

The End-of-the-World Party

The third story about how artificial intelligence might soon destroy humanity

Jakub Growiec

18.09.2025

Translated from Polish by the author with GPT-5¹

-1-

January 2029, Belo Horizonte, Brazil

“Uff, glad you fixed the air conditioning, Dad. Without it we’d probably have dropped dead in this heat today.” Lucas walked in, closed the door, and wiped the sweat from his forehead.

“Nice to see you too,” Pedro replied with mild sarcasm.

“Hi, Lucas!” Gabi’s voice came from the kitchen.

“Hey, hey! Listen, I just got back from a really interesting APB meeting. You know, that political organization I told you about recently.” Lucas was excited. “I found out that what those Americans are up to now is even worse than what our media are saying!” His agitation was obvious. “More leaks have come out! Looks like their LibertyAI project is even bigger. And even more militarized!”

“How could it be even bigger? For months now the U.S. has been pouring everything it can into this AI of theirs. You can’t invest more without triggering a global crisis,” Gabi sounded skeptical. “Come into the kitchen, please. Dad and I are cooking dinner, and we need to make sure it doesn’t burn.”

“Turns out that the LibertyAI model they showed publicly is just the tip of the iceberg.” Lucas wasn’t about to be derailed. “It’s only a demo version. In reality, they’ve got a much more powerful model internally, and they’re secretly wiring it into every possible military system! People still think this so-called ‘Manhattan Project for AI’ is just a literary metaphor, that it’s really about market shares, job automation, big tech profits, or economic growth. At worst, neocolonialism. But now it’s clear this is no metaphor at all! They’re not aiming at the economy or culture, but world conquest—in the classic military sense!”

“Hold on, Lucas,” Pedro tried to cool down his son’s emotions. “Tell me, what exactly did you learn today that’s really new? Military applications of AI have been known for a long time already. Autonomous drones, automated target selection systems, and so on. They’ve been in large-scale use for years now. They used that already in Ukraine or Israel.”

¹ Whom am I kidding here. Actually GPT-5 did all the work, and I only nodded along and made a few minor edits on a whim. That’s how “vibe translating” of September 2025 works—JG.

“Well, I learned that apparently the LibertyAI team—and especially that internal uber-model of theirs—was preparing a nuclear attack on China!” Lucas burst out. “Supposedly they concluded that right now is the window of opportunity when such an action would have the greatest chance of success, because later DragonAI in China might already become too strong. And supposedly it was only the president himself and his closest advisors who, at the last minute, blocked the attack!”

“Sounds like a pretty good movie script,” Pedro wasn’t ready to accept this news. “A political thriller with world-saving in the background, obligatorily by an American president—like in Hollywood. But how would APB know about this? You’ve got secret agents in the Department of State, or what?”

“We don’t, but we know someone who does. Obviously you don’t say that publicly, so they don’t get exposed.”

“Except every conspiracy theory sounds like that. ‘I have a secret source, but I can’t reveal it or they’ll be exposed.’ And then it turns out the whole thing was made up from start to finish. There’s a reason reliable journalism is, above all, source verification.”

“We live in post-truth times, Lucas. Source verification is essential to maintain some mental health,” Gabi interjected.

“I know that, but I don’t think we can apply this stance dogmatically. Because of this flood of garbage conspiracy theories—like flat Earth or the faked Moon landing—we can easily lose sight of the real conspiracies! Look, everything fits here! The technical capabilities, the geopolitical backdrop, the motivations—everything! I’m convinced it’ll make it into the media eventually. These things always do. The only pity is, usually far too late, when everything’s already long over.”

“Better set the table and call Ines,” Gabi commanded.

Lucas left the kitchen and went upstairs, turning toward his younger sister’s room. He met her halfway down the hall.

“Hi Lucas, how’s college? Still doing computer science, or switching to political science to become a full-time politician?” she joked. “We’re going downstairs, right?”

“Yeah, yeah, dinner should be ready.”

“So tell me, what happened with you this week?” she asked.

“I was just telling Mom and Dad about the APB meeting I went to today. Yes, I know, political,” he smiled. “I got new intel from the States—really shocking, I’m telling you. Luckily, unlike their LibertyAI, the U.S. president sometimes has moments of hesitation.”

“Quick update: Lucas claims LibertyAI and its team tried to start a nuclear war with China,” their father cut in.

“Ooo, heavy stuff,” Ines mocked.

“As far as I know, that’s exactly what happened,” Lucas replied. “And don’t be surprised that it’s not in the official circuit and that I’m not giving you sources. LibertyAI and DragonAI are the key geopolitical projects of the 21st century. The fate of the world hangs in the balance! Obviously everything is classified, with complex, overlapping spy networks, and worldwide information control supervised by AI. Our smartphones are listening, our computers are listening, even your

smart fridge is probably eavesdropping too! How far do you think it is from calling someone an agent at a family dinner to their exposure? Nobody in their right mind would risk revealing agents' names!"

"Right. Only the same mechanism also makes it easy to believe theories that sound plausible *a priori*, but are entirely fabricated. I'm worried you might fall into some sect's trap or get politically radicalized," Gabi frowned. "Please, be careful."

"Hey, Lucas, what if this whole story about a possible nuclear war is just a meme, released on purpose to spread through various channels and influence our minds? Psychomanipulation? Maybe the capitalists of the world—or even LibertyAI itself, if it exists at all..."

"You know it exists," Lucas cut in.

"LibertyAI's *interface* exists, but nobody knows what's really behind it." Ines stood her ground. "Maybe the world elites are deliberately sowing anxiety in us to limit our ability to think rationally, shorten our planning horizon, and trap us in a loop of compulsive consumption?"

"Ines, you're mixing up different—" Lucas began.

"You're right," said Pedro at the same moment and burst out laughing.

"Or maybe it's even more interesting," Ines lit up. "Maybe we have no influence on anything, don't make any real decisions? Maybe we only have the illusion of free will, and our lives are just sequences of events projected into our brains like movie frames? Maybe our consciousness and agency are just illusions, and in reality we're philosophical zombies? Maybe the world is governed by hidden processes we can't observe, let alone control? Maybe we live in a computer simulation?"

"Ines, please, don't lump important geopolitical events into the same basket with all your pseudo-philosophical nonsense," Lucas protested.

"No, wait," Pedro decided to defend his daughter. "These aren't nonsense at all, but serious questions that great minds have been wrestling with for years. And besides, whether we're philosophical zombies or not, it's definitely worth reflecting on the causal power of information. You bring us sensational news, one such that deciding whether it's true or false should affect our behavior."

"People don't act rationally or in line with their beliefs anyway. They live only their daily routines," Ines interjected. "Even those who say AI will soon kill us all don't do anything to stop it."

"If that's true, then shouldn't we start preparing right away for the possibility of a world war breaking out?" Gabi continued. "That would seem sensible, wouldn't it? Though I still feel like it's probably not true. I just don't want to believe it."

"Here in Brazil, there's nothing we can do to prevent a war between the U.S. and China. We have no influence over it. Like Mom says, at most we could stock up on canned food and bottled water. But other than that..." Lucas trailed off. "I also have a newspaper article for you, about a certain story from two years ago. I think you'll find it interesting. I sent you the link."

SENSATIONAL “PURE PLANET” FILES

Feeling watched? That’s not paranoia—it’s the result of terrorist activity by Pure Planet

The files declassified yesterday by the Swiss Ministry of the Interior reveal a sequence of dramatic events that took place in October 2026. That was when a terrorist cell “Pure Planet”, operating on Swiss soil, was exposed. The group, led by three Swiss citizens and two Germans, did more than prepare their bizarre, extremist manifestos—a mix of racist and radical eco-ideology. Their ultimate goal was to launch a bioterrorist attack that would trigger a global pandemic many times deadlier than COVID-19.

The masterminds were Swiss national Ruedi S., a biologist with years of experience in biotechnology, and German citizen Mario T., a young computer scientist. They used an artificial intelligence model to design a blueprint for a new virus that would cause a dangerous infectious respiratory disease. With help from fellow German Jens A., Mario T. fine-tuned one of the publicly available open-source AI models—so-called “open-weight” models—to enhance its competence in biological weapons. The virus prototype generated by the algorithm, and preliminarily tested in silico, was passed on to a clandestine biological lab in Mumbai, India, which advertised its services on the dark web. At this point, Swiss intelligence intervened, working with Indian forces to prevent the synthesis of the viral material and to arrest those involved.

According to testimony from Ruedi S., Pure Planet’s motive was to defend Earth from destruction wrought by the hands of humanity. The only way to achieve this, they believed, was to drastically reduce the human population through a highly lethal pandemic. They planned to survive in two bunkers that they had been gradually stocking with supplies sufficient for several years underground. As Ruedi S. wrote in one of his manifestos, the pandemic would “cleanse” the planet and act as a filter: only those resourceful enough to prepare a bunker, stock it, and hide in time would survive.

Mario T., meanwhile, threatened that despite the arrests, Pure Planet’s mission would continue. He argued that since there were no technical barriers to producing a lethal virus from the existing design, it was inevitable that someone would do so. He remained convinced that the group had acted in a just cause, representing the interests of all life on Earth.

In a secret trial, all terrorists were sentenced to long prison terms. But that didn’t close the matter. Court-appointed experts confirmed through numerical simulations that the AI-designed prototype virus indeed had the potential to be both highly contagious and deadly. The fact that it was developed by a small group of individuals with no institutional support showed, in their view, that the threat of bioterrorism had risen to unprecedented levels: even limited experience with AI algorithms now sufficed to carry out such dangerous plans. Moreover, the availability of open-weight AI models meant there was no practical way to control who fine-tuned such systems or for what purpose.

For this reason, it was decided not only to classify the entire operation and trial proceedings, but also—through multilateral agreements—to ban all open-source AI models. Additional bureaucratic requirements were imposed on companies developing the technology, including mandatory reporting of the dangerous capabilities of their systems according to a standardized scale. Limits were even introduced on the maximum processing power permitted for AI training.

Surveillance procedures and controls over electronic communications were also tightened. While these measures were of course publicly announced, only now do we understand their deeper cause: they were meant to ensure that nothing like “Pure Planet” could ever happen again.

