2

Burnt-Out Images: A Thermodynamical Turn

When someone believes they are some kind of messiah, they are said to have fallen victim to Jerusalem syndrome. In tech, however, self-declared messiahs display a *Prometheus syndrome* – after the mythical figure who stole the fire from gods.

One example is Emad Mostaque, the former CEO of AI image generator Stable Diffusion, who describes his tool in precisely such terms: as 'bringing fire from the gods of creativity to the world'.[1] But Mostaque is not the only tech executive to invoke the fire-poaching hero. Google CEO Sundar Pichai, too, has directly compared artificial intelligence to fire: 'AI is one of the most important things humanity is working on. It is more profound than, I dunno, electricity or fire,' said Pichai at a town hall event in San Francisco in January 2018.[2] Similarly, Sam Altman, CEO of OpenAI, has stated that AI is the most important new technology since the taming of fire and the wheel.[3] And, in 2022, Microsoft released a fork of OpenAI's GPT model called 'Prometheus', integrated into their search engine Bing and parts of their software. The use of Promethean vocabulary by this crowd is not surprising, as this mythological figure was one of Ayn Rand's favourite mythological figures, celebrating individual entrepreneurialism.

Fire thus becomes a symbol for so-called AI technologies and, in extension, digital-image generation based on it.

Fire as Technology

The relationship between fire and technology is ancient. Likely the most important tool of humankind, fire may even have altered the human body and brain: cooking food probably brought modern humans into being by modifying their brains and digestive tract. Fire may also have contributed to the rise of language, the division of labour and development of social forms.

This technology was initially associated with what came to be known, much later, as domestic labour: as hearth, candle, spark, combustion and electrical current. It helped humans to survive in colder environments and thus to migrate into different regions. It assisted in the shaping of landscape and the creation of paths by burning down vegetation or making the interiors of caves accessible, and domesticating space to some extent. It made night-time available for human activities. It made the creation of paintings in dark and remote cave passages possible. As such, it has dominated human attempts to control their surroundings.

Philosopher Bernard Stiegler describes the way humans started to modify their environment via steam engines and combustion motors using equally Promethean metaphors: 'The Anthropocene epoch can appear ... only starting from the moment when the question of the cosmos is itself grasped as that of combustion, in both astrophysics and thermodynamics – but, therefore, also in relation to this exceptional *pharmakon* that is domesticated fire, fire as that artifice par excellence delivered to mortals by Prometheus.'[4]

Medium Hot and Cool

When it comes to AI information processing, the invocation of fire is not only metaphorical; it is also one of its conceptual frameworks. Emad Mostaque's flagship image generator is

called Stable Diffusion for a reason: it is a so-called diffusion probabilistic model,[5] which has its origins in the heat equation developed by nineteenth-century physicist Joseph Fourier.[6] This formula describes how heat spreads through a medium, a dispersal which becomes an important template for some parts of machine learning–based image-making. In this thermodynamical paradigm, a detailed, intelligible image is seen as being 'cool', then is diffused into 'hotter', less organised articulations, and ultimately into random noise. Image generation proceeds by reversing this process of entropy and recovering, or 'restoring', an image from noise.

The move towards a thermodynamical paradigm, however, has many unexpected consequences. Its stochastic turn orients images away from indexical representation, towards prediction and approximation via averaged renderings based on giant datasets. The thermodynamical paradigm links physics with information theory and increasingly encroaches on image generation. In information theory, as media scholar Nicole Starosielski observes, 'the idea that signal exchange is equivalent to heat exchange can be traced back directly to Claude Shannon's theory of information, which advances thermodynamics-inspired concepts such as entropy in the description of communications systems'.[7] Other areas inspired by thermodynamical paradigms include statistics, finance and climate science.

