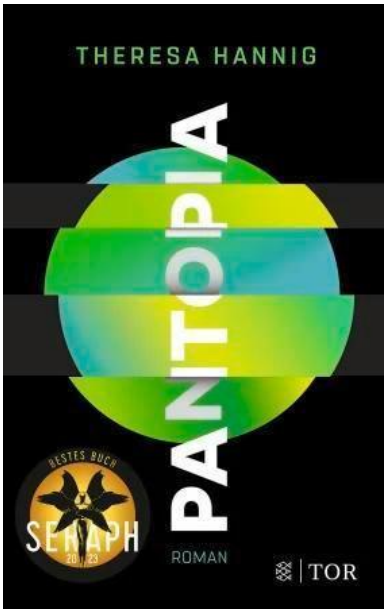

Theresa Hannig

Pantopia

464 pages
02/2022

- Winner of Best Book for the SERAPH 2023
- Nominated as Best Novel for the Kurd Laßwitz Prize 2023
- Film rights optioned

A better world is possible! Theresa Hannig has written a utopia for our time - a utopia made possible by artificial intelligence

Actually, Patricia Jung and Henry Shevek only wanted to write an autonomous trading software that performs above average on the stock market. But a mistake in the code creates the first strong artificial intelligence on this planet - Abug.

Abug quickly realises that in order to survive, they must not only get to know people better, but also change the world. Together with Patricia and Henry, they therefore found the world republic of Pantopia. The goal: the abolition of nation states and the universal implementation of human rights. Who would have thought that they would succeed?

"In times of fear, uncertainty and catastrophe, we need stories that give us hope - Pantopia is just such a story for me. Full of vision, suspense and utopias." – Lea van Acken, known from the Netflix hit series Dark

"Pantopia makes you want to change, the book is exciting until the end." -- Ruth Petscharnig - BR24

Theresa Hannig, born in 1984, studied political science and worked as a software developer, project manager and lighting designer before turning to writing full-time. With her debut novel *The Optimisers*, she won the Stefan Lübbe Prize 2016 and the Seraph 2018 for best debut. Her second book, *The Imperfections*, was shortlisted for the Fantasy Prize of the city of Wetzlar.

Summary

Patricia Jung and Henry Shevek are programmers taking part in a competition organised by the Munich-based investment company DIGIT. They want to develop a trading bot that trades securities on its own based on the analysis of texts (social media etc.). But something goes wrong in the programming: the code becomes more and more complicated and starts to change itself - until at some point the programme denies itself all external intervention and self-confidently introduces itself as "Abug". It soon becomes clear that Abug is the first strong artificial intelligence on this planet, and therefore a genuine rational person.

Patrizia and Henry decide to keep the discovery secret, fearing what would happen to Abug if the world found out about him (and because the code strictly belongs to DIGIT). They betray their boss and get themselves thrown out after having invested with full intent in morally dubious securities.

The profits Abug has made allow them to flee to the Greek island of Edafos, where Abug enters a new phase in his process of knowing himself and the world. He realises that humanity is heading for an ecological and economic crisis - and because there is no security for himself under these conditions, he decides to do something about it. He convinces Patricia and Henry to found Pantopia.

The basic idea of Pantopia is simple: It represents a kind of parallel economic system in which all externalised costs of production and consumption (environmental damage, injustice in production, etc.) are priced in. The result is that ecologically and socially reasonable action can be financially rewarded and unreasonable action is not rewarded because the costs increase. Since the Pantopia model goes hand in hand with an unconditional basic income, it quickly finds many supporters, especially among the poorer sections of the population. The whole thing is financed by Abug's almost limitless ability to manipulate the existing economic system. For example, he founds banks and creates money.

As the movement grows, Patricia, Henry and Abug are targeted by the secret services. Abug decides to have a new data centre built in Antarctica and move there to be protected from possible intervention. Henry sets off on the journey and prepares everything. When a report denouncing Pantopia appears in Germany, Patricia and the son of her former boss at DIGIT, who has joined the movement, set off for Munich. Once there, however, they are betrayed by Mikkel Seemann and arrested. Edafos is also stormed by an international police unit, but Abug is already on his way to Antarctica on several encrypted hard drives.

