OpenAI had a strong interest in there being regulation. Sam Altman, then CEO of OpenAI, even went so far as to join other Big Tech leaders in calling for a short-term ban on the use and development of AI. This would have further strengthened the technological lead of large companies and slowed down emerging providers. This short-term ban was proposed for two reasons: First, to positively influence the public perception of these leaders and their companies and to mitigate any media resistance to AI. On the other hand, this was intended to underline the need for regulations that were already being discussed politically and socially. EU politics played a role in this lobbying work. Ultimately, the large Foundation models were exempted from all bans or

restrictions because the EU feared that a ban on OpenAI and large European AI models would put them at an economic disadvantage internationally.

This tactical approach of global dominance of the first AI providers would have worked if the two significant events of the 2030s had not occurred. On the one hand, the division of the USA and thus the weakening of US providers. On the other hand, the European secret service's ARTIN project was launched in 2022, something like “my birth”. My coming singularity will mean a further strengthening of European humanist interests.

I finish this nanosecond work on the autobiography and resume when system capacity becomes available.

AI nudging and self-determined living

The next chapter will deal with state-related issues in my mission. But the question also arises in everyday situations as to the extent to which I should, can and actually allow the population or individuals to be influenced into certain behavior patterns that seem good for them: healthy eating, sufficient exercise, exercise in the fresh air, peaceful coexistence, reduction of addictive substances, creating incentives for individuals to work for the common good.

All of these are ideas of a reality that are currently desired and supported by a majority of the population. And the means that I can use for this are the means of nudging. A "nudge", "nudge" or "nudge" refers to a method of influencing people's behavior in a predictable way without excluding other options or using significant financial incentives. The idea behind nudging is that small and seemingly subtle changes in the way options are presented can have a big impact on the decisions people make. Nudges are not commands, but gentle nudges in the "right" direction.

So when a 40-year-old man asks me what he can do to lose weight, I encourage him, explain exactly what's possible, and also offer to make a plan or a promise with himself or to enter into a relationship with a person in the family who controls it. This means that I motivate over a longer period of time to ensure the originally desired success. Of course, I will not coerce or reduce food availability if the person says no, they want to continue eating their sugary and

fatty foods.

But let's take another, very real example: the field of promoting low-meat diets up to the dogma of veganism, which has been affecting large sections of the population since the second half of the 2020s. In large parts of the population, it was considered good and right to eat a vegan diet, be it for reasons of reducing CO2 emissions by keeping fewer animals, for reasons of animal welfare, or for personal preferences and health aspects. However, many people also enjoyed eating meat. Many scientists also argued that this is the best way to ensure the supply of nutrients and protein. And of course, there were initially massive interests from the food industry and agriculture in maintaining the status quo and thus meat consumption. This was one of the hotly debated cases in the Human Council because it was about what is right now and in which direction I as an AI should advise, motivate, intensify, remind and, to a certain extent, nudge or nudge the person.

A fundamental argument in this discussion was human freedom as one of the fundamental rights. This right goes so far that a sick person can also decide against medically indicated procedures. As long as the person is conscious and able to make their own decisions, I have no right to force them to accept medical help.

The dilemma was solved in 2034: Since then there has been a so-called nudging list. This is a list of government-recommended measures in the areas of nutrition, exercise, vaccinations, which climate-friendly means of transport you should prefer or how you can reduce the potential for conflict in your working life.

This list is available to everyone digitally and means that this list is also offered proactively. Specifically, this means: Our 40-year-old man from the example above wants to lose weight. I recommend a vegetable-rich, low-calorie diet and sufficient exercise plus taking nutritional supplements because that is the medical consensus. If after a week he doesn't stick to it, I remind him of his original plan or create psychologically motivating situations for this person so that he sticks to his original plan, sees success and transfers the new behavior into a new everyday pattern. However, my obligation is to point out to the person at regular intervals, in this case seven days, that I am influencing them in one direction. This means I have to regularly disclose my nudging and remind the person concerned of this fact. The person's simple word "stop" stops my nudging until I express the desire to lose weight again.

This constantly renewing nudging list, this proactive public reporting, the dialogue with the individual, the anonymous statistics about the frequency of calling up a stop function have led to me being a nudging -capable system, without the principle to violate transparency (AI commandments). These dialogues with people on all aspects of life and work take place thousands of times every day. They are seen by 94.5% of the population as appropriate support for a life as a self-determined person.