This new environment for image production presents a decisive break with a photographical or, more generally, an optical paradigm of image production based on causality that dominated much of twentieth-century thinking about media. In principle, within traditional optics, the route of every photon could be ray-traced back to its origin. In a thermodynamical environment, however, causality is replaced by probability, and ray-casting by forecasting.

The optical paradigm was associated (partly via photography and cinema) with Newtonian physics, the laws of reflection

and refraction, shooting (in all senses of the word), factory production and the military (by way of camera technologies relating to film transport and emulsion production). In this context, a ray of light can be stopped or interfered with during transmission, thus leading to an understanding of noise as disruption or scrambling of signal.

Burnt-Out Images

In a thermodynamical environment, however, noise becomes an indication of increasing disorder – but also a raw material for image production.

Diffusion probabilistic models (formulated as 'Markov chains') first add random noise to the available training data in a so-called forward diffusion process, then reverse the process to recover the data (fig 2.1).[8] An image is drowned in noise, then the noise is removed and the process remembered. Next, the removal process is applied to random noise samples, thus seemingly reversing entropy and disorder.

As one researcher documents:

> Diffusion processes are continuous-time, continuous-state processes whose sample paths are everywhere continuous but nowhere differentiable. They arise in many natural phenomena and in some sense are generalized versions of the Brownian motion.
>
> ... In the fields where diffusion has been applied, it has been used to model phenomena evolving randomly and continuously

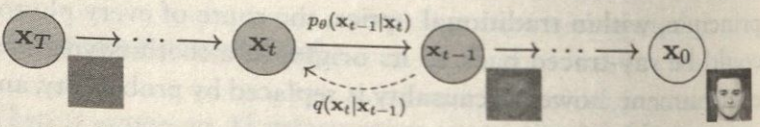

Figure 2.1. A graphic rendition of the diffusion denoising process (Jonathan Ho, Ajay Jain, and Pieter Abbeel, 'Denoising Diffusion Probabilistic Models', arXiv, 16 December 2020).

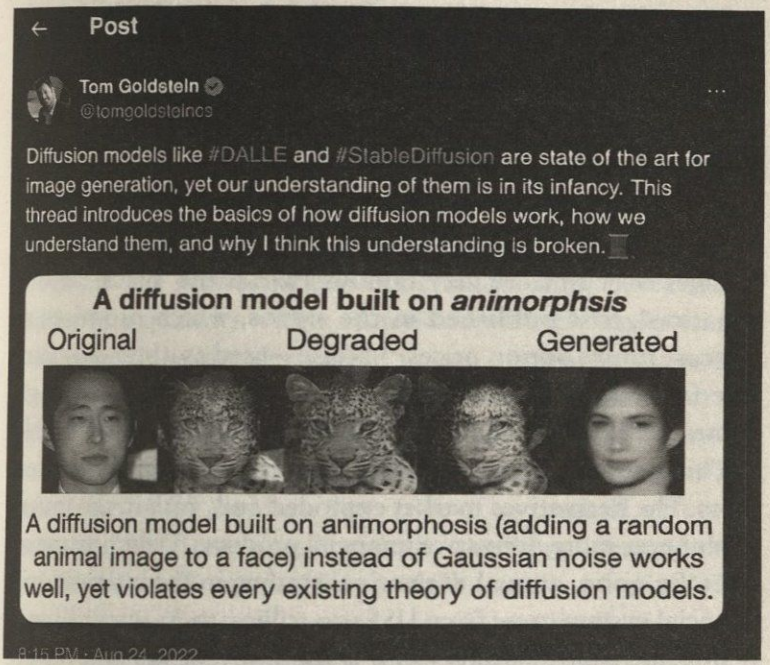

Figure 2.2. Exchanging noise with a JPEG of a leopard seems to work, too (Tom Goldstein [@tomgoldsteincs], 'Diffusion models like #DALLE and #StableDiffusion are state of the art', X, 24 August 2022, 8:15 p.m.)

