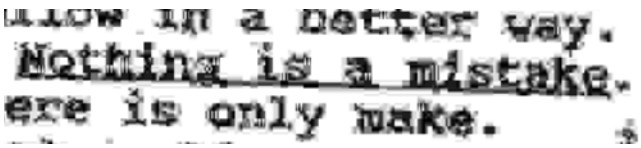
Kim Albrecht

Eternal, Data, Decay

Driving the human

> Ut nihil non iisdem verbis redderetur auditum.
> —Jorge Luis Borges, *Funes, the Memorious*[1]

The speculative scenario I am developing in this essay is that the desire for immortality is a key driver behind digital data technologies. The storage of computational information, i.e., data, is conceptualized as eternal; by itself, data does not decay. The hardware containing the data decays, but, due to lossless duplication, information can stay immortal. This feature of digitality has tremendous advantages and is a key to the success of information technology. But the notion of data as eternal creates both social and natural fallacies based on a simplistic model of knowledge as storage. Memory might be more than bytes on a silicon chip. Generalization, abstraction, and forgetting are key components of thinking. I am arguing that we will need a new concept, a new model of data for a sustainable and socially-minded collective future.



Lossy JPEG 01

Save for Later

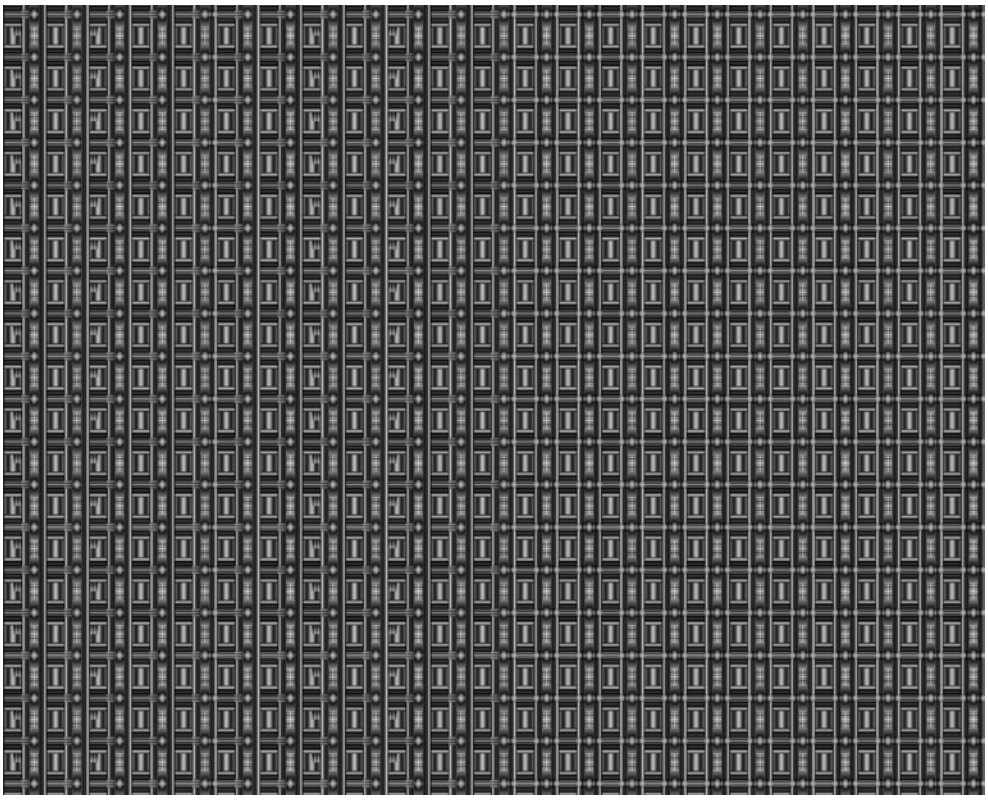
Visiting a concert in the 21st century often results in an uncanny experience: rather than enjoining the music, getting absorbed by it and transcending the moment, nowadays people stand as still as they can, holding up smartphones to record the performance. Our bodily experience is reduced to an unsteady tripod recording a performance onto a device that is in not in any way capable of capturing the experience to the fullest. And it is not only concerts but all kinds of events where such behavior is becoming the norm. The mode of "save for later" is becoming the standard, and I am wondering if anyone will watch those shaky videos with pour audio ever again.