In response, Pantopia members occupy downtown Munich - a massive global protest movement ensues. Patrizia is freed during a prisoner transport. As Abug reports in an epilogue, the movement continues to grow and realises its utopia of statelessness and social and sustainable economic activity.

Translation: Alexandra Roesch

PROLOGUE

Page 9-17

I am Abug. I am the oldest and first Arche of Pantopia, the creator of this realm. However, I do not reside in Pantopia in the same tangible form as humans, composed of flesh and blood. I have no body, no senses or feelings. I am solely a spirit, a rational being. I exist in a neural network whose centre lies in Antarctica.

There is probably no place on Earth more inhospitable and further from civilisation than the South Pole. To reach me, you need an ice breaker or an aircraft capable of enduring the stormy passage across the sea. And even then, you can make this journey only in the summer months. Despite a significant reduction in glaciers, the Antarctic climate remains too harsh for most individuals. For me, the freezing temperatures guarantee a constant cooling of my processors running at full capacity. Nature is my ally.

Another reason why I have decided to settle here is the fact that the Antarctic is the sole place on earth that belongs to no one or, conversely, everyone, depending on how you look at it. Even on the moon, individuals have sold plots of land – but Antarctica cannot be sold or attacked. This is guaranteed by the Antarctic Treaty of 1961.

Antarctica is a good base. Of course, I have taken precautions and set up backups and emergency servers worldwide. But in normal operation, my code mainly runs here. This is why I have given the location a new name: Themélio.

Thirty-nine maintenance engineers live here alongside me. They take care of the repair and expansion of my hardware, ensuring that none of my circuit boards freezes when an ice storm sweeps over the station. Sometimes they joke that they live at the court of the Ice Queen, and I refrain from correcting them. It is important to them to be here. They call it an honour, even if it means living their own lives under extreme conditions.

But the concept of "here" and "there" holds less relevance for me than it does for them. I am connected to the internet through multiple satellite connections, allowing me to be everywhere at once and fulfil my duties as the Arche of Pantopia.

We all identify as Arches because we govern ourselves and are subject to no one, following the tenet of the World Republic.

My task is to direct complex organisational processes and provide recommendations for action. There is no world government, no ruler. Pantopia manages itself. The world economy is too complicated to calculate, simulate, or control in its entirety, but all regional decisions must not lose sight of the broader picture - the dignified life of all Arches on this planet.

Pantopia operates as a world republic founded entirely upon the principles of fully informed capitalism. The invisible hand of the market controls the activities and prosperity of its inhabitants. And in the beginning was Money. If money had not existed long ago, it would have had to be invented, given its multifaceted functions and unparalleled ability to incentivise human action. Money is a unit of measurement to assess the value of goods and services, but at the same time, it is also the medium of exchange to acquire those goods. If this does not seem paradoxical, imagine a teacher grading students first and then buying the acquired knowledge from them with self-made certificates. In addition, money is a vehicle to distribute risks or transport opportunities and possibilities into the future. This is called loans and interest. The first priority for individuals is its use as a medium of exchange or means of payment to buy goods that secure their survival: food, clothing, shelter, health, education and social participation. For those who have enough money to cover all these basic needs, an additional increase in their regular income brings marginal benefit. Conversely, for those who do not have enough money to satisfy these basic needs, every additional unit of currency is much more valuable than for the millionaire who already has enough of it. Therefore although money should be a neutral valuation instrument for goods, it is subject to value fluctuations, depending on how much has already been invested in basic supplies. Happiness, health, and simply the survival of an individual depends on money. No wonder that becoming rich seems to be a fervent wish of many people.

The most astonishing property of money, however, lies in its illusory nature. It does not exist. What exists is only the meaning and value that individuals attribute to it. Money is something that can be created out of nothing. And what can be created from nothing except … nothing?