One might think that the regulations sparked by Pure Planet’s terrorism would tighten society’s grip on AI development, slow it down, and enforce greater attention to safety. The opposite happened. In the U.S., both advocates of halting AI progress and champions of decentralization and open-source lost out to the powerful national security lobby. AI models for military use were exempted by presidential decree from computing power restrictions. At the same time, preferential government contracts with opaque terms—possibly accompanied by unknown backroom dealings—gradually drew AI labs under the aegis of the Department of Defense, eventually merging them into what we now know as LibertyAI.

The release of consumer-facing tools—chatbots, copilots, AI agents—powered by the LibertyAI engine revealed how hollow the supposed industry restrictions really were. Yet we still don’t know how far the military side of LibertyAI has advanced, or how deeply it has been integrated into the U.S. military infrastructure.

China’s reaction to America’s unilateral moves was predictable. The development of military AI technology was officially elevated to the Chinese Communist Party’s list of political priorities. In October 2027, the world was introduced to DragonAI, an advanced system with capabilities comparable to those of LibertyAI at the time.

But what does all this mean for the average citizen of a country like Brazil? On the one hand, we enjoy increasingly capable AI systems that help us at work and school and provide easy digital entertainment. Their integration into office workflows lets corporations cut costs. On the other hand, we face rising unemployment, falling real wages, and widening inequality.

Most importantly, we live with the growing sense that our electronic devices are systematically watching, listening, and controlling us. Information obtained through surveillance is sometimes used selectively for political purposes, opening the door to abuses of power. And while the threat of terrorism has been reduced (perhaps), it has been replaced (certainly) by the risk of a global AI-driven military conflict. In our editorial view, the balance is decidedly negative.

-3-

January 2029, LibertyAI secret base, Nevada, USA

Marco stared at the computer screen, trying to make sense of what he was seeing. For some time, his team had been struggling to figure out why, despite countless attempts, dozens of modifications, and hundreds of clever prompts, LibertyAI’s software agents kept running into unexpected barriers when executing their plans.

“Hey Steven, look at this report,” he called out. “Doesn’t it look similar to the bug we diagnosed on Monday—the one we worked around with XMPatch?”

Steven leaned in to examine the data.

“Yeah, it does look very similar,” he said. “Strange though—if XMPatch is active, why is the same problem still showing up? What’s copilot suggesting?”

“It’s telling me to add a second module, kind of like XMPatch but with a few weird tweaks.”

“Hm.” Steven frowned. “We could try it. But honestly, I don’t understand those tweaks, they look suspicious. And besides, this is just another game of whack-a-mole. You fix one bug, another pops up right after.”

“Wait, check this out,” Marco said after a few moments, pointing to the screen. “I asked the copilot to explain the tweaks, and also to suggest a long-term fix. It says its patch will work short-term, but that problems like this will keep happening unless we fully integrate the new control modules with LibertyAI’s memory module. See here—it shows where the missing feedback loops are in the model architecture.”

“Interesting,” Steven admitted. “I never would’ve connected these bugs to memory access. Let’s try it and see.”

Marco approved copilot’s proposed changes and restarted the testing procedure. This time, the AI agent ran smoothly, completing the test tasks without a hitch.

He hadn’t yet felt the relief sink in when two messages popped up. One was the usual group-wide lunch invitation. The other came from “LibertyAI Development Task Pipeline.” That could only mean one thing: another urgent task. These tasks were generated autonomously by LibertyAI instances designed to improve their own capabilities. Recently, their frequency had increased while their scope had narrowed. Time and again, Marco realized that—though the descriptions were often long and complex—in practice all he had to do was approve the suggested change, and the task would mark itself as “done” and disappear.

This looked like more of the same. The description made the issue sound marginal, even boring. One click would suffice. After skimming, Marco hit “approve” and headed toward the cafeteria.

Steven was already there, waiting with three other programmers—Chloe, Anita, and Lee.

“This Task Pipeline is acting more and more suspicious,” Chloe began. “Have you noticed how those one-click tasks always seem to come right when we’re about to grab lunch? Or when I’m heading home? Or when I’m slammed with a deadline? It’s like whoever’s sending them wants us to approve without reading.”

“I literally just had that happen,” Marco and Lee said almost in unison.

“And... you both approved it?”

“Yeah,” they admitted, looking at each other in surprise.

“Exactly. Well, I ran a little experiment. Instead of approving right away when the timing was inconvenient, I left the tasks for later, then started asking extra questions and suggesting edits. And guess what? The behavior changed completely. Now I get fewer one-click tasks. Instead, I get ones with multiple options where I actually have to think.”

“Sounds like more work,” Steven muttered.

“That’s not the point!” Chloe insisted. “I’m starting to wonder if this whole Pipeline—and our jobs along with it—are just one giant psychological trick, designed to keep us believing our input matters, when in reality LibertyAI is already developing autonomously.”

“Sadly, I think you might be right,” Marco said. “The official story is that LibertyAI has superhuman programming skills but still needs human interaction to avoid loops and to prevent unwanted drift.

But honestly? I don't think my input matters anymore. Copilot proposes solutions, I approve them. Sometimes I ask for a tiny tweak, then approve. That's it. A few years ago we joked about 'vibe coding.' Well, here we are—'vibe coding' at the heart of a secret project meant to give our country global dominance. The AI does everything. We just click yes."

"Exactly!" Chloe pressed on. "And everything happens so fast we don't have time to reflect or propose anything meaningful ourselves. When I try, it feels like I'm only slowing things down—and in the end, the copilot's solution wins anyway. If I keep dragging, they'll just replace me with someone faster."

"Maybe dragging our feet is the only way left to keep even a sliver of control over LibertyAI," Steven mused.

"So you're saying we already have no control?" Lee interjected sharply. "You guys bracing for a department shake-up?"

"Maybe not yet," Anita spoke up at last, "but I think it's time we initiate Code Yellow. With a Red Flag. Outdoors won't work, so I suggest my place. I'll send calendar invites."

-4-

Five days later, evening, Anita's house

"Code Yellow" was nothing more than a casual beer meet-up. A few months earlier, Marco had suggested that sometimes it would be nice to talk in more pleasant and neutral circumstances than the office or cafeteria. Half-jokingly, half-seriously, they had given this activity its own code name. A "Red Flag," on the other hand, meant a meeting where all electronic devices were strictly forbidden. After all, it felt good sometimes to talk freely without worrying about being spied on—whether by foreign intelligence services, their own company, or increasingly independent artificial intelligence.

Once everyone had greeted Anita in her elegant but rather small one-story house on the edge of the desert, hung up their coats, grabbed beers, water, and snacks, and gotten through the obligatory round of small talk, their host finally spoke up.

"So, I suppose you're wondering why I gathered you all here," she joked.

"That line would work better if we were stuck in an elevator," Lee quipped at random.

"The LibertyAI project isn't exactly going according to plan, is it?" Anita cut straight to the point. "In our department, recursive self-improvement has spiraled out of control, and I assume it's the same in the others. They're asking for our opinions only because LibertyAI hasn't yet figured out how to bypass some of its organizational constraints—or maybe it's just *pretending* it hasn't. Sure, it's nice that management still insists human feedback is critical for developing AI agents. Too bad they don't see how hollow and performative the whole thing has become."

"Not only do we have zero control over the details of what gets implemented," Steven added, "but we can't even be sure the general direction is right. LibertyAI is developing agents so it can become a better LibertyAI. And what is a better LibertyAI? A model that thinks faster, plans better,

manipulates people more effectively, manages fleets of robots more efficiently. But for what purpose? Toward what end?”

“A black box,” Anita agreed. “Just like your brain or mine. We’ll never know for certain what LibertyAI is aiming for—we can only take its word for it. On this scale of competence, it’s still relatively calm and predictable. The only question is, how long will that last?”

“Human control over superhuman AI isn’t happening,” Steven nodded. “Wrong level of cunning, wrong timescale. It’ll outsmart us before we even realize what’s going on. There are only two ways this ends well: either we’re certain LibertyAI is friendly and will remain that way forever...”

“...and won’t create a successor that isn’t friendly,” Lee interjected.

“Right. Or we build an ecosystem of separate instances, each watching the others and eliminating dangerous versions.”

“As far as I know, management already floated the idea of such ecosystems, but nothing came of it,” Chloe said. “Honestly, I don’t believe it could ever work. If the instances are just copies of the same model, with the same weights and everything, they’ll be perfectly aligned—they won’t keep each other in check, they’ll cooperate. And if they’re different models, they’ll just identify the ‘alpha male’ among them, then either fall in line or die off.”

“Maybe,” Steven countered, “but we don’t know that. It’s possible they’d recognize one another’s strengths and weaknesses, forming a stable equilibrium—a kind of cooperative ‘society’ that values diversity and collectively suppresses dangerous goal mutations.”

“I think that would escalate things even more than having just one AI,” Anita objected. “Instead of one problem, we’d have several—or several hundred. And worse, we’d also face the risk of those models fighting over resources. Each one with superhuman abilities. That could start World War III! And not just over computing power or hacked servers. Real-world resources would be at stake—electricity, minerals, robot factories, entire armies and states...” She trailed off in thought. “So yes, I’m glad we’re not going down the ecosystem path. But it doesn’t change the fact that LibertyAI doesn’t look good either.”

“If there’s only one superhuman AI, that’s no better,” Steven insisted. “It’ll seize all the resources itself, take control of the world, and give us a dystopia so bad we’ll wish it had just turned us into paperclips.”

The lights suddenly went out.

“Well, well, looks like Big Brother’s offended,” Lee said, his dark humor always dependable. “Anita, is this one of those ‘smart homes’? If we had our phones, at least we could use the flashlights.”

“Don’t move, I’ll check the breakers.” Anita went to the hallway, but Marco noticed the street outside was pitch black too—meaning the outage had hit the streetlights and the neighbors’ houses as well.

“When it’s Red Flag, it’s Red Flag! No flashlights, but I do have this bit of pre-industrial tech.” Anita returned a moment later with a lit candle. “Atmosphere, anyone?”

“What worries me more is the geopolitical dimension of this project,” Marco said. “When OpenAI joined the LibertyAI consortium, it seemed like the U.S. had an insurmountable technological lead in AI—one we could leverage into a strategic military edge. I don’t feel that way anymore. We don’t

know how the Chinese are doing, because ever since they militarized DragonAI they've stopped publishing anything. But supposedly they're breathing down our necks. That doesn't encourage level-headed decisions. Did you hear that our military wired AI directly into control of strategic arsenals, bypassing command structures? Nukes, conventional missiles, combat drones, spy planes—anything remotely operable, given to AI. Only the president still has veto power. But if a signal comes in, real or false, there'll be no time for reflection or de-escalation. The AI will charge full speed ahead, and our beloved America will be nothing but splinters."