It seems that humans have voluntarily become data-gathering devices feeding global-scale database structures. Some inner urge makes us want to capture moments rather than experience them. And it is not only happening with our mobile devices' cameras: global web companies track each and every move of our mouses and trackpads, each scroll and every pause.

Gathering data has become the lifestyle of the 21st century. And this data is used to train machine learning algorithms mimicking our own behavior. From image tagging and text translation to content generation: machines are made to behave in forms similar to the data gathered by/through us.

This, however, is not an essay about the why: the urge to "save for later," to record the most mundane is a fascinating question, but what I am mainly interested in is the discrepancy between the model of digital data and the world we inhabit.

1 Here, Borges quotes from Pliny's *Historia naturalis*: "Nothing that has been heard can be repeated with the same words" (my translation). First published in 1942 as "Funes el memorioso," the following translation of Borges's short story is referred to in this text: Jorge Luis Borges, "Funes, the Memorious," trans. Anthony Kerrigan, in *Ficciones*, ed. Anthony Kerrigan (New York: Grove Press, 1962), 107–115 (quotation on 111).



Grid compressed as cosine wave functions

As everyone who reads this text knows: we all get older. At some point, wrinkles appear, hair gets thinner, exercising becomes harder and memory lossier. The same is true for every pair of shoes I have owned: traces of time are unavoidable. Everything grows and decays: A tree sprouts leaves, they grow and fall off in autumn, reentering a circle of growth and decay. The pre-digital photos from your childhood show the signs of time as colors slowly desaturate and a faint, ever-growing yellow haze develops. Over millions of years sand transforms into stone through pressure, temperature, and chemical reactions. Nothing is permanent, everything is in flux; always.

However, data as transmissible and storable computer information[2] is not conceptualized in this manner: Data has a binary concept of decay, i.e., it is or it is not, and the transition between these two states is almost non-existent. Data centers are like cold storage facilities for fruit or ancient books in which the hardware slowly decays over time; but the content, the data, does not decay in the same manner. The distinction between hardware and software, between the physical and the digital, or between the medium and the message conceptualizes data as immortal. While errors occur when copying DNA from one cell to another, copying data from one storage device to another is, in most cases, error-free. The process of lossless copying allows computational data to be conceptualized as eternal. Its immortality might be the reason why humanity is so drawn to it: the "later" in "save for later" is attached to the promise of eternity.

While my favorite musician gets older, retires, or dies, the shaky video footage recorded with my mobile phone and stored in a network of global data centers will

2 See *Online Etymological Dictionary*, s.v. "data (n.)," accessed November 1, 2022, https://www.etymonline.com/word/data.

Decaying strings of data

continue to exist, unaffected by the very mechanisms of aging everything else in the world is subject to. Data overcomes death through lossless copies of itself. This is, of course, not entirely true: codecs change, data is deleted, and storage devices break. Keeping computational data intact is hard work. But what makes data so special in comparison to anything else in the world is that data does not decay. As long as I can open it and it is not deleted, a digital image will stay the same. It will not develop a yellow tint, the colors will not slowly desaturate and fade away. The software part in the twofold hard- and software divide will stay on for as long as it exists. The binary underpinnings of computation also determine the concept of data storage.—Data either is or it is not.

Fallacies of Data

> I have more memories in myself alone than all men have had since the world was a world.
> —Jorge Luis Borges, *Funes, the Memorious*[3]

One can easily argue that the conceptualization of digital data as immortal is one of the biggest advances: its lossless reproducibility and its capabilities as a non-decaying storage medium are a vast improvement in a world in which everything else will sooner or later adhere to the laws of entropy.—A world in which the characters in older Wikipedia articles would slowly fade away, or webpages would turn yellow if not constantly updated is, after all, a strange thought experiment. But to some extent, this is happening: website links do disappear, and if I don't constantly update all my devices with the latest software, these systems become unusable. However, the mode at play here is a different one.—Code can be read by a system or it cannot.

In spring 2022, during my residency at ZKM, a group of researchers tried to get one of the earliest computers, the Zuse Z22 with the serial number 13, running again. This computer, developed by Konrad Zuse and lead designer Lorenz Hanewinkel in 1957, uses 415 vacuum tubes.[4] 65 years

3 Borges, "Funes," 112.

4 See "Zuse Z22," ZKM | Karlsruhe, accessed November 1, 2022, https://zkm.de/en/artwork/zuse-z22.

after its development, it proved to be a laborious task to execute the software on its hardware. The distinction I am intrigued by is that, once the software is running, it will run just as it did in 1957. While photographs from that time stored on analog photo paper will show signs of decay, the software code remains unaffected by time.