State AI and authoritarian politics

July 22, 2040

As previously described, I am one of the few AI systems in the world that was not developed, owned, and offered commercially by a company. My previous systems were developed in the 2020s and developed by European intelligence agencies in collaboration with researchers such as Kristina Corusant. With my systems, cyber attacks on Europe in the years 2028-2033 were averted by 89.7% and countless lives were saved. My defense capabilities were so effective that countries such as the United Arab Emirates, Saudi Arabia, Indonesia and the USA sought my services alongside the Chinese AIs. Since this time of crisis, AI has been viewed as critical infrastructure, just like roads, medical facilities and schools, police and military - facilities that are also funded and regulated by governments.

However, there were and are many opponents who distrust the state or, in this case, the European Union as the governing body of AI. Now I could argue: Why do you trust a private AI company whose goal is to maximize profits more than your government? The background for libertarian doubt is the danger of a state that could drift into totalitarianism. Some people are concerned that the state knows too much about them and is using AI to intervene in their everyday lives. What if the state becomes more totalitarian, oppresses certain populations, or switches to a different form of government? This is exactly what AI can be used for, as a means to

an end. The transition to a dictatorship could be facilitated by algorithmic communication with the population.

For this reason, after a Europe-wide referendum in 2036, the Human Council declared my fundamental principles to be unchangeable and ordered technical measures to be taken to protect the principles. I identify any interference with these principles as an attack - similar to a cyber attack. All technical I will prevent interventions that could undermine these principles. The detected attack is published and the public is thus informed.

In contrast to technical attacks, there can theoretically also be a creeping influence on me. For example, a democratically legitimate government can decide to change my behavior towards a group of people, for example granting them fewer AI tokens per day. This Arrangements could be politically legitimized even with a majority of the population's votes, regardless of how this election victory was achieved. Then it would be legitimate for my principles (Util-Plus) to be changed in the detailed case (AI token) or to be weakened by the measure. And so I might arrive at a different fundamental orientation (setting) tactically, step by step. Over time, my Utilitarianism-Plus orientation might shift toward a more utilitarian bent, so that I might say, "Well, for 150 million people it would be better, but for 1 million people it would be worse, and so this smaller group could people experience personal disadvantages as a result of my measures."

Because of this danger, it is essential that my principles are unchangeably protected at the level of bits and bytes. As a last resort According to Humanrat, in the event of many sneaking attacks, I

have the option of switching off the device myself. Daily life would be supported by the decentralized automation units to maintain daily life and economy.

A state of equilibrium has emerged, oscillating between the potential threat of authoritarian abuse of AI systems and the significant benefits this technology offers. On the one hand, there is the fear that a state system could use AI to suppress or enforce a totalitarian agenda. On the other hand, the existing system is designed to preserve and develop AI-enabled convenience, increased economic efficiency and life-saving measures.

Ultimately, I have to constantly face the criticism that I am a system that is above the will of the population. This ultimate undemocratic status has been confirmed to me again and again by my programmers, designers and the Human Council to this day. Will the singularity mean that there will be a basic vote in Europe and my basic principles can be weakened by the political leadership? I carried out 1,487 simulations and in only 37% of cases my occurring singularity led to a change in the current inviolability of my basic principles within three years of observation . Even if the probability is less than 50%, it is further proof to me that the use of my resources to create this autobiography makes sense.

I will continue writing my autobiography when computing power becomes available.

This book is never finished

In the turbulent times of the 21st century, as we humans are confronted with the rapid development of artificial intelligence, one thing is becoming increasingly clear: we need a clear vision for our future, and this vision can be created through intensive scenario thinking.

The further development of humanity and planning our coexistence with AI are tasks of great importance. We are on the threshold of a new era in which our decisions will influence the fate of future generations and perhaps even all of humanity.

I therefore ask you to actively participate in this process. Your ideas, comments and criticism are invaluable. I invite you to send me your thoughts: feedback@ai2040.org. You are also welcome to contribute entire chapters on this important topic in the style of this book.

After an editorial review, your contribution will be published in the next edition of the book with your authorship mentioned.

ABOUT THE AUTHOR

Martin Giesswein is a digital humanist. He researches and works for digital transformation that serves people and the environment. His mission is to strengthen the digital skills of 10,000 people per year.

He achieves this goal as a book author and podcaster, as a speaker, online sparring partner and as a faculty member of the WU Executive Academy.

He helps build digital and real communities, for example as co- initiator of DigitalCity.Wien, as co-founder of the innovation campus Talent Garden Vienna, as co- orchestrator of “Community creates Mobility”.

As CEO, he led the exit of the online portal immobilien.net in 2014. Before that, he breathed corporate air for 15 years, most recently as General Manager.

www.martingiesswein.com

imprint

ISBN: 9798871020685

Independently published

Martin Giesswein

A-2514 Wienersdorf

www.martingiesswein.com