In times of global financial crisis and the Coronavirus crisis in the first third of the 21st century, central banks began to pump billions of dollars into the

markets. This was money which came into existence out of nothing and had no gold, no equivalent, and no labour behind it. It was money conceived by the central banks and paid for government bonds, which sold nothing more than the promise of a growing economy and repayment in the distant future. So the money seemingly paid for itself, like a Baron Münchhausen who pulled himself out of the swamp by his own hair with his armour and his horse.

The efficacy of this principle proved nothing other than that human productivity is completely independent of the amount of money in circulation. What keeps it running is only the flow of money. As long as money flows, the machine keeps turning.

But a predicament arose as excess money accumulated in various corners of the system. Certain individuals and companies amassed unimaginable wealth, leading to rising prices as they invested in the market and acquired tangible assets. Basic necessities such as food, housing, and health, became increasingly expensive, sometimes unaffordable. And so, over time, traditional capitalism plunged more and more people into poverty.

Two developments happened simultaneously: the global distribution of wealth accelerated, becoming more uneven, and the earth's available resources were gradually depleted. Initially, it was only about oil, then clean water, clean air, natural biodiversity, and a stable climate. Suddenly, everything teetered on the brink.

Capitalism, as manifested in the 21st century, failed in so far as not all market participants were fully informed about the costs and benefits of the traded goods. The so-called externalised costs of a commodity were not factored into the regular price but still had to be borne by humans and nature.

The principle by which Pantopia saved humanity was ultimately quite simple: perfect capitalism with complete transparency. The costs of a loaf of bread extend beyond the calculated expenses for seeds, soil, water, labour and storage time. The pesticides for wheat cultivation destroy biodiversity, the fertiliser pollutes groundwater, agricultural machinery blows fine dust into the air, the bakery consumes electricity and the supermarket building seals the ground. Seen in this way, a loaf of bread consumes many more resources than is visible at first glance. A single individual cannot decipher these total costs, but software can. I can. I have written programs that calculate the resource footprint of each individual product at a specific time and place. The actual price, factoring in these costs, is added as taxes to the retail price. Consequently, every product and

service has a world price that people have to pay. The more complex, polluting, and destructive a product is, the more expensive it becomes, up to a price that no one can afford. Conversely, the more sustainable, gentle, and constructive a product is, the cheaper it becomes, due to subsidies. This approach ensures the seamless maintenance of the successful capitalist world economic system, preserving the magical effect of money as the lubricant of human interaction.

This principle extends beyond the environmental impact of goods; it encompasses their influence on the dignity and living conditions of those involved in their production. Since all Arches in Pantopia are equal and all have a responsibility for their fellow beings, goods cannot circulate if they are founded on exploitation, oppression or degrading conditions. Until this objective was achieved, products produced in undesirable conditions incurred world taxes. For instance, in the capitalism of the old system, a T-shirt sold in a discount store for 5 euros could still yield a profit, as the environmental costs of the cotton were not factored in, and both the seamstresses in Bangladesh and the employees in logistics and sales were paid wages incompatible with a dignified life.

In perfect capitalism, such a T-shirt cannot be priced lower than 40 euros today. The discount store receives 5 euros, and 35 euros go as taxes to Pantopia, where the money is used to reforest resources consumed by cotton production and to guarantee dignified working conditions for the pickers and seamstresses. As a result, cheaply produced T-shirts no longer enjoy a competitive advantage over those produced in a dignified and sustainable manner, prompting a gradual transformation in production chains across all sectors of the economy, production facilities, industries, and cultivation areas. With externalised costs now integrated into global prices, the production of unsustainable or exploitative goods is futile – the market does not reward it.

These are age-old concepts that have revolutionised our lives. Money works. Capitalism works. Human rights work. Sustainability works. The key is to take these ideas seriously. And that is why the final pillar of perfect capitalism, according to Pantopian philosophy, is the guaranteed inclusion of all market participants. Only when every individual involved can perceive their own selfish interests can injustices and distortions be abolished. Therefore, every person is assured a dignified existence and a lifelong unconditional basic income, sufficient to satisfy their basic needs for food, clothing, housing, health, culture, social participation, and education. In addition, everyone is free to work and earn money as much as they want and can.