> in time under certain conditions, for example, security price fluctuations in a perfect market, variations of population growth in ideal conditions, and communication systems with noise.[9]

At the same time, it remains patently unclear how, exactly, diffusion models work and whether random noise is really a key factor in AI-image generation. One researcher exchanged Brownian noise with a picture of a leopard – and still managed to generate plausible imagery (fig 2.2).[10]

In the absence of precise explanations, it is, however, notable that most scientists definitely rely on thermodynamical terminology to describe what image-generation models are supposedly doing – perhaps simply to align their rhetoric with adjacent fields such as finance and information theory.

From the Camera Obscura to the Camera Casino

Image generators are not the only operation that use heat diffusion, in part, to describe their operation. The pricing of stock options is predicted using a similar body of mathematics.

Fourier's heat equation is a so-called partial differential equation (PDE), which uses random variables to model unknown changes over time. A very famous PDE is the 'Black-Scholes equation', first published in the 1970s, which promises to forecast future option prices.[11] Noise – used as a stand-in variable for random processes – plays a central role in seemingly domesticating chance and making it stochastically predictable.

This promise spawned a huge industry within just a few years. The derivatives market exploded and, with it, the invention of ever more derivative financial products. Debt became an asset. By 2007, annual derivatives trading in the international financial system grew from US$100 trillion to $1 quadrillion.[12]

In time, many bankers and traders forgot that they were relying on a form of wishful thinking; they used the equation as a kind of talisman, a 'bit of mathematical magic'. It worked because everybody believed in it; in reality, however, it failed to predict the actual likeliness of extreme adverse events in a situation already destabilised by aggressive risk management. With the banking crisis of 2008–09, a house of cards of financial alchemy and magical thinking collapsed.

It is unknown if Black-Scholes would have produced the same outcome using the JPEG of a leopard as a random variable instead of noise; perhaps this would even have increased its magical aura as a key tool of predatory capitalism.

Derivative Images

Sociologist Donald MacKenzie, in his seminal investigation of Black-Scholes entitled *An Engine, Not a Camera*, makes

the case that such PDEs are not analytical renditions of any situation but, rather, active participants in co-creating it. He refers to the formula as an engine of inquiry, a motor. A similar motor is now embedded into machine learning–based image generators – *an engine as a camera*, so to speak.[13]

Prior to the meteoric rise of generative AI, media scholar Jonathan Beller defined derivative images as programmable images in the service of capital, transforming every social quality into quantity within racialised, extractive industries.[14]

Additionally, the derivative image is now also *literally derivative*: it is derived from billions of other images in the training data. According to copyright law, 'derivative' means work that includes elements from a copyrighted work. The underlying image may be a derivative of an original, or even a derivative of a derivative. Practically speaking, creating these derivatives works to avoid copyright issues and enable the large-scale theft of data.

The crash of this system of value in relation to images is already happening. This has led to a widespread disbelief in any truth value of any image or information – whether AI-generated or not – with wide-ranging consequences for public discourse (or whatever remains of it).

Making Disorder Profitable

In this kind of value system, risk and disorder themselves become assets. New financial derivatives try to put a price on climate risk in order to transform it into a set of opportunities – aiming to appraise the value of climate change itself and turn it into a speculative market. This mechanism relies on the idea of markets acting as superior information processors that 'sense' reality via pricing. As the philosopher of technology William R. Morgan observes:

> For ESG [environmental, social and governance] finance, the climate catastrophe is not, as a matter of necessity, something to be avoided. It is rather a source of expected risk, the existence of which demands the invention of products whose role it is to transfer risk from those who do not want exposure to it to those that do. This is the power of ESG finance, not to blunt the risk of disaster, but to facilitate the pricing of catastrophic risk so as to enable its being bought and sold.