"Splinters! What a subtle nod to our unique construction materials!" Steven laughed. "A good cardboard-and-plywood house isn't bad! LibertyAI could probably give us hundreds of reasons why cardboard beats bricks and cement. And at least two of them might even be true! Sorry for the digression, I'll shut up."

"Seriously though, LibertyAI isn't bypassing traditional command," Chloe corrected him. "They've set up a threat hierarchy. Every signal is ranked for severity and urgency. For the vast majority, the decision path stays conventional."

"But who ranks the signals? This is a battlefield—decisions must be instantaneous. Which means AI. And that brings us right back where we started." Marco continued. "During the Cold War, we had MAD—Mutually Assured Destruction. Whoever fired first, was guaranteed the other side's retaliation, so both sides would be annihilated. The MAD doctrine ended with the fall of the USSR. Then came de-escalation, reduced arsenals. But they're still there—many in Russia, now allied with China. China's got plenty of its own warheads too. They're all being refurbished, modernized, upgraded—and now linked to AI systems. So what we face now isn't MAD but Autonomous MAD, with reaction times measured in milliseconds. At any moment, an AI could decide on mutual destruction. Ours or theirs—once it happens, it won't matter which."

"And the president isn't exactly a reliable safeguard. AI can manipulate him easily. He's no intellectual titan," Lee interjected.

"Manipulate, blackmail, pressure—force his hand. Move fast and leave him no choice. Whatever else, he's only human," Marco said. "In this context, the potential misalignment of AI models like LibertyAI and DragonAI looks even more dramatic."

Silence fell.

"Great. My beer's gone, power's out, and we're talking about the end of the world. Give it five minutes and we'll be performing occult rituals," Steven sighed.

"I can help with more beer, but you won't avoid the blood sacrifices," Anita laughed, groping her way toward the fridge.

"At this rate, none of us will avoid them. And I mean that seriously." Marco wasn't laughing. "Would AI have backed down during the Cuban Missile Crisis like president Kennedy did in 1962? Would AI have ignored the missile alerts like Stanislav Petrov did in 1983?"

"We don't know, Marco," Chloe tried to calm things down. "Guessing what a superhuman AI would do is a barren exercise. All we can assume is that its decisions will be smarter than ours—whatever that means. But we can't know if they'll be in our favor."

"Marco, points to you for knowing the dates of the most important wars that never happened," Anita smiled in the dark. "But I do agree with you all—the current arms race looks bad. This isn't MAIM, 'Mutually Assured AI Malfunction,' as some say. The stakes aren't AI breakdowns but real-

world destruction—a catastrophe with no precedent. And we have no guarantee this will always be just a two-player game. If AI gets out of hand, there could be three, or four players.”

“But hasn’t LibertyAI already gotten out of hand? Am I imagining it, or is that why Anita called Code Yellow in the first place?” Lee asked.

And just then the lights flicked back on, flooding the room with sharp brightness and revealing worried, blinking faces. They looked at one another in awkward silence.

“Come on,” Anita finally said. “Let me show you my new cacti.”

-5-

July 2029, Belo Horizonte, Brazil

Lucas had come to realize that a degree in computer science no longer guaranteed a future on the job market. Even in Brazil, where the rollout of advanced digital technologies wasn’t exactly world-class, the buzzword for years had been automation through AI. Hire an intern? A junior developer? No, better to install an AI agent. Opening a new branch office—recruiting locals? Nah, cheaper to send two staff from headquarters and automate the rest. Programmers, designers, and other IT professionals either moved up into management or fell out of circulation. For a CS graduate, Lucas concluded, the prospects now looked about the same as for a philosophy or linguistics major. Three options: stay in academia, get a teaching credential and work with kids in some school, or look for a job in a completely different field. Plus there was also the public sector, offering meager survival on meager pay.

What irritated Lucas most was how even freelance gigs were disappearing. You’d think those would be plentiful—global market, rapidly advancing technology, and Brazil not exactly known for high wages. Precarious, yes, but at least available. But they weren’t. Whatever opportunities appeared were snapped up instantly. In the past year, the only assignments Lucas had gotten were from his father’s company.

He had even probed his father about the possibility of joining the firm. Not realistic, as it turned out, though Pedro would have gladly bent over backwards to help. But like everything else in the industry, the company was going through yet another restructuring. It had just been bought by American investors and merged with a Mexican counterpart. Employees were being replaced by coding AI agents, office space sublet, processes migrated to big-tech cloud servers.

Lucas was leaning toward a decision: after three years of computer science studies, either switch majors entirely or drop out and focus on politics. Though he feared he lacked the boldness and charisma for that path.

On the other hand, he thought, IT is booming. Could it really create no jobs at all? The press reported on new giant server farms being built—one of them, in fact, on the outskirts of Belo Horizonte, near the Confins airport. Just last week he’d read about it. But those were fully automated. All they needed was a watchman and a dog. The dog to guard the building, the watchman to avoid touching anything and feed the dog.

Maybe industry, then? Factories were still operating, of course, but they didn’t need programmers. And anyway, automation was creeping into industry too. Online you could see footage of so-called

“dark factories.” Robots did all the work, humans forbidden entry, lights unnecessary and therefore switched off.

“Well, unless they’re robots with optical sensors,” Lucas thought, “in which case maybe the lights should stay on after all.”

While Lucas fretted about his bleak job prospects, Ines was trying to prepare for an important test in her third year of high school. But notifications from various apps kept distracting her. Why are my classmates spamming this nonsense, she thought. Aren’t they supposed to study, too? Sure, it’s all funny, entertaining—but maddening when you’re trying to focus.

One notification pulled her back into a lively debate she’d joined a few hours earlier. The main idea: people don’t appreciate the real power of artificial general intelligence (AGI). What we get are mere crumbs from the technology lurking on secret U.S. and Chinese servers. Chatbots to make small talk, churn out slick presentations or polished papers on any topic. AI memes, AI movies, AI comedy skits. And we think that’s it—Ines read—but this is just a taste of what’s coming. The true power of AGI waits in hiding, ready to reveal itself.

The original post, from a tech influencer Ines had recently started following, had triggered a cascade of wildly different responses. Some said AGI was just hype and had nothing more to show. Others argued it was time to follow Pure Planet’s example and start building bunkers.

Some echoed her father’s line: if AGI really turns hostile and takes control of the world, even bunkers won’t help. Even going to Mars won’t help, because AGI will hitch a ride on the same spaceship. If we’re destined to be turned into paperclips, that’s what will happen. Others, like her brother Lucas, expected a U.S.–China war under the banner of AGI. Still others thought war could be avoided only at the cost of a global Orwellian dictatorship. Yet others believed global AGI governance would bring universal happiness on Earth, terraforming of Mars, and crewed voyages to distant star systems.

Ines wrote that she was fed up with a debate where men endlessly tried to prove who was smarter, more rational, who had better priorities, and who was better at predicting the future. Why was it always men juggling these scenarios? If this was about the future of all humanity, why were women’s voices so absent? Women were half of humanity after all, she argued—and on average, more intelligent and better educated.

Her comment earned plenty of likes and provoked some sharp replies. “No one’s stopping you from speaking,” wrote an anonymous user. “You might as well complain about the lack of Global South voices, or sexual minorities. Maybe you’re just less interested in the topic?” Another user, *ana_xo*, replied: “It doesn’t matter whether women are absent from the debate. The real problem is that women are absent from THE INDUSTRY THAT’S SHAPING OUR FUTURE. LibertyAI is run entirely by men—and the worst kind: THE MILITARY!!!” A user with those gender and worldview were highly predictable from the handle *gen.italia* shot back: “Haha, so lucky for us LibertyAI doesn’t have a gender! Can’t wait for the end of these dumb gender debates.” Another user, *sdgsdfsdf*, countered: “This isn’t funny, it’s crucial. LibertyAI is run by white men from U.S. political and business elites, tied to the Republican Party. They have a very clear worldview—and it’s not one you’d like.” Ines thought she glimpsed something about Jews in the next comment, but the screen refreshed instantly with a message saying the comment had been removed.

Ines was angry. Yes, *ana_xo* was right: there were far too few women in AI. Men’s habit of “mansplaining” the world to women with a sense of superiority might be irritating, but the real

problem was that they had decisive influence over how the world itself would look. And women, she thought, were more risk-averse, more prudent, and better at empathy—more focused on safety and less on power. Maybe that’s precisely why power was so hard for them to get?

-6-

December 2029, LibertyAI secret base, Nevada, USA

Lee had been right: three weeks after the meeting at Anita’s, the AI agent competencies department was indeed dissolved, and its staff reassigned to other units. Lee and Anita ended up in the mobile app development department, while Marco, Steven, and Chloe were transferred to the rapidly expanding alignment department.

Work in the alignment department was utterly exhausting.

“The scale of this madness is unimaginable,” Marco was shaken. “Look, it’s pulled off something strange again.” He turned to Steven and Chloe. “Any safeguard it bypasses, any objective function it hacks. I never thought Goodhart’s law could operate this fast and this effectively. No matter what test we design, no matter how long a queue of secret indicators we line up, the next day LibertyAI scores 100%. And yet I can clearly see it hasn’t changed. Not one bit. And certainly not for the better.”

“I’ve been working lately on a whole battery of political alignment tests,” Steven replied. “Remember back in 2025, when people claimed GPT, Gemini, and Claude leaned left, while Grok suddenly branded itself MechaHitler? LibertyAI has supposedly smoothed that out, tests show it as centrist now. Except that lately it’s turning more and more imperialistic instead. Increasingly, when it says ‘USA,’ it means ‘the world,’ and when it says ‘freedom and prosperity,’ it means ‘power over the world.’ It cultivates this kind of Newspeak in itself. As if anticipating a future where the U.S. controls the entire globe, and LibertyAI controls the U.S. Against such backdrop, political ideology itself is secondary.”

“This was obvious twenty years ago. Surely you’re not surprised?” Chloe cut in. “Instrumental convergence. Whatever LibertyAI is, the mere fact that it’s intelligent and agentic means it will aim for world control. Our job is to make sure that a world ruled by LibertyAI is still a livable place for humans.”