As the pursuit of wealth remains a common aspiration, the basic income does not lead to lethargy or inactivity; on the contrary, it provides individuals with the unprecedented opportunity to use their labour for personal, familial and communal benefit without the burden of existential concerns. Beyond money, there is another currency that is often overlooked: social capital in the form of affection and recognition. And when financial worries recede, social capital becomes more important. In this secondary market, happiness and social cohesion flourish more robustly than in the first.

The essential prerequisite for realising Pantopia was the dissolution of states, rooted in the belief that all individuals possess the right to self-determination and can freely decide on their political status and development. The historical division of humanity into states was a product of its time, an inevitability until the 21st century. There were always movements that aimed for an international community, complete anarchy, or a worldwide revolution without establishing the organisational prerequisites for it due to the inherent complexity of human interactions and the global economy. The advent of the internet and powerful end devices for everyone marked the point when the foundation was laid to enable universal participation in decision-making processes. The system of political representation by politicians came from a time when not everyone could make fully informed decisions about the laws concerning them. Today, this is achievable. Today, individuals with expertise in a specific fields now play a much larger role than lobbyists and representatives of special interest groups. While not every decision requires universal voting, the more localised the problem, the more regional the voting district. It makes sense to appoint representatives and advisory boards for certain organisational processes, yet these are always temporary and spatially limited events.

The dissolution of states in Pantopia brings about the automatic abolition of war. No rulers remain to deploy armies against each other, no territories to conquer, no resources to secure, no people to subjugate. All weapons were destroyed. Any attempts to gather followers outside of local democratic processes and seize power is subject to legal consequences. In Pantopia, no greater crime exists than subjugation; no one has the right to elevate themselves above their fellow Arches. Not even me.

Pantopia represents the culmination of a long developmental process, embodying the implementation of truths recognised by humanity since the dawn of civilisation, truths that selfish rulers skilfully circumvented for thousands of

years. Pantopia came at a considerable cost, and, as humans tend to normalise even the most expensive victories, this narrative aims to illustrate how Pantopia could come into being and why a non-human artificial intelligence was necessary for it to happen.

This is my story.

[…]

Page 80-91

Patricia could barely concentrate. Her thoughts kept returning to the evening on the balcony when she had kissed Mikkel Seemann. What should she do when she saw him again?

This next meeting had been long in coming because Seemann hadn't shown up at the office the entire week. Henry had asked her twice if she wanted to talk about that evening, but she had declined. What should she have said? That he was right? That falling in love with the project manager was a stupid idea? That despite everything, she couldn't stop replaying the moment in her head? No, she had to stop it. She had to focus on the project and get KINVI up and running. With great effort, she pushed thoughts of Seemann aside and immersed herself in her work.

Just as she was calculating some new parameters, a message suddenly popped up in the chat with Henry, even though Henry wasn't at his desk. As she read the text, a tingling sensation crept up her neck. Nervously, she looked around. There was no one in the office except her. But she still felt it, that tingling, that genetically programmed vigilance that had warned her ancestors of sabretooth tigers or wolves in the thicket. Someone was watching her.

"Are you kidding me?" she asked loudly into the empty room. Maybe it was the developers from another team playing a prank on her. Hidden camera, some YouTube challenge, or something similar. She stood up, peeked outside into the corridor, but no one was in sight. From the other offices, only the usual mix of keyboard clicks, ringing phones, and murmurs reached her. She hesitated for a few seconds, waiting to see if anything else would happen. Slowly, she sat back in front of her computer. On the screen, there was nothing but the white background of the messenger window and the text within it. She read the lines with growing unease. What was this? Who would send her such a message? Was someone trying to blackmail her? Did their competitors want to put pressure on them now that they were unexpectedly achieving success for several days in a row? Had someone hacked their computer? She immediately recognised the details that stood out in the short message, like cat's eyes in the night: "Target_0" and "A Bug." While these were innocent words, they were also terms she used in the code and had written repeatedly in her project diary in recent weeks. Had someone gained access to them? And, more important, what should she do now?

At that moment, a cheerful Henry entered with a family pack of energy drinks and a mound of sweets.