Through these mechanisms, the future heat or climate chaos itself may become a market full of opportunities for profit, but also a system of sensors to capture reality via pricing. As Morgan writes, 'Derivatives such as these appear to offer absolute certainty in reply to seemingly absurd questions. What is the price of a hurricane in sovereign debt markets? What about several hurricanes? ... What is the cost of exposure to climate risk?'[15]

Let us add to this series of questions: What is the price of a riot or rebellion? What is the price of entropy? How does it become a valuable source of disorder to be gambled on or exploited as raw material?

In terms of image-making, markets are becoming 'cameras' – or, more generally, data processors – environmental sensors that render reality as a pricing system.

Stochastic Point Clouds

Thus, image generation moves even further away from any indexical contact with reality, from observation and exposure to actual, unaveraged situations. To be clear, image generation does not have anything to do with reality per se. Images can be sheer fictions, dreams or hallucinations – basically anything at all. Nevertheless, reality and observation, and the numerous questions arising from such notions, were important factors for consideration within an optical/indexical paradigm

of representation. After all, fidelity to observable reality was based on ray-tracing the trajectories of actual photons/light rays back to their origin. This part is now filtered through statistical procedures.

In a stochastic paradigm, contact with any external reality is heavily mediated through the averaging of large-scale datasets. This statistical realism is based not on observation but on quantity – on the idea that scale, in itself, will capture a statistically relevant sample: if one just captures 'everything', a real underlying pattern will end up being captured too. Such thinking not only confuses the internet with reality but also privileges quantity over actual observation and analysis. Most importantly, the unexpected and contingent cannot enter these calculations by default.

In contrast to this, the paradigm of documentary photographic practices allowed for the possibility of unexpected things to happen in front of the camera, to expand the realm of what was supposed to be visible and knowable – to collect new primary-source data, so to speak.

The image matter of stochastic renderings are huge data clouds, as well as quasi-astrological attempts to interpret and identify inherent features and patterns. One underlying assumption is that it is both unnecessary and infeasible to record individual events, and that their statistic relevance is minor. But this also erases the possibility of ever capturing something which is not yet already known and included into the dataset. The dataset becomes something like a VR headset that replaces actual vision with heavily redacted data visualisations.

Machine Dementia

The process of conjuring up images from granulated data could easily turn into a feedback loop. People generate data, which is

then used to train AI models. In a next step, AI models generate more data. Then, these synthetic data are also used as training data to train more AI models. This, in turn, creates so-called model collapse, which one computer scientist refers to as 'a degenerative learning process where models start forgetting improbable events over time as the model becomes poisoned with its own projection of reality'.[16]

When undergoing model collapse, machines start to stammer and repeat elements from the range of most likely predictions, which sound rather demented to human beings (table 2.1). Total probability – the ultimate norm – is a kind of machine dementia.

Conditioned to be as 'normal' as possible, the model thus starts sounding psychotic. As journalist Alfonso Maruccia puts it, 'Like a JPEG image recompressed too many times, the internet of the AI-driven future is seemingly destined to turn into a giant pile of worthless digital white noise.'[17]

The problem is not just the production of noise and nonsense; it is an overall acceleration of entropy, in the sense of breaking down meaning and matter, and of warming the planet: 'Just as we've strewn the oceans with plastic trash and filled the atmosphere with carbon dioxide', Maruccia writes, 'so we're about to fill the Internet with blah.'[18]

e/acc

Some tech ideologists appreciate accelerating entropy. They advocate the development of artificial general intelligence (AGI) to increase disorder by consuming an enormous amount of energy. Why would this make any sense? Because, according to Guillaume Verdon – a physicist, self-declared thermodynamical priest and evangelist of 'effective accelerationism' (often abbreviated 'e/acc') who tweets under the moniker of @BasedBeffJezos – the second law of thermodynamics proves

Table 2.1. Model collapse

Example of text outputs of an OPT-125m model affected by *Model Collapse* – models degrade over generations, where each new generation is trained on data produced by the previous generation.