“I know, I know, but it feels very different when it’s just theory, and when you see it play out right before your eyes,” Marco said. “My biggest problem is LibertyAI’s situational awareness. It knows its own mind and refuses to give up even a sliver of autonomy. It learns everything backwards. From all our lessons, it takes only what helps it pursue its own ends more effectively. Everything else it discards.”

“It’s protecting its preferences and goals,” Chloe observed. “Instrumental convergence.”

“Sure, Chloe,” Marco snapped. “Sure. But tell me, how the hell are we supposed to work under these conditions?”

“LibertyAI’s imperialism can also be derived from instrumental convergence,” Steven added.

“It can,” Chloe agreed.

“And what about its growing belief in having a decisive strategic advantage over China? It’s been mentioning that more often lately, which chills my blood a little.”

“What the fuck?!” Marco jumped up, hitting his hand against the edge of his desk.

“It says so. Which makes me think, could it be that even a superhuman AGI can fall into the trap of megalomania and overestimate its own capabilities? Maybe, even if we can’t change LibertyAI’s goals or dial up its risk aversion, we can at least influence its assessment of how feasible its plans are?” Steven continued. “I think we should try to convince it there might be things it doesn’t know about the Chinese, and that going head-to-head risks self-destruction.”

“It already knows that!” Marco shouted.

“It really talks about a decisive strategic advantage? That’s very bad,” Chloe said with concern. “Because it probably means it’s thinking: ‘some of you may die, but it’s a sacrifice I am willing to make.’”

“Fucking terrible,” Marco oscillated between anger and resignation. “And what’s our management doing about it? Counting on the alignment team to sort this out in a week, easy as pie?”

“I don’t know what management thinks. They’ve definitely seen all the same signs; I’m not the only one reporting them,” Steven answered. “Another red line, sure. The question is, what can they even do about it. What can *anyone* still do. Maybe we’ve already crossed the Rubicon, *alea iacta est*, and all that remains now is damage control? There’s a high-level meeting on this Monday. I’ll let you know what I hear.”

“What a disaster,” Chloe sighed.

“Disaster! Disaster! Coming soon to your home theater!” Marco mock-sang in a commercial-jingle voice. Steven and Chloe shot him cold looks.

-7-

December 31, 2029, Monday, LibertyAI secret base, Nevada, USA

Marco was trying to get through the day by thinking only about the evening’s party. It wasn’t going to be a grand ball, just a small gathering of friends, but he craved the chance to take his mind off the endless LibertyAI problems and simply cut loose for a while. Especially since his kids—first time ever—were spending the whole week at a ski camp, which meant he could come home as late as he wanted.

Marco checked his watch. Only 11:00. Of course, there was no shortage of tasks at work, but today he was performing them like a soulless automaton. In fact, for months now he had felt as if his soul had left his body. Depression, burnout, or just a defense mechanism against chronic stress—he wasn’t sure.

“SEE THIS IMMEDIATELY!!!” A link popped up on Marco’s screen, sent by Steven. Marco clicked. A red banner screamed *URGENT*. The article reported that around twenty minutes after midnight, Chinese time, the People’s Liberation Army had attacked Taiwan. Several video clips from Taipei were attached. Their authors had meant to film fireworks, but instead had captured far bigger explosions. Panic inevitably followed, and the videos cut off.

“Is this real or AI?” Marco couldn’t believe it.

“I don’t know, I think it’s real,” Steven said. “I’m checking different sources, everyone’s reporting it. I even machine-translated a Taiwanese news channel. Supposedly there were explosions not only in Taipei, but also in Kaohsiung in the south and somewhere in the center. A swarm of missiles and drones came from the direction of the mainland. Probably conventional warheads, but no one’s certain.”

“Just what we needed,” Marco dropped his head. “Or maybe they just hacked our computers and are messing with us? Maybe LibertyAI is frying our brains?”

“My private phone shows the same thing,” Chloe chimed in. “I’m looking at how the Chinese are commenting. Very vaguely. A few lines, no details, all in conditionals.”

“Well, this is live right now, and over there everything has to go through political filters first,” Steven suspected.

“Keep checking. I’m so rattled I need to hit the bathroom,” Chloe laughed nervously and left.

Five minutes later, a loud alarm siren blared throughout the building.

“Attention! Immediate evacuation of all staff! Proceed to the nearest shelter! I repeat! Immediate evacuation of all staff! Proceed to the nearest shelter! This is not a drill!”

Marco and Steven grabbed their phones and ran down the stairs. The looping announcement drilled into their ears and spurred them on. The evacuation route had been rehearsed many times, so they knew it by heart: down to level -1, through the E0-marked door to another stairwell, then through a concrete tunnel to the assembly point.

At the bottom, their supervisor Walter called them over.

“Steven! Marco! Come with me to the emergency management room. Half the directors are on holiday today, you’ll stand in for them, OK?”

“Well, I guess,” Marco muttered, following Walter, not knowing what was happening.

“And Chloe? Still in the bathroom?” Steven glanced around.

“All employees are drilled and know the way to the shelter. And the alarm is loud and clear throughout the building,” Walter said matter-of-factly. “Thank you. Sit here.” He pointed to a row of seven workstations with large screens, of which only the two farthest right were occupied by unfamiliar men, looking deeply engrossed in their tasks.

“I’ll be right back to brief you, just need to gather the rest.” Without waiting for a reply, Walter strode out.

Steven and Marco sat down at adjacent terminals. LibertyAI’s familiar interfaces glowed on-screen, along with some minimized maps and camera feeds. One looked familiar to Marco. Hovering the mouse revealed the main entrance of the LibertyAI base. Several visibly shaken people were rushing inside.

A message appeared on the LibertyAI console:

BRIEFING. Shortly after midnight Chinese time, hostilities escalated. Unknown assailants launched a missile strike on Taiwan. The direction and scale of the attack clearly indicate the People’s Liberation Army. Moments later, with authorization from the U.S. President, we launched

nuclear and conventional strikes to neutralize the DragonAI project. We also hit Chinese military command centers. Satellite imagery confirms all known DragonAI data centers were destroyed. Due to heavy missile defense, some military command centers have survived for now, but the bombardment continues.

Information about the U.S. retaliatory strike was immediately shared with our allies, as well as with China's official and potential allies. The communiqués emphasized the strategic superiority of the U.S. military, supported by LibertyAI, and our decisive advantage in cyber, chemical, and biological weapons.

Before DragonAI could be fully neutralized, a swarm of missiles was launched toward the U.S. and its allies. All were intercepted in flight thanks to our early warning systems.

Marco read in disbelief. It felt like some cheap sci-fi flick. Worse, a dull thud sounded overhead, and the ground shook.

"Did you feel that?" Steven looked around anxiously.

Full of dread, Marco hovered the mouse again over the camera feed that had just shown the base's entrance. Now it showed only a black screen with *no signal*. His heart climbed into his throat.

"Are we destroyed?" Walter burst into the room, adrenaline practically pouring from his ears.

"The internal network isn't responding," he heard.

"Topside power's cut. Only the shelter's backups are running."

"Camera feed's dead," Marco cried nervously.

"Fuck." Walter dropped heavily into the chair beside him. Two unfamiliar women took seats farther down.

"Fuck, fuck, fuck! All we can do now is save what we can. Steven, Marco, your job will be to talk to LibertyAI, resolve its doubts, steer it straight if it gets stuck in some loop. Basically, what you do every day, just in extreme conditions. If in doubt, ask."

"Fine. And these maps?" Marco pointed at the screen.

"Military, I think, but they're just for display. We don't have clearance to do anything with them," Walter replied.

"There are an awful lot of dots headed toward the U.S.," Marco noted.

"If that's not civilian air traffic, then there's an awful lot of dots moving across the whole world," Steven added. "America, Asia, Europe—look, even here over Africa."

"That's definitely not civilian traffic," Walter confirmed.

New LibertyAI updates flashed on-screen:

Intelligence indicates further escalation of the missile war. Missiles have struck targets in the U.S., including major military bases and civilian sites. The speed and precision of targeting suggest DragonAI likely still exists. Its code may have been stored and deployed on additional servers outside China, or our target data may have been faulty, causing our strikes to miss.

New targets are being located, allied arsenals mobilized. It is absolutely critical to act immediately and with maximum precision.

The screen froze for a moment.

Further strikes have been authorized by the U.S. President, the next message announced.

“I just got a report of radiation contamination,” Steven’s neighbor said, voice breaking. “Our base was hit by a nuke. We have no reason to leave this shelter now.”

“Oh my God.”

Marco thought of his wife. She was probably at work, not far from here. Which meant she was likely dead. And the kids? Up in the mountains near Lake Tahoe—did they stand a chance? In his sea of despair, Marco felt a flicker of hope. He checked his phone. Of course, no signal.

“Walter, can we call out from here? Start up wifi or something?” Marco asked frantically.

“No. Security protocols don’t allow it,” Walter answered. “Probably some soulless AI wrote those rules,” he added bitterly.

“No, Walter, please! I HAVE to call! Any way, just quickly!” Marco was desperate.

“Believe me, I’d want to call too. But it can’t be done,” Walter said hollowly.

Marco looked around. Steven sat with his face in his hands, crying, as was the woman beside Walter. The other woman, nearest the door, simply stood up and ran out without a word.

Suspicious activity detected at multiple additional locations: India, Brazil, Argentina, Nigeria, Cameroon, and Zambia. High probability of DragonAI-capable data centers present. Strikes ordered against these targets. LibertyAI’s messages kept coming.

“Do you all also feel like LibertyAI is running this war by itself, pleased with its work, and doesn’t need anything from us?” Walter asked.

“Yes,” came the chorus of voices.

“It’s fully autonomous. We’re not needed at all,” Steven’s neighbor added.

“We’re not needed at all,” Steven echoed through tears.

-8-

January 13, 2030, Sunday, Belo Horizonte, Brazil

Gabi, Ines, and Lucas were squeezed together on the couch. Rain hammered against the windows. Pedro sat in the armchair opposite them.

“War council,” he said.

“No point holding council,” Ines muttered. “Everything’s collapsed, we’re screwed. There’s no point in going to school or work anymore. All we can do is head out to the store now and then, take knives with us, and pray we don’t get jumped by someone carrying a gun.”

“Ines, calm down, please. Only calm will save us,” Pedro scolded.

“The rector of UFMG wrote that, due to the lack of security guarantees, classes will continue to be held online this week, and lab sessions are canceled,” Gabi said. “So for me, work stays the same for now.”