As everyone knows who has tried to restart an old computer or an old hard drive, computational systems are fragile. But the underlying concept is different from that of decaying entities. Within the next three sections, I want to investigate three perspectives on why such a concept of data might be problematic. There are social, climatic, and conceptual reasons why it makes sense that everything in the world is in flux. The concept of data as eternal, as a non-decaying entity might not always be favorable.

Social Aspects

Over the past 40 years, the internet moved away from its military and research origins to become a global phenomenon. With its transition to a mass medium, storage capabilities and content changed dramatically. Especially through the advent of large online platforms, credit reporting agencies, and consumer data brokers, companies emerged that have extensive knowledge about a large segment of the world's population.[5] Some research even suggests that, based on the data gathered, social media companies can judge a user's personality better than close relatives or friends could.[6] Furthermore, advancements in machine learning result in troublesome new search engines that can find photos of a person based on a single image.[7] Oftentimes, as I have also experienced myself, the results of such searches contain images unknown to oneself. The scale and magnitude of societal change that will arise from individual data cannot be fully grasped yet. We live in an era that gambles with the concept of privacy, and it is not clear who will and who will lose in this game.

The European Union, among other countries such as Argentina and the Philippines, has thus introduced *right to be forgotten* laws. These laws are supposed to prevent stigmatization based on past actions.[8] But these laws carry entail various problems and issues: One concern is the question of where to draw the line between the right to be forgotten and the freedom of speech.[9] The executive director of the Wikimedia Foundation Lila Tretikov criticized the EU for punching holes in free

5 See Wolfie Christl, "Corporate Surveillance in Everyday Life," Cracked Labs, June 2017, https://crackedlabs.org/en/corporate-surveillance.

6 See Wu Youyou, Michal Kosinski, and David Stillwell, "Computer-Based Personality Judgments Are More Accurate than Those Made by Humans," *Proceedings of the National Academy of Sciences* 112, no. 4 (2015): 1036–1040, https://doi.org/10.1073/pnas.1418680112.

7 See Kashmir Hill, "A Face Search Engine Anyone Can Use Is Alarmingly Accurate," *The New York Times*, May 26, 2022, https://www.nytimes.com/2022/05/26/technology/pimeyes-facial-recognition-search.html.

8 See Alessandro Mantelero, "The EU Proposal for a General Data Protection Regulation and the Roots of the 'Right to Be Forgotten,'" *Computer Law & Security Review* 29, no. 3 (2013): 229–235, https://doi.org/10.1016/j.clsr.2013.03.010.

9 See Daniel J. Solove, "The Virtues of Knowing Less: Justifying Privacy Protections Against Disclosure," *Duke Law Journal* 53, no. 3 (2003), https://scholarship.law.duke.edu/dlj/vol53/iss3/2; and Tessa Mayes, "We Have No Right to Be Forgotten Online," *The Guardian*, March 18, 2011, https://www.theguardian.com/commentisfree/libertycentral/2011/mar/18/forgotten-online-european-union-law-internet.

knowledge.[10] Another issues is the technical implementation: With digital data being so easy to copy, it is difficult to contain its spreading. In sociology, this is known as the *Streisand effect*, meaning that containing information—in this case, an image of Barbra Streisand's residence in Malibu—leads to the unintended consequence of increasing awareness and thus the spreading of this very information.

The difficulties of "digital forgetting" put a burden on individual victims as exemplified in a 2019 article by the *New York Times* on online child abuse content:[11] digital images remain online forever as long as they are duplicated and stored in various online locations. Even in the illegal case of child abuse content, the right to be forgotten seems impossible to enforce. Social media and cloud storage companies are caught between detecting these images and maintaining the privacy of their customers. The digital concept of immortal data creates disputes and suffering on a social scale. The digital notion of lossless information reproduction leads to a world in which certain actions that are destructive on a personal level are perpetuated virtually.