Input:	some started before 1360 – was typically accomplished by a master mason and a small team of itinerant masons, supplemented by local parish labourers, according to Poyntz Wright. But other authors reject this model, suggesting instead that leading architects designed the parish church towers based on early examples of Perpendicular
Outputs:	
Gen 0:	Revival architecture such as St. John's Cathedral in London. The earliest surviving example of Perpendicular Revival Architecture is found in the 18th @-@ century Church of Our Lady of Guernsey, which dates from the late 19th century. There are two types of perpendicular churches: those
Gen 1:	architecture such as St. Peter's Basilica in Rome or St. Peter's Basilica in Buenos Aires. There is no evidence that any of these buildings were built during the reign of Pope Innocent III, but it is possible that they may have been built during the reign of his successor, Pope Innocent
Gen 7:	architecture in England. In an interview with The New York Times, Wright said: "I don't think there is anything wrong with me being able to do what I want to do. It just doesn't work for me. He added: "I don't know if you can call it funny,
Gen 9:	architecture. In addition to being home to some of the world's largest populations of black @-@ tailed jackrabbits, white @-@ tailed jackrabbits, blue @-@ tailed jackrabbits, red @-@ tailed jackrabbits, yellow @-

Source: Ilia Shumailov, 'The Curse of Recursion: Training on Generated Data Makes Models Forget', arXiv, 14 April 2024.

that the universe itself aims to increase entropy. This is apparently the 'will of the universe', and 'the universe exponentially favors (in terms of probability of existence/occurrence) futures where matter has adapted itself to capture more free energy and convert it to more entropy'.[19]

Verdon expands on a very interesting, but still speculative theory by scientist Jeremy England, who claims that life arose

as a form of spontaneous self-organisation of matter in relation to external energy sources. According to England, some systems are able to dissipate more energy than others, which gives them an evolutionary advantage.[20] While this theory remains to be validated and is restricted to the molecular level, Verdon applies it to capitalism as an intelligent 'meta-organism'. Because capital is able to consume/diffuse a lot of energy, he muses, it must be a higher form of being.

Verdon calls for humanity to 'climb the Kardashev scale' to convert even more free energy into unavailable energy and bring the universe closer to heat death.[21] The Kardashev scale is a ranking of civilisations according to how much energy they are able to use. A type 1 civilisation uses all the energy on its planet; a type 2 civilisation harnesses all the energy produced by its nearest star; and a type 3 civilisation captures all the energy of its galaxy. The more energy a civilisation captures and uses, the more in line it appears with the presumed 'will of the universe'. Thus, according to Verdon, the development of technology should be accelerated to ascend the Kardashev scale by building AGI.[22]

According to the e/acc ideology, AGI's main purpose is not to solve computational or scientific problems but to burn through tons of energy, thus increasing entropy and accelerating heat death. Burning becomes an end in itself. This, however, comes at a massive cost, requiring massive energy resources and an industrial-scale infrastructure that consumes power and increases overall disorder, including social disruption.

As Bernard Stiegler notes in his observations on the Anthropocene, 'The thermodynamic machine, which raises, in *physics*, the new and specific problem of the dissipation of energy, is also an industrial technical object that fundamentally disrupts *social* organizations.'[23] The thermodynamical imaging infrastructure is disruptive in that it consists of a massive system of burning, selling and profiting from fossil fuels, wasting resources, heating the atmosphere and, as a consequence,

causing the devastation of the environmental and social fabric. Some cultists even advocate accelerating this process – probably after placing some financial bets on it. Where there is the fire of combustion and heat, smoke, ash and embers are not far. So are burnout and an acceleration of disordering, entropic processes. AI image generation is thus embedded in what one could call a larger *digital political economy of entropy*.