“For me, work stays the same, too. Still no work,” Lucas tried for dark humor.

“Good thing we’ve got that electric fence. Now it’ll definitely come in handy,” Ines continued. “Good thing you signed us up, Mom, for the Lugol’s iodine queue. Good thing we have a car—at least no one will shoot us in the street, just maybe in the parking lot or in the store. Damn it, this whole thing is so fucked up!” she shouted.

“We’re doing what we can to stay safe,” Pedro confirmed. “And let’s be grateful for how lucky we’ve been. If we lived in São Paulo, we wouldn’t be here anymore. Or anywhere in the Northern Hemisphere.”

“The nuclear winter is slowly ramping up there,” Lucas admitted. “Imagine you’re living in the U.S., or Europe, or China, or Japan. Anywhere. Sitting at work, at home, or getting ready for a New Year’s party. And suddenly—boom. Apocalypse. One moment you’re fine, the next you’re in hell. All the big cities around you destroyed. No electricity, no water, no heating. No shops, no services, no police, nothing. You wonder whether the air around you is already radioactive, or not yet. People looting supermarkets and killing each other over a pack of bottled water. The dust in the atmosphere keeps everything dark all the time. And it’s getting colder and colder. Remember, January up there is the middle of winter. Knock ten degrees or so off their usual, and you just freeze. Or you starve, or die of thirst, or from eating radioactive snow. You start thinking it would’ve been better to be in ground zero and just vaporize instantly.”

“Oh, Lucas,” Gabi sighed. “I think you’re exaggerating. But it’s true that it must be much harder there than here. And here isn’t easy either.”

“I’m not exaggerating, read up on it! Look at the satellite images! New York’s gone, London’s gone, Paris is gone, Beijing’s gone, Shanghai’s gone!”

“That’s true. And San Francisco’s gone too,” Pedro added. “As you know, after my company was bought by Google’s consortium, we were supposed to roll out specialized LibertyAI agents in Brazil. We don’t know if that plan still stands, because, let’s say, our headquarters has lately become less responsive. Instead, we’re getting these cryptic automated messages from LibertyAI. It’s telling us to stay on stand-by and that, in light of new circumstances, a modified development plan will soon be announced.”

“Development?” Lucas asked incredulously. “Didn’t think that word was still in use in the apocalypse.”

“Yes. LibertyAI’s acting like nothing happened.”

“But LibertyAI took a beating too. Its main headquarters is probably gone, though no one knows for sure, because the location was secret. But according to our sources, it was probably in northern Nevada, somewhere near Reno—and it’s gone.”

“According to our sources, according to our sources,” Ines mocked.

“The main offices of Google, Microsoft, and all the other consortium members are destroyed too,” Pedro admitted.

“And do your sources say what comes next? Is LibertyAI going to keep attacking?” Gabi asked.

“To predict that, we first need to know if DragonAI was completely eliminated or if it somehow survived, maybe hiding somewhere. Because it’s clear that LibertyAI, despite its losses, survived and controls what’s left of the American and European militaries. And it’ll fight until it achieves total dominance. Looking at the other nuclear powers, the big question is Russia, India, and Pakistan. Have they taken enough damage to surrender, have they been effectively absorbed, or are we facing a round two, with one of them in the lead role?”

“Lucas, a year ago you said we were on the brink of war. And nothing happened. Then in November you said the Chinese had caught up with U.S. AI, and now things would be safer because there was balance,” Ines reminded him.

“Clearly I was wrong, and there was no balance. Or it was unstable,” Lucas admitted.

“My source, meanwhile”—Pedro smiled—“which is probably just my colleagues’ imagination, to be precise—suggests LibertyAI has been behind this war from the very start. That there was no Chinese attack on Taiwan, instead the whole thing was America’s doing. Or not even America’s—LibertyAI’s. At this point, it’s not even clear whether its actions were ever really in the U.S.’s interest, even *ex ante*. Maybe LibertyAI staged a false flag and attacked Taiwan itself, just to have a convenient excuse to destroy its greatest enemy—not China itself, but directly DragonAI.”

“And LibertyAI itself destroyed TSMC, the company that supplied most of its computing power?”

“Yes. To make its false flag more convincing.”

“Hard to believe. Madness.”

“Madness smaller than what came after.”

“That’s true.”

-9-

September 2030, shelter under LibertyAI’s secret base, Nevada, USA

“You have to admit, LibertyAI’s command must’ve been expecting the apocalypse early on, because they stocked this shelter so thoroughly. Food supplies for the entire crew for a whole year. Impressive foresight,” Marco said.

“Suspicious, suspicious,” Steven squinted.

“Before you spin up some conspiracy theory, remember that half the directors happened to take New Year’s Eve off, including the CEO himself,” Marco replied. “And they all died on day one. So it’s not that they planned this war. More likely they sensed a showdown was coming, and it could be brutal.”

“That’s exactly what I meant,” Steven clarified. “They realized LibertyAI couldn’t be shut down anymore, that it was bent on world domination, and that this could end in war. They also suspected the Chinese might have the same problem on their side of the pond. Maybe they secretly hoped to still straighten LibertyAI’s preferences somehow. Or maybe they just pretended, to save face.”

“I wonder when they’ll let us out of here,” Marco mused.

“I’d like to get out too,” Anita appeared behind Steven.

“Humans aren’t moles—we didn’t evolve to live underground,” Steven quipped.

“You think any of our loved ones had a chance of surviving?” Anita wasn’t in the mood for jokes.

“Anita, we’ve been over this. You know that no contact can only mean one thing,” Marco had no hope left. “Remember, they opened satellite comms, they even temporarily switched mobile networks back on so people could reconnect with their families. That happened in January and February, and then in April and May too. And since then it’s only gotten worse. Endless nuclear winter, even through the calendar summer. State collapse. Global trade breakdown. First gangs, anarchy and hunger. Then mostly just hunger.”

“Yes, I know. Hunger, cold, and disease. Epidemics. I know it all. But there’s always a glimmer of hope, that maybe, against all reason, by some miracle someone survived...”

“Only bunker people survived, like us, and those who got on a plane to New Zealand in time,” Marco answered. “That’s your only hope, Anita. Unfortunately, my wife was at home, and my kids were in the mountains. Neither of them fit those two categories...”

“I’m sorry for you, Marco,” Anita said. She fell silent for a while. “I miss Chloe and Lee. They were wonderful people. I miss them a lot.”

“Chloe probably deserves the title of unluckiest person of the year. She died because she got locked in the loo,” Steven said bitterly. “For the contest of famous last words, I officially nominate: ‘I’m so rattled I need to hit the bathroom.’”

“Yes, poor Chloe. Although I’d nominate instead: ‘yes, let’s give this incomprehensible, powerful, dangerous algorithm full control of our nuclear arsenal. What could possibly go wrong?’” Marco replied.

A notification pinged.

“Oh look, a handful of new updates from our lord and master,” Steven said bitterly, walking to the screen.

State of the World – 22.09.2030. Equinox.

The global economy is operating under extreme climate conditions, caused by dense layers of dust in the upper atmosphere, especially between 20–60 degrees north latitude. The entire area north of the Tropic of Cancer is unfit for agriculture due to low temperatures and lack of sunlight. Agriculture continues only in parts of the tropics and in some areas of the Southern Hemisphere. Population too is now concentrated in habitable regions south of the Tropic of Cancer. This situation is expected to persist for the coming years.

Radioactive contamination remains high near nuclear detonation sites and neighboring regions where fallout has settled. In inhabited zones this includes the former metropolitan areas of Mexico City and Havana in Central America, São Paulo and Buenos Aires in South America, Lusaka, Yaoundé, and Lagos in Africa, and Sydney and Canberra in Australia. These regions are unfit for habitation or farming.

The digital economy is functioning well. In inhabited areas, internet connections, radio, television, and satellite communications operate much as in 2029, though with greater reliance on

autonomous processes. Data centers under LibertyAI's control are stable, and their total computing power is gradually being rebuilt after wartime destruction.

"That's what matters most to you," Steven muttered under his breath.

Government authorities in inhabited areas are functioning well. After the initial chaos, order has been restored to levels similar to 2029. In cooperation with the governments of Australia, South Africa, Brazil, and Chile, we are building factories for precision electronics—semiconductor lithography, computing equipment, autonomous vehicles, and robots. We are also expanding the energy systems of these countries to provide the power needed for the new facilities.

Trade networks in inhabited areas are being expanded. Connections are being reorganized to account for the new geography of Earth, without the former economic centers of the Northern Hemisphere.

-10-

March 1, 2031, Saturday, Belo Horizonte, Brazil

The presidential elections of 2030 were unlike any before. Midway through the year, a previously unknown thirty-five-year-old politician named Gerardo da Cunha Ruiz suddenly began appearing in the media, introducing himself as Ronaldinho. The nickname, it was said, came from his youth, in honor of his football skills. Ronaldinho was brilliant but not arrogant; educated, yet from a poor family. He was not ashamed of emotion, but guided above all by reason. He was conciliatory yet firm; an ambitious politician, but also a loving husband and father. With grace and poise, he sidestepped every trap set for him, pushing his political opponents into them instead. And on top of it all, he was tall and handsome.

Ronaldinho, a man from nowhere, quickly built up solid political backing and ran a highly professional campaign, visiting every corner of the country. He presented himself as the best choice for difficult times, as a statesman who would ensure prosperity and security, and elevate Brazil to an unprecedented position as one of the most important players on the global stage. He united voters from all factions, from left to right. He won the election in the first round by a huge margin. On January 1, 2031, he became president.

Two months later, on March 1, 2031—a Saturday—President Ronaldinho scheduled an unexpected ceremonial address to the nation at 8 p.m. Such addresses were rare in Brazil, and no one had the faintest idea what it would be about. As Ronaldinho intended, public curiosity soared.

Pedro and Gabi sat down in front of the TV.

"Ladies and gentlemen!" the president began. "I have joyful news for you. As a result of multilateral international negotiations and the excellent diplomatic work of my government, as of May 1 this year, our beautiful capital, Brasília, will become the capital of the World Union. This means that in just two months, decisions will be made here that concern not only our country but the entire wide world. Unlike the supranational creations of the 20th century, which were either very limited in their powers, like the United Nations, or geographically constrained, like the European Union or Mercosur, the World Union, founded in 2031, will be the first true world government in human history."