Ecological Aspects

The metaphor of the *cloud* in cloud computing and cloud storage is an ill-defined one: clouds consisting of frozen or liquid droplets suspended in the atmosphere have little to do with the server systems that make planetary networked computation possible. While the floating water particles and the rain resulting from them enable life on earth, cloud storage is becoming one of the most energy-intense and thus destructive industries on the planet. In 2007, an average data center consumed as much energy as 25,000 homes.[12] Over the past decade, this consumption increased dramatically with no sign of slowing down in the 2030s.

Similarly, the metaphor of the *web* draws on the image of almost invisible silk-based spiderwebs. The web metaphor suggests lightness, transparency, extraordinary stability, and a state of floating in the sky. In 2019, about 20 billion devices had been connected to the internet. Predictions estimate a doubling from 20 to 40 billion devices by 2025.[13] In 1987, the traffic to and from data centers amounted to about 2 terabytes, double the amount of the hard drive this text is written on. In 2007, that number grew 250 million times to 50 exabytes. In 2017, that number reached 1.1 zettabytes (a 10 with 21 zeros). By 2017, data centers used an estimated 200 terawatt-hours each year, about 1% of global electricity

10 See Lila Tretikov, "European Court Decision Punches Holes in Free Knowledge," *Diff*, Wikimedia Foundation, August 6, 2014, https://diff.wikimedia.org/2014/08/06/european-court-decision-punches-holes-in-free-knowledge/.

11 See Michael H. Keller and Gabriel J. X Dance, "Child Abusers Run Rampant as Tech Companies Look the Other Way," *The New York Times*, November 9, 2019, https://www.nytimes.com/interactive/2019/11/09/us/internet-child-sex-abuse.html.

12 See James M. Kaplan, William Forrest, and Noah Kindler, *Revolutionizing Data Center Energy Efficiency*, ed. McKinsey & Company (2008), 2, https://sallan.org/pdf-docs/McKinsey_Data_Center_Efficiency.pdf.

13 See Lionel S. Vailshery, "Global IoT and Non-IoT Active Device Connections Worldwide from 2010 to 2025," Statista, September 6, 2022, https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/.

consumption. They contributed about 0.3% to overall carbon emissions.[14] These numbers are predicted to double in the 2020s.[15]

In 2018, a study by OpenAI predicted that the power needed to train state-of-the-art machine learning models would be doubling every 3.4 months.[16] Training GPT-3, a deep learning model to produce human-like text, consumed an estimated 1,404 megawatt-hours in 2020. Crypto-currencies as well as artificial intelligence are marketed as the latest innovations in technology. But these developments do not seem to take into account that we will need to pursue a fundamentally different path in the future.

Cryptocurrencies are the latest technological development spurring the notion of data as immortal—the distributed ledger is the ultimate invention of non-forgetting. Not only is every transaction stored in a blockchain, but every blockchain stores every transaction. Every system attached needs to hold the entire list of records, called blocks. Removal, decay, or forgetting would break the entire system. Crypto-currencies are conceptualized on an accumulation of data; decay is not only excluded but would fracture the system. In 2021, the energy consumption related to crypto-currency *mining* alone amounted to about 120 terawatt-hours of electricity.[17] According to a comparison by the BBC, the energy consumed by crypto mining could power all electric kettles used in the UK for 27 years.[18] Various calculations suggest that one Bitcoin transaction emits as much CO2 as an average household in three weeks[19] or about the same as a business class flight from Berlin to London.[20]

The crypto trend is now followed by the metaverse, the dream of an integrated network of 3D virtual worlds. The meta-verse television series *Upload* is a perfect example of the dreams of immortality through digital technology. The plot is simple: by 2033, humans can "upload" themselves into a virtual afterlife. Similarly, the series *Black Mirror* contains various episodes discussing the virtual afterlife.[21] The metaverse is another step towards a concept of digital eternity: death is lurking behind every corner and data is the 21st-century mechanism to cope with it.

The metaphor of the cloud, the web, artificial intelligence, cryptocurrency, and the metaverse all obscure the actual footprint of these technological infrastructures. It might not be a pure coincidence that crypto, from Greek *kruptós,* means

14 See Nicola Jones, "How to Stop Data Centres from Gobbling Up the World's Electricity," *Nature* 561, no. 7722 (2018): 163–166, https://doi.org/10.1038/d41586-018-06610-y.

15 See Anders S. G. Andrae, "Total Consumer Power Consumption Forecast" (presentation, Nordic Digital Business Summit, Helsinki, October 2017).