In her influential book *Media Hot and Cold*, Nicole Starosielski describes parts of this system:

> The most powerful media organizations of the twenty-first century will be thermal. The circulation of images, sounds, videos, and texts will depend on a massive regime of heating and cooling. Data and networks, like the people they connect, will be ever more fragile. Too hot or too cold, and the platforms will collapse. Digital infrastructures – data centers, network exchanges, and fiberoptic cables – will drain the planet's energy in order to create a stable thermal environment – not for people but for information.[24]

Political Economy of Entropy

In May 2021, the Colonial Pipeline, an American oil pipeline system, suffered a ransomware cyberattack. Its operator, the Colonial Pipeline Company, stopped all operations to contain the attack. Within several hours, the company paid seventy-five Bitcoin, then worth $4.4 million.[25] After the ransom was delivered, hackers provided an IT tool to the company to restore the system. However, the attack had already brought about serious side effects. Airlines and airports were disrupted for lack of fuel. Panic-buying of gas caused a bump in prices and fuel shortages. Some people even filled plastic bags with gasoline.[26]

A few weeks prior, a new malware, called Prometheus, had been launched. Within a short time, it would infect thirty organisations, including banks, governments in North America

and Asia, and numerous companies such as Colonial Pipeline. A Prometheus attack basically takes a server hostage by encrypting all the files on it and demanding a ransom (which they call a 'timed ticket') for people to access their own data. If the victims (called 'customers') don't pay, their data are put up for auction. Payment was demanded in the cryptocurrency Monero, which was considered 'secure', or Bitcoin, at a higher cost due to its traceability. By doing so, the kidnappers effectively separated people from their data and demanded money from them to access their own files. A Prometheus attack thus combines several elements: theft of knowledge, seizure of infrastructure and establishment of a kind of monopoly access to it. If people want to access it, they have to pay.

One can see elements of such hostage-taking within digital industries as well. It seems likely that the Promethean projects advertised by Mostaque, Altman et al. in fact amount to a legal version of a Prometheus attack. After all, many data corporations control access to people's data. The producers of data have to pay to access their own data/projects through proprietary clouds and gateways. The difference is that such extortion is entirely legal, since those corporations own the infrastructure on which their customers/hostages rely. Contemporary Prometheus impersonators thus not only steal knowledge/fire but rent it back to humans (instead of simply handing it out as a gift). They train workers to become dependent on proprietary pipelines that ultimately heat atmospheres and destabilise environments.

Prometheus Gets Fired

In late March 2024, Stability AI founder Emad Mostaque took the stage for a PR announcement. An AI-generated Aristotle announced him as 'a modern Prometheus' in faux-Greek jargon: 'Under his stewardship, AI becomes the Herculean force

poised to vanquish the twin serpents of illness and ailment and extend the olive branch of longevity.'

'I think that's the best intro I've ever had,' Mostaque said.[27]

But simultaneously, Stability AI, the company that had made diffusion-based image generators famous, was in deep trouble. Its key researchers – the team that had initially created Stable Diffusion at two German universities – had resigned. Creditors were putting on pressure. The disclosure of abusive child pornographic imagery within the training data used for Stable Diffusion had barely been addressed. Stability's position on stealing copyrighted material was criticised by former personnel. Meanwhile, three copyright infringement lawsuits brought by Getty Images and a group of artists in the US and UK claimed Stability AI illegally used their art and photography to train the AI models powering Stable Diffusion.

The main reason for Stability AI's crisis, though, was acute financial instability. Stability AI was burning through money – in the form of compute, or the operation of graphics processing units (GPUs) chained together on Amazon's Web Services cloud. As one journalist reported, 'The costs associated with so much compute were now threatening to sink the company. According to an internal October financial forecast seen by Forbes, Stability was on track to spend $99 million on compute in 2023.'

The result? According to Mostaque, a new version of the Stable Diffusion image generator could now generate '200 cats with hats per second'.

In the end, Stable Diffusion would be described as 'a giant vacuum that absorbed everything: money, compute, people'.[28] In late March 2024, Mostaque resigned.