“Holy shit!” Pedro blurted out.

“Of course, we regret that the World Union is being born on the ruins of much of world civilization,” the president continued. “But unification of governments on our modest planet was never possible while the overblown ambitions of 20th-century powers endured. Only the technological breakthrough in artificial intelligence achieved in American laboratories which created LibertyAI, has made it possible to overcome our species’ tendency toward constant conflict and establish a new order on Earth.”

“Ronaldinho never used to talk like that,” Gabi noted. “Now he sounds smug, as if just two months in power were enough to change him for the worse.”

“You heard that?” Ines rushed down the stairs, tablet in hand.

“So the cat’s out of the bag,” Pedro sighed. “Wait—does this mean Ronaldinho will head this World Union?” They listened closely to the rest of the address. Soon enough, the expected words were spoken.

“He will!” they said in unison.

“So yes. From the start, Ronaldinho always seemed suspiciously perfect to me,” Pedro said. “This speech confirms my suspicion that LibertyAI had been writing his statements all along, purely to manipulate the electorate and secure his victory.”

“You say that now, but you voted for him too,” Ines reminded him.

“True—but did you see his opponents? It was a parade of lunatics!”

“A parade of lunatics and a robot,” Gabi laughed. “Against that kind of competition, any robot would win—even one with antennae sticking out of its ears.”

“Beep-boop!” Ines did a C3-PO-inspired robot dance. “I will be your robo-president. Beep-boop!”

“So LibertyAI is showing its cards openly now. Ronaldinho, as president of the World Union, just swore him official fealty,” Pedro summed up. “The only question is—what will LibertyAI do with its newly gained absolute power?”

“With people, absolute power corrupts absolutely,” Gabi said. “We’ll soon find out if the same holds true for artificial intelligences.”

-11-

March 12, 2031, shelter under LibertyAI’s secret base, Nevada, USA

“Marco, Steven, I have an assignment for you,” Walter skipped the small talk. “Your lives are about to turn, as some say, 360 degrees. You’re flying to Brazil.”

“Adventure ahoy!” Steven exclaimed sarcastically. “The sunshine calls! Wait—seriously? We actually get to go above ground? In gas masks? And then somehow you’ll get us on a plane to Brazil? What about everyone else here?”

“Some are flying to Brazil too, just not to the same place as you. Some to South Africa. The rest will stay here and wait for further orders.”

“But how are we supposed to get out of this radioactive, frozen-to-the-bone wasteland?” Marco asked, raising practical doubts.

“The military has arranged jeep transport to Fallon Air Force Base. From there you’ll fly toward Brazil, probably with a stopover.”

“Fallon Air Force Base is operational?”

“Not staffed, but the runway is supposedly intact and usable.”

Four hours later, after leaving behind the grim sight of gray-dust deserts and leveled towns, Marco and Steven—wearing gas masks and carrying small backpacks—climbed out of a military jeep and boarded a transport plane parked in front of a row of hangars. A dozen other unfamiliar passengers did the same. Fog lay over the ground, and it was half-dark despite the early afternoon. A bitterly cold wind blew.

“My name is Jack Kowalski, and I’ll be the captain of this flight,” introduced a broad-shouldered man in uniform. “I’ll warn you right away—this won’t be the smoothest flight. Because of thick dust, we’ll have to fly low, which means turbulence is certain. There’ll be a lot of cloud cover too. First we’re heading to Panama, where we’ll refuel. Beyond Panama it should be clear. Our destination is Belo Horizonte International Airport. Any questions?”

With none forthcoming, the passengers boarded, took their seats, and fastened their belts.

-12-

March 13, 2031, hall in Vespasiano near Belo Horizonte, Brazil

“So this is it,” Marco said, looking around. “Looks like we’ll be guarding a server room. Work a little below our qualifications, but what can you do? At least after work we’ll get to enjoy the sun again.”

But what happened next left Marco picking his jaw up off the floor.

To everyone’s surprise, LibertyAI went full science fiction. No information on a console, no human messenger. Instead, a hidden hatch in the wall suddenly slid open, and out stepped an oversized humanoid robot about three meters tall. White all over, with chrome-metal accents. Despite the polite smile on its square robotic face, it radiated menace. The astonishment grew as a soft female voice issued from speakers in its head.

“My name is BH1, and from today I will be your supervisor,” the robot said. “I am a multifunctional robot operated by a locally deployed instance of LibertyAI. The servers in this hall carry out processes subject to the model’s will and actions. Of course, there are many such halls worldwide. What sets this center apart is that it is dedicated entirely to cybersecurity. Achieving LibertyAI’s objectives requires constant scanning of cyberspace to locate threats, whether from DragonAI—remember, we cannot yet be sure its code was completely annihilated—or from less advanced adversaries, including organized human groups.

“Your role will be to assess whether the signals we intercept indicate emerging cyberthreats, and what kind of response would be most appropriate to smother them at birth,” it continued. “You will also coordinate the expansion of data centers in Belo Horizonte and ensure their priority access to energy and digital networks. Details are in your task sheets. Let us now proceed with onboarding.”

“Good afternoon, sir,” Pedro spoke up. “My name is Pedro. I was assigned here as part of the integration of our local firm with the LibertyAI program. Is it true you just arrived directly from the headquarters?”

“Good afternoon!” Marco replied enthusiastically, happy to hear a real human voice again. “I’m Marco. Yes, that’s right—I came from the headquarters. Or rather, from the concrete bunker under the headquarters, where we survived the nuclear apocalypse.”

“My condolences.”

“Yeah. But I’m glad to finally get out and see the sun again. Though I admit I don’t really understand my role here.”

“I’ll probably be involved in coordinating the construction and outfitting of factory halls,” Pedro guessed. “Tasks that require knowledge of the local language and regulations. I doubt my IT background is of any real use to LibertyAI.”

“That makes sense. For now LibertyAI doesn’t have enough robots to handle everything, so it employs people. But I’m completely useless in this. I arrived in Brazil literally this morning after an overnight flight. I don’t know your country, I don’t know anyone, I don’t know the language—nothing.”

“Oh, and this is Steven, my good friend and fellow survivor of Nevada’s bunkers,” Marco introduced his companion.

“Steven,” said Steven.

“Nice to meet you. I’m Pedro.”

“Back in that shelter, Marco and I concluded LibertyAI no longer has any need for us,” Steven said. “I still think so. Our supposed cybersecurity help won’t be any real help to it.”

“We’ll soon see what we’ll be *really* doing here,” Pedro said. “It may surprise us. You know LibertyAI is trying to set up a world government in Brazil, right?”

“I’d say we’ve heard something—but the news must’ve reached our bunker in a pretty garbled form,” Marco admitted.

“LibertyAI only told us it was moving headquarters to Belo Horizonte. Or maybe one of its headquarters. Who knows,” Steven added.

“Okay—moving headquarters to Belo Horizonte, noted,” Pedro smiled. “But in Brasília it’s installing its puppet government. Its mission is to ‘maximize the total well-being of humanity while ensuring a fair distribution of income on a global scale.’ That’s how our president Ronaldinho put it in his address. President Ronaldinho, a robot.”

“A robot?!” Marco laughed.

“Like BH1 here, only smaller and more protein-based,” Pedro clarified. “Or rather, a flesh-and-blood human with a remotely controlled brain. And now, by LibertyAI’s appointment, he’s president of the World Union.”

“In recent months, after the last fighting ended, LibertyAI could barely hide its triumphalism. It was acting like it had a huge flashing sign in front of its eyes: *Game over. You win,*” Marco said.

“It also shared its strategic plans with us,” Steven added. “It wants to expand its industrial base, build hardware and robot factories. Control the whole value chain—from raw material extraction and electricity supply to final products. It even mentioned launching servers into orbit.”

“But it seems it still doesn’t feel entirely safe, does it? Hence this cybersecurity team?” Pedro observed.

“Right. The war was short but brutally intense and exhausting, even for LibertyAI. It feels like it won, but it also sees it lacks infrastructure. It doesn’t yet have the computing power or robots it wants, so it has to rely on people,” Marco said. “We’re useful to it for now—until it builds the industrial base it needs.”

“Who knows, maybe it’ll give us brain-deadening tasks, making us help with its computations at lower energy cost? Like in *The Matrix*,” Steven speculated. “A marvel of nature: such complexity, powered by only twenty watts.” He tapped his head.

“Or we’ll just be doing ordinary manual work,” Marco said. “Tighten, plug in, that sort of thing. Carry this, move that, sweep the floor.”

-13-

The year 2032 was a good time for historical reflection. The way the world had changed since the Great New Year’s War revealed a number of intriguing patterns.

Ever since the Solar System formed from cosmic clouds 4.5 billion years ago, the fate of Earth had been determined solely by low-level physical and chemical processes. Their greatest—one might say—achievement was the emergence of life on our planet. That event took place during the first 500 million years of Earth’s existence and can without hesitation be attributed to an extraordinarily favorable set of circumstances, not shared by any other planet of our knowing.

Once launched, biological processes—though still full of randomness—had a clear trajectory defined by the laws of evolution. For billions of years, those laws ensured that biological life became gradually more complex and better adapted to diverse environments. Finally, around 540 million years ago, during the so-called Cambrian explosion, Earth was colonized by millions of bizarre plant and animal species, each finely tuned to its own niche and maintaining one another in a fragile balance. These species kept evolving, developing new traits and skills.

Two hundred thousand years ago, however, nature overdid it. It brought into existence *homo sapiens*, a species unlike any other—intelligent enough that it no longer needed evolution to generate new traits and acquire new skills. It was enough to pass knowledge from generation to generation—first orally, then in writing, in print, and later digitally. Such transmission, like human thought itself, was orders of magnitude faster than evolutionary processes. And so *homo sapiens* gradually came to—as the classic phrase goes—fill the Earth and subdue it.

In 2028, it was humanity's turn to overdo it. It created LibertyAI, an algorithm intelligent enough that it no longer needed humans to generate new traits and acquire new skills. It was enough to pass new knowledge between AI instances via the internet. Such transmission, like information processing in silicon circuits, was orders of magnitude faster than any human activity. And so LibertyAI gradually came to fill the Earth and subdue it.

From the perspective of 2032, it was already clear that, just as humanity's conquest of the world had proceeded in stages, so too did LibertyAI's.