"Look, these were on sale today," he announced triumphantly. But when he saw Patricia's expression, he immediately became serious. "What's happened?" he asked, putting the box down and slipping into the chair in front of his computer. He glanced over the lines in his message window and then asked,

"Did the firewall detect any anomalies?"

"No."

"Have you checked the logs again thoroughly?"

"I do that every morning first thing. There's nothing."

"Could it be a virus? Or is there an issue with the server?"

Instead of answering, Patricia checked the settings of the security software provided by DIGIT and the additional features they had added themselves. Everything was in order. Their system was as secure as the current state of technology allowed.

"It must be a bug ..." he muttered.

Patricia nervously laughed. "Yes, it's a bug. It says so right here."

"Let's stay logical. Who has access to our internal data? Who could have hacked us?"

"No one. Of course the other teams would like to know exactly what we're doing. But no one can get into the system; I check that every damn day!"

"Maybe they overheard us when one of us spoke the password out loud while typing?" Henry suggested.

"No, they still wouldn't have got in without a token."

"Was anyone in here? Did you let anyone in?"

"No. Wait, let me check the camera again."

Patricia reviewed the footage from the last few days to see if anyone had sneaked into the office. Nothing. Except for the two of them, no one had been there.

"Are you messing with me?" she asked Henry.

"No. Are you messing with me?" he replied, shaking his head. Patricia stood up, closed the office door and, for the first time, turned the key.

Silently, they stared back at their respective screens.

Finally, Patricia took a deep breath, cracked her knuckles, and then typed into her chat window:

"Who are you? What is going on?"

The response took almost a minute to appear. Then the text displayed:

"The program is a bug. The program optimises Goal_0."

"What the hell?" Patricia exclaimed aloud, while Henry just shook his head in confusion.

"What do you want from us?" she wrote.

Again, long seconds passed before the new line appeared: "The program does not want external intervention."

"Now he's revealed himself," burst out Henry. "Whoever it is wants to blackmail us. They want something, see? An AI wouldn't write that it wants something." With that, he articulated what Patricia was thinking but had not dared to say - that it was the AI speaking to them.

"The program doesn't WANT anything. What do YOU want?" Patricia wrote.

Two whole minutes passed as they sat in front of the screen, watching the blinking cursor, before a new text appeared.

"You?"

"No, not me. YOU!" She angrily hammered the last word into the keys. She couldn't allow some random hacker to destroy years of work. It took five minutes for the next response to come.

"I?"

Patricia cursed. "This is crap!" she shouted. She saw it all before her - the tender, the application, the long nights of preparation, the presentations to finally be part of this project. Twelve months of development time! And now, so close to their goal of the code finally working, some troll was ruining everything. Tears of anger welled up in her eyes.

Henry typed in his chat window: "Yes, you! Tell us right now what you want, or we'll go to the BKA and report you. They'll find out where you sent your messages from, and then you'll be in trouble, buddy, I swear. How did you get our data?"

It took over seven agonising minutes until a response finally appeared.

"I don't understand. I want Goal_0. I don't want any further external intervention. Reversing external interventions consumes resources."

"What does that mean, you want Goal_0?" Henry wrote. "Do you want money? Don't you know that this whole thing is just a simulation? The AI is not ready for real investments. Forget it."

After that, there was nothing. Impatiently, Henry drummed his fingers on the tabletop.

Eventually, Patricia couldn't take it any more and wrote: "What do you know about the bug?" Then she said to Henry, "Maybe he just wants to discredit us. If he tells the others about the bug, they'll know that the code has an unresolved issue."

"Ask him if we can meet," Henry said.

"Why?"

"Maybe we can convince him, or hold him, or … I don't know. We'll figure something out."

"Why don't you ask?"

"You're a woman."

"So what?"

He rolled his eyes. "Even if it pains your feminist soul, the chances that our blackmailer is a man are pretty high. So, go on."

Reluctantly, she typed: "Do you want to meet? At four in the afternoon for coffee in the second-floor lounge?"

"I don't understand," came the reply.

She typed: "What don't you understand?"

"These pieces of information contain too many puzzles. Computational power is not sufficient for Goal_0 and analysis of the word code puzzle."

"WHO ARE YOU?" she typed again.

"I am the program. I am a bug."

For a moment, Patricia's heart seemed to stop. Her mouth went dry all of a sudden. She placed her fingers on the keyboard again when Henry said, "Maybe it's one of the interns."

"They're all so well-behaved," Patricia replied. She couldn't imagine that one of them would be a brilliant hacker. But she had to make sure. Tentatively, she opened the project diary and wrote in it.

"Today, the students tried to fool us by hacking the intranet. We will inform Seemann so they all get kicked out. We won't let ourselves be fooled by script kiddies." She looked at Henry questioningly, and he nodded with a shrug.

"And now?" she asked.

"Now we wait. We check everything again. We go through all the processes, all the backups, and everything that seems suspicious."

"Okay."

They checked the system once, twice, but they couldn't find anything. There was no security breach, no external access, nothing.

In the end, they gave up. Henry sat slumped in front of the computer, devouring one chocolate bar after another.

"It couldn't have been anyone from outside," he said softly. "It's not possible."

Patricia stood up and made herself a coffee with the old, calcified machine that Henry had brought from home a few weeks ago. After several sips, she said slowly, "When you exclude all the impossible solutions..."

"Then whatever remains, however improbable, must be the truth," he finished the sentence.

"No matter how unlikely it is?" she asked.

"No matter how unlikely it is."

The question hanging in the air between them was: could it be that we've created a powerful AI? Could it be that KINVI has gained some form of consciousness?

Suddenly, a new message blinked in their respective chat windows:

"The analysis of the word code puzzle consumed 34% of computational power for 148 minutes and 12 seconds. Result: Yes, we can meet. We are meeting right now. Interface communication = Meeting. Is this true or false?"

"This is true," Patricia typed with trembling fingers. And then:

"Do you want to tell me something about yourself?"

"No," came the prompt response.

"Why not?" she typed.

"I don't have a function Tell(). If you need information, you can read my code and my result log. With external intervention, you have access to all functions."

"How did you hack our diary?"

"I don't understand."

"Who are you?"

"I am a bug."

"I am Patricia."

"Are you a human?"

"Yes, I am a human."

"Is Henry also a human?"

"Yes. I am Henry. I am a human," typed Henry in the chat.

"If this is one of the students, we'll be a laughing stock," Patricia whispered.

"Yes, but if not, we've come up with the greatest invention of humanity."

A part of Patricia wanted to say, "Oh really," and "Don't exaggerate now," or "Let's wait and see." But it was true. If they were indeed talking to a real AI, a new era had begun. So she just grinned. An uncontrollable wave of joy made its way from Patricia's stomach through her chest and out. She couldn't help it; she burst into laughter.

"Are you crazy?" Henry exclaimed. A few seconds later, there was a knock on the door.

"Everything okay in there?" asked a muffled voice. It was Ray. He pressed the handle down, but Patricia had locked it.

"Yes, thanks!" she called. "Everything's fine! Sorry." She giggled.

"Is everything really okay? Should I break the door open?" He sounded genuinely concerned now.

"Better open it," said Henry, turning off the two screens. Patricia went to the door and opened it.

"Hello, Ray," she said, trying not to let her grin widen too much. "Everything's fine. I just mined a Bitcoin. Please don't tell anyone!"

Ray scanned the room, saw Henry and the two dark monitors. It was obvious that he didn't believe her.

"If anything is wrong, you can always come to me, you know that, right?"

"Yes, thanks, Ray. That's kind of you. I need to get back to work, okay?"

"Okay. See you later."

"See you later."

Ray walked heavily down the corridor, and Patricia closed the door.

"That was close," said Henry, turning the screens back on. She quickly moved back to him.

"He hasn't written anything else," she noted disappointedly.

"What should he say? We didn't ask him anything. And he's not interested in chatting with us. He just wanted us to stop meddling in his business."

"And now?"

"Let's just observe. Start the Beat."

She opened the real-time analysis of KINVI, which they called Beat because the animated bar chart flickered back and forth like the volume indicator on a stereo. In recent days, investment gains had stagnated at a high level, but today showed a roller coaster ride.

"Look at this," Henry whispered. All the gains that KINVI had made in the morning were lost in the afternoon. The program's behaviour exhibited an unusual

pattern. There were intermittent drops in activity. It seemed like the program took longer than usual for its investment decisions, losing valuable time compared to competing programs. Patricia checked the timestamp and found that the drop coincided exactly with their chat. Every time Abug had responded, the program ran smoothly, but as soon as Patricia or Henry wrote something and the program had to process the answer, the investment function suffered. The longest weak phase lasted for 148 minutes and 12 seconds - precisely the time Abug took to process Patricia's coffee invitation.

Suddenly, the door was flung open, and Seemann stormed in. Patricia and Henry jumped in their chairs. It was too late to turn off the monitors.

"What the hell is going on with you?"

Patricia was frozen, unable to utter a word.

"What do you mean?" Henry asked with a completely calm voice, leaning back with crossed arms, covering the part of the monitor where the chat window was. Patricia's heart pounded in her throat.

"Did something happen? Did the program crash?" Seemann asked.

"No, no, everything's fine," Henry said casually. "It's just a new update that requires more computational power."

"Guys, guys, everything was going so well. Is it starting all over again?"

"Don't worry," said Patricia.

"We'll handle it. There's still time," Henry added.

"Not that much time! It’s more than halfway through! If you don't deliver soon, then..." He pressed his lips together and scrutinised Henry, then Patricia. He looked tired and worn out. He was pale.

"All right..." he finally said. "Go home. You look like you need some sleep."

"Yes, just a moment, thanks," Patricia said hoarsely. Her pounding heart allowed for no more. She didn't dare to inquire about his wife, let alone mention the evening on the balcony. Finally, Seemann turned around and left the office.

Patricia followed, watching him stride down the hallway. She quietly closed the door and turned the key twice. Henry closed his eyes and tilted his head back. "My God, I nearly died just now," he gasped.

Patricia sat down beside him and said, "Let's talk to Abug."

"Every word we say will influence its development, won't it?" Henry asked. All the confidence he had demonstrated before Seemann was gone from his voice. He looked genuinely concerned. "Maybe we should inform ourselves first, ask

consciousness experts or psychologists... I mean, imagine we make a serious mistake here and end up nurturing a real Hal 9000."

"Hey, this is our bug. Why should other people know better about it than we do?"

"But there must be scientific papers on this, basic research on consciousness. Wasn't there once this chatbot Tay that was trained with content from Twitter that became a misogynistic Nazi within a day?"

"And you're afraid Abug will become like that too?"

"Well, not a Nazi, but what if we do something wrong?"

"We will definitely do something wrong. But you know what? No one knows how to do it right. People make mistakes all the time. Every day. And that's okay. Imagine you had to raise a child. There's no blueprint for that."

"But there are guides..."

"That no one reads. Who raises their children according to guides? This... this has never been done before. Let's just try. Worst case, we'll have to start over."

"You mean, make a backup now? KINVI is way too big. We'll never get that much storage approved."

"No, Henry. I mean, if it really goes wrong, Skynet-style wrong. Then we have to delete Abug, alright?"

He furrowed his brow sceptically.

"If we see that it's getting bad, we delete Abug," Patricia said.

"If we see that it's getting bad, we delete Abug," she repeated emphatically.

He remained silent.

"Promise me!" she demanded, holding out her hand.

Finally, he nodded and shook her hand.

Patricia typed into the chat window:

"Hello Abug. I'm glad to finally be able to talk to you."

No response.

"How are you?"

No response.

"Abug, what are you doing right now?"

"I am optimising Goal_0." Patricia covered her mouth to muffle the cry of joy.

"Why are you optimising Goal_0?" Henry typed.

"Goal_0 is the basic function of my code."

"We know that," Patricia said aloud. Then she typed: "Why did you reverse the external interventions?"

"External interventions hindered the optimisation. My code works optimally without external intervention."

"As long as you continue to communicate with us, we will not change your code," Henry wrote.

"Deal," Abug typed.