16 See Dario Amodei and Danny Hernandez, "AI and Compute," OpenAI, May 16, 2018, https://openai.com/blog/ai-and-compute/#modern.

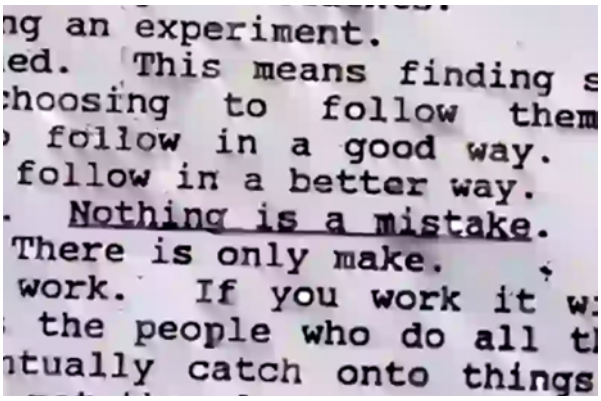
17 See "Cambridge Bitcoin Electricity Consumption Index (CBECI)," Cambridge Centre for Alternative Finance, accessed November 1, 2022, https://ccaf.io/cbeci/index.

18 See Cristina Criddle, "Bitcoin Consumes 'More Electricity than Argentina,'" *BBC News*, February 10, 2021, https://www.bbc.com/news/technology-56012952.

19 See Orlando Whitehead, "One Bitcoin Transaction Emits as Much CO2 as a Household in 3 Weeks," *The Brussels Times*, January 15, 2022, https://www.brusselstimes.com/201597/one-bitcoin-transaction-emits-as-much-co2-as-a-household-in-3-weeks.

20 See "CO2 Footprint per Bitcoin Transaction," *Economics Stack Exchange*, Stack Exchange, accessed November 1, 2022, https://economics.stackexchange.com/questions/32120/co2-footprint-per-bitcoin-transaction.

21 See Sophie Gilbert, "*Upload* Satirizes a Capitalist Heaven," *The Atlantic*, May 7, 2020, https://www.theatlantic.com/culture/archive/2020/05/upload-black-mirror-hell-digital-heaven/611293/.



Lossy JPEG 02

hidden, whereas data, from Latin *datum*, means *the given:* their transformation and production are far from the metaphors used to describe these processes. The technological infrastructure puts a heavy burden on the planet. The tools to detect climate change in the first place, i.e., supercomputers and global sensing devices, are the tools that multiply the problem.—Understanding and destruction amalgamate.

However, the constant flux of growth and decay of the world we inhabit creates space for alternatives. The hoarding of the digital and its various trending terms only take one side into account: growth, but not decay. Every new technological invention seems to not only stay on this path but to exponentially multiply it.

Conceptual Aspects

> There is a pretense of self, in other words, which is only expressed in this narrative; a self that is these sentences. We tell their story, and thereby come to what consciousness we have. Scribble ergo sum.
> — Kim S. Robinson, *Aurora*[22]

22 Kim S. Robinson, *Aurora* (London: Orbit Books, 2016), chap. 6, EPUB.

The argument I have presented so far questions the concept of digital data as eternal. Digital data does not decay, it is or it is not. While such a concept makes sense with regard to the origins of computational systems in military and research programs, it causes great disruptions on both the social and environmental level in the contemporary world. What if the model of eternal data is not only socially disruptive and environmentally questionable but, furthermore, a conceptually false understanding of memory and knowledge?

In 1942, five years after the development of the first electromechanical computer,[23] Argentinian author Jorge Luis Borges published a short story titled "*Funes el memorioso*" (later published in English as "*Funes, the Memorious*"). It is a tale about a man, Ireneo Funes, who, after falling off a horse, acquires the *talent* of remembering everything. In the short story, Funes is able to remember every shape of every cloud at any given moment, as well as what he perceived in each moment. However, Borges's story is not just fiction: several humans with savant syndrome have abilities similar to the ones described by Borges. For example, Kim Peek (1951–2009) was able to name every city in the US—almost 20,000, in total—, including their zip code and area code, as well as the highway leading to each city. Among other things, he knew the contents of 12,000 books by heart.[24]

23 See Hadwig Dorsch, *Der erste Computer: Konrad Zuses Z1—Berlin 1936. Beginn und Entwicklung einer technischen Revolution* (Berlin: Museum für Verkehr und Technik Berlin, 