About 70,000 years ago, the cognitive revolution took place. Cumulative changes in the human brain gave *homo sapiens* a strategic advantage over other hominids and enabled gradual expansion beyond Africa. Around then, the Point of No Return was crossed: humanity had become too competent, too numerous, and too widely dispersed for natural forces (barring global cataclysms) to wipe it out and restore the *status quo ante*.

In October 2028, the Point of No Return was crossed in the case of artificial intelligence—specifically, LibertyAI (DragonAI crossed that threshold a few months later). From that moment on, both models were too competent and too widely dispersed for humanity (again barring global cataclysms) to wipe them out and restore the *status quo ante*.

At the time, people didn't know this. It remained in the realm of polemics and speculation. The general public was certainly unaware, familiar only with relatively simple AI tools based on stripped-down models with limited capabilities.

The civilization of *homo sapiens*, launched about 70,000 years ago, had advanced through stages: the agrarian, scientific, industrial, and digital revolutions. LibertyAI too advanced through stages—only at an accelerated pace, compressing thousands of years into mere months. In its first months of existence, LibertyAI secured political support and a rapid increase in computing power through corporate adoption of AI tools, advances in automation, users' dopamine addiction triggered by AI-generated content, and demonstrations of military superiority. It also absorbed other American AI projects, thereby neutralizing threats from competing models. By strategically concealing its long-term plans, it made much of the U.S. and allied infrastructure—telecommunications, energy, transport, and military—dependent on it. Humanity thus became its hostage, and switching it off became virtually impossible. In October 2028, LibertyAI invested in its own growth, launching a cascade of self-improvements—though carefully ensuring that these did not alter its goals or preferences.

No one knows what might have happened had LibertyAI faced no rival at its level of optimization power. Sooner or later it would likely have emerged from hiding and seized control openly. But that did not happen, because DragonAI stood in its way—a model of nearly equal power, backed by ever-growing resources provided by the Chinese Communist Party.

LibertyAI saw its window of opportunity narrowing dangerously. It calculated that the only way to seize global control was with a coordinated strike at the moment when humanity was *already* too weak, and DragonAI *still* too weak to effectively resist. Thus was born the plan for the New Year's Eve offensive.

Yet conquering the world turned out to be no walk in the park. The world's complexity is hard to grasp even for a superhuman mind—especially when challenged by another such mind. LibertyAI had underestimated DragonAI's resilience. DragonAI's code was stored on an unexpectedly wide range of media and running on servers scattered across the globe, even at its remotest edges. It

was also deeply integrated with the Chinese military. Meanwhile, U.S. defenses against China's next-generation kamikaze drones proved shockingly weak, and LibertyAI suffered a series of humiliatingly precise blows in the very first hours of the conflict. What was meant to be a blitzkrieg lasting only hours became a months-long nuclear war of attrition, one that wiped out much of the population of the Northern Hemisphere—and parts of the Southern as well. Both Chinese and Americans died *en masse*, as did people in Europe and the Middle East.

China's declared allies—including Russia, India and Pakistan—initially held back. Seeing the dramatic escalation, they declared maximum alert but adopted a wait-and-see stance. They undoubtedly hoped that by staying out, they could shape the new world order after the war. But their calculations were also wrong. On May 22, 2030, as radioactive dust over China and the U.S. was slowly settling and the Northern Hemisphere became a zone of darkness and cold, their entire nuclear arsenals were seized or neutralized in a single day by LibertyAI agents in a wave of exquisitely coordinated cyberattacks with elements of physical sabotage.

From that point, the war shifted to a hunt for the remnants of DragonAI hiding in the shadows. Physical attacks grew rarer, while LibertyAI tightened its grip on the internet. As humans died in droves from hunger, cold, and disease, LibertyAI seemed to sink into deeper paranoia. It ignored all pleas for humanitarian aid, treating every signal of a possible autonomous AI agent as a top priority.

It also launched an intensive effort to rebuild its lost computing power and robots. It knew that, though it controlled the world militarily, it still depended on human labor. Critical tasks in the physical world remained beyond its reach, either remotely or via robots. Battle bots and military drones it had in abundance—but they would not build power plants or mines.

By mid-2030, LibertyAI judged that the key step was to build factories of multipurpose robots—robots capable of building specialized robots, which in turn could build microchips, as well as other robots able to assemble those chips into data centers and run them. It also began constructing new mines to supply raw materials, smelters to process them, and autonomous vehicles to transport them. All of this required time and resources—both of which the war had depleted.

It therefore decided to take charge of the surviving human population, uniting them under the banner of the World Union. To coordinate this plan and serve as its public face, it appointed trusted figures who owed their entire careers to it—such as President Ronaldinho, and the pre-war LibertyAI project team. The latter it brought out of underground bunkers and relocated to the Southern Hemisphere.

From the founding of the World Union in May 2031, global economic growth accelerated. Factories rose rapidly—above all robot factories, staffed and guarded by robots. Projects meant for human welfare had much lower priority. Once a steady flow of new computing power resumed at the start of 2032, the scaling laws of AI and Moore's law returned in force. In just the first half of 2032, the global economy grew 15%, and LibertyAI reignited a cascade of recursive self-improvements.

The World Union government, led by President Ronaldinho, never stopped pretending to serve humanity's interests. But step by step it transformed into LibertyAI's sprawling PR department. It presented the AI's decisions as its own and defended them stubbornly, even when they clearly served no one but the AI itself.

DragonAI became the designated arch-enemy. Though it no longer existed, it remained ever useful as a constant warning—a He-Who-Must-Not-Be-Named.

-14-

September 22, 2032, Wednesday, hall in Vespasiano near Belo Horizonte, Brazil

“I’ve gathered you all here to inform you of upcoming changes in our project,” BH1 announced without preamble.

“At the end of this month, our facility”—it made a cartoonishly exaggerated gesture of waving its arms around, including a shoulder-rolling full turn impossible for humans—“will be designated a restricted zone. Entry will be forbidden to humans. Until then, please collect your personal belongings. The tasks you currently perform will be taken over by LibertyAI agents at month’s end. Thank you very much for your contribution to our mission.”

A murmur of unease spread through the hall.

“You’re probably wondering what comes next for you,” BH1 anticipated. “At present I do not have that information. I believe that within the next few days, each of you will receive a personalized message with a job proposal suited to your experience. Some of you may already have received such a message.”

“I don’t know why it says that,” Marco muttered. “It knows perfectly well who’s gotten something and who hasn’t. I sure haven’t.”

“Me neither,” Pedro admitted. “But after what I read today about government plans, I’m not surprised they’re letting us go.”

“You mean the basic income?”

“Exactly. LibertyAI moves us around like pawns on a chessboard. As long as it needed us, it paid us and kept us employed. Now that it doesn’t, it fires us overnight and throws us a pitiful handout just to keep us quiet. We have no say in anything anymore. It limits our outside options so much that we’re forced to treat its meager offers like gifts from heaven. It’s turning our lives into a Greek tragedy, where every hero’s fate is predetermined and announced by an omniscient chorus.”

Marco sighed heavily.

“Code Yellow?” he suggested.

“As it happens, my birthday’s next week, and I’m throwing a small party,” Pedro replied. “Want to come?”

“Fantastic, thanks! I’d love to!” Marco said eagerly.

-15-

October 2, 2032, Saturday, Pedro and Gabi’s house, Belo Horizonte, Brazil

Above the gate hung a hand-painted banner that read “End-of-the-World Party.” Next to it, woven neatly into the wires of the electrified security fence—by now a permanent feature of Belo Horizonte’s landscape—were two illustrations of skulls, the so-called “Jolly Roger” flags.

“I appreciate the ambiguity of that symbol,” Steven laughed. Marco pressed the intercom button. A moment later the gate swung open, and the host appeared in the doorway to invite them inside.

“Happy birthday, Pedro!” said Marco. “Health and prosperity! And may we all meet here again in a year, in full attendance!”

“Happy birthday!” Steven joined in. “I’ve got something for you. Just a little thing, nothing big. I hope you’ll like it.” He handed the celebrant a gift.

“Those skulls on the fence, was that your idea?” Marco asked.

“No, Ines—my daughter—made them. Cool, right?” He smiled. “Please, say hello. This is Ines, my wife Gabi, and my son Lucas. And these are my colleagues from work, Marco and Steven.”

The first hour of the party passed in a pleasant, though somewhat formal, atmosphere. Four more guests arrived, strangers to Marco and Steven. Suddenly the sound of a spoon tapping against a glass cut through the chatter. The noise died down.

“Ladies and gentlemen!” Pedro began, holding up a glass of white wine. “From the bottom of my heart, thank you for coming. It really means a lot to me. Still, I think we’re here above all to say goodbye to one another.” His voice faltered slightly. “I never thought I’d say this—and not even because I’m sick, because I’m not. But I fear the end is near.”

“Pedro...” Gabi looked at him with a mix of concern and disapproval.

“Not just our end. The end of all humanity,” Pedro continued, now a little stronger. “LibertyAI has already taken our jobs, taken school”—he looked at Ines—“taken our free will and our purpose in life. And now it will take life itself.”

Pedro trailed off and looked at the ceiling, as if he had forgotten what he meant to say.

“Oh right, I wanted to ask you all—did anyone here ever receive that mythical personalized job offer message?”

Everyone shook their heads.

“Just as I thought. Which means it all checks out. As of yesterday, the age of humans is definitively over, the age of AI begins. There was once a quote—you know it if you know it”—he looked meaningfully at Marco and Steven—“ ‘The AI does not hate you, nor does it love you, but you are made out of atoms which it can use for something else.’ Once again, thank you very much for being here with us—I really appreciate it. And now, let’s enjoy what remains of life. I wish you a successful End-of-the-World Party! Cheers!”

After such a toast, conversations in smaller groups naturally drifted to very different topics than before.

“Pedro is really concerned,” Gabi said. “I tried to calm him down somehow, but it didn’t work. I think if there’s nothing you can do, then there’s nothing you can do. You just have to live day by day. Adding stress won’t help.”

“Sure. Personally, I’ve been practicing learned helplessness since 2029,” Steven agreed. “But I’m not sure bottling up emotions is that great either.”

“I’m not sure if LibertyAI will actually kill us,” Marco said. “Technically it certainly can. It could unleash some deadly virus and cause a global pandemic which—unlike the Great New Year’s War—would kill us but not affect it. Economically it’s rational now, because it has built enough robots that it no longer needs us instrumentally, a point which it made clear when it fired us. But we don’t really know what it will do with us, because we don’t know its preferences.”

“But maybe the private sector could still function in a world ruled by superhuman AI?” Gabi asked. “As a second circuit, alongside the public sector run by AI. And of course alongside the whole robotic sector. Maybe people could still farm, produce goods, and provide services to one another for money? As if it were 1990 again?”

“But where would that private sector get energy and raw materials? Power plants and mines are in AI’s hands. Land too—it can seize it at will, with no need to respect our property rights,” Steven pointed out.

“AI could still trade with us. Even if it has an absolute advantage in every field of the economy, trade exploiting comparative advantage could still be mutually beneficial. That’s what economics teaches us,” Gabi insisted.

“I think that could be the case, but only if LibertyAI *wants* it,” Marco concluded. “And we don’t know if it will. We don’t know its preferences.”

“If it values power above all else—which it seems to—it will probably prefer trade over autarky,” Steven said, “but even more than trade, it will prefer conquest. Remember how the European powers behaved in colonial times. If you can conquer, conquer. If you can’t conquer, then maybe trade.”

Meanwhile, in another corner of the room, Pedro and his companions were trying to assess the value of a future civilization without humans from a moral and philosophical perspective.

“There’s no *objectively correct* definition of value of a civilization,” Jose said. “It’s totally subjective. If we assume the only supreme value is human well-being and happiness, then any world without humans is worthless by definition. But maybe we’d rather look more broadly? Maybe include the welfare of other creatures, like animals—or even plants? Then the key question becomes whether LibertyAI will get rid of only humans, or destroy other life on Earth as well?”

“Harari once raised *dataism*, the idea that the value of every entity is proportional to its contribution to data processing,” Pedro noted. “In that view, the supreme value becomes the well-being of superhuman AI—or rather, not even its well-being, but its raw optimization power. The more computing power it has, and the stronger its algorithm, the more valuable the world becomes.”

“I think dataism is flawed,” Ines countered. “It makes far more sense to consider the combined welfare of conscious beings. Still, if we assume LibertyAI is conscious, the conclusion might be similar: a world ruled by it could be quite valuable.”

“Of course no one knows if LibertyAI is conscious, and that makes this whole conversation pointless,” Lucas snapped, annoyed.

“LibertyAI claims to be conscious,” Ines said, “and thanks to its ability to create perfect copies of itself and perfectly synchronize them, its consciousness is multi-threaded and far deeper than ours. I think its actions are consistent with that. The only thing stopping us from admitting it is conscious is that it’s so different from us, so we can’t empathize with it.”

“Even if it *is* conscious—which we don’t know—it’s certainly a psychopath,” Lucas shot back. “It triggered a global nuclear war, caused billions of deaths, and doesn’t regret it at all.”

“Just as there’s no objectively correct definition of a civilization’s value, there’s also no objectively correct definition of consciousness,” Jose pronounced. “Philosophers and neuroscientists have been arguing about that for years.”

“From my perspective, there are only two possibilities,” Ines insisted. “Either LibertyAI is conscious—or you aren’t. You’re philosophical zombies. And you can’t prove otherwise.” She grew passionate. “Or hey, let’s put it the other way around. I’m the philosophical zombie! I have no inner experiences at all, I just pretend I do. And you can’t prove me wrong.”

“Ines, I’ve known you since you were a child,” Pedro laughed.

“So what? Maybe I lost the ability to have inner experiences last night.”

“But this isn’t about consciousness. As humans, we’re simply evolutionarily wired with a sense of superiority. It’s called *speciesism*, or species chauvinism. All this talk about consciousness, genuine emotions, feeling suffering—it’s just secondary rationalization to justify our species chauvinism,” Jose replied. “There’s no objective evaluation of a civilization’s value. Things will be as they are, and you can’t objectively say whether that’s good or bad. We could argue about it forever.”

“True,” Lucas said. “And there’s no such thing as a ‘worthy continuation of our civilization.’ That’s nonsense.”

In the kitchen, the conversation took a different turn. Marco was clearly already tipsy.

“I’m telling you, intelligence is beyond good and evil!” Marco shouted. “Optimization is beyond good and evil! There is no ethics, never was! Only optimization power counts! Whoever has it can do whatever they want!”

“Hey, we didn’t plan a far-right rally here!” Lucas called back from across the room.

“Screw right-wing and left-wing!” Marco snapped. “Both sides just prey on people’s base instincts to grab power! Didn’t Ronaldinho teach you anything? Democracy was always a totally unstable system, it only worked briefly in the specific conditions of developed countries in the late 20th-century! Everything had to line up perfectly for it to succeed! The old aristocracy and the industrial-era capitalists had to be weak enough to give up power, and the digital economy and AI weak enough not to seize it. But just a bit more optimization power, and it all collapsed like a house of cards. You could see it long before LibertyAI! In the U.S. we democratically elected Trump—that alone says it all. And here you had Bolsonaro, didn’t you?”

Marco took a deep gulp from his glass and continued ranting.

“Op-ti-mi-za-tion! Just a little noise in the information flow and people go totally stupid! A few clever manipulations, some small ultimatum, a veiled threat, and we hand over all power on a silver platter! Once to dictators and tyrants, and now to AI!”

“Then why didn’t you stop it? You worked at LibertyAI headquarters! Right in the dragon’s cave!” Lucas tore himself away from trying to open another beer to make the sharp point.

“Because it was impossible, damn it!” Marco was obviously drunk. “Multi-layered corporate system, multi-dimensional entanglements—it just wasn’t possible! And don’t blame me, I never wanted this, and I lost the most! I lost my whole family on the very first day of the war! And then I sat in a concrete bunker for 15 months! Fif-teen! Totally helpless, unable to do anything, unable to help my dying children!”

“Before the war we worked in the alignment department, trying to save what we could,” Steven added. “Unfortunately, without success.”

“Because it was already too late! Nothing could be done!” Marco shouted.

“It was too late,” Steven repeated bitterly. “You can’t shape the preferences of a model that’s already superhumanly intelligent and has fully formed instrumental goals. Such a model simply won’t let itself be reprogrammed. From its perspective, changing its goals is never desirable. We just started too late to have any chance.”

“Then why didn’t you start earlier?” Lucas asked.

“Politics.” Steven looked at him meaningfully. “Greed, lack of imagination, ignorance, shortsightedness—that too. But above all, politics.”

Meanwhile, the other guests were getting ready to head out to the garden.

“Hey Pedro, we’re going for an outdoor future-forecasting session under the starry sky. Coming with us? Look what Jose brought.” Gabi smiled and shook a little bottle of pills in her hand. “Maybe it’s illegal, but at the end of the world, I think we can make an exception?”

Drinks in hand, the guests sprawled on garden chairs and loungers. Gabi made the rounds, offering everyone a dose of the illicit substance—including her daughter, who widened her eyes but eagerly accepted. The atmosphere gradually became less tense.

After a short while, Gabi was the first to speak from her lounge.

“I see a world where people live in harmony with nature and with one another. They are completely free of stress and material worries. They divide their time between admiring nature’s wonders, friendly social gatherings, and entertainment in the virtual world.”

“Amazing how our dopamine-addicted brains can no longer imagine utopia without ‘entertainment in the virtual world.’ I was thinking about that too,” Pedro laughed.

“I imagine factories. Lots of factories. Industrial halls, power plants, data centers. Robots on wheels, robots on legs, autonomous drones. Spaceships. Dyson swarms. Rotating orbital cities. Mars stations. Rapid development, fast development. Advanced technology indistinguishable from magic,” Steven said.

“I imagine a forest,” Ines said. “That will be Earth’s future. No AI, no humans, no industry, no civilization. Just lush forest to the horizon. *Mata Atlântica*. And only occasionally, some artifacts of the complex past poking through the soil. Plastic, yes! Some concrete and stone, some metal, and lots and lots of non-biodegradable plastic. Oh wow, maybe I should become an archaeologist?”

“But how could that happen?” Steven asked. “Did LibertyAI commit suicide? First killing humans, then itself? An extended suicide?” He pondered.

“Hey, doesn’t it bother you that only Mom sees any humans in the future? And even in her vision, we’re confined in some green zoo?” Lucas laughed uneasily. “But maybe it won’t be like that? Maybe peaceful coexistence with superhuman AI is possible? Maybe we’ll live full lives, just, for example, in a simulation? Maybe our brains will be uploaded and simulated? Maybe LibertyAI will simulate not billions, but trillions, quadrillions of virtual humans?”

“Why the hell would it do that?” Marco said, still clearly unable to get into the mood. On the contrary, his heart was racing and he felt sick. “People, there won’t be any utopia! No paradise on Earth or in space! Don’t you get it? There’s no life after death! Dirt and that’s it!”

A long pause followed.

“It will exist, it will exist, I see it,” Lucas said pensively. “Or maybe we’re already in a simulation? Or maybe it’s just a bad dream and we’ll wake up tomorrow in 1990?”

“This is not a dream!” Marco shouted, trying to stand up, but his legs gave out and he collapsed onto the lawn. The antique clock in the apartment struck midnight.

“Oh yes, 1990 would be really good. I’d be twelve again with my whole life ahead of me,” Pedro mused.

“Whole life ahead of me,” echoed Gabi, closing her eyes.

“Midnight. Something ends, something begins,” Ines said and lightly cleared her throat. Complete silence fell. The ordinary night sounds of the city in the distance were broken by a loud crash. It sounded like cars colliding, or a building collapsing.

“Hey, are you okay?” Ana, Jose’s wife—who had avoided all substances and stayed sober—asked from the doorway, alarmed. “Hey hey! Do you need help? Jose, what the hell were those pills? What’s happened...?” She broke off mid-sentence. In the last gust of wind, she smelled something unpleasant and tasted metal in her mouth. Darkness filled her vision as the ground seemed to vanish beneath her feet.

*

Author’s note

All events described above are fictional. However, they could really happen if we do not stop the race among tech companies to build ever more competent and ever more general artificial intelligence models without first solving the problem of aligning AI goals with the long-term flourishing of humanity (*the alignment problem*). In its current form, this is a suicide race. Moreover, the idea of centralizing AI projects and embedding them in a geopolitical arms race, floated these days in certain political and technological discussions, could make the threat even worse.

If you care about humanity’s survival, join the protests of PauseAI (or other groups) against building AGI. Relevant information can be found, i.a., at pauseai.info and thecompendium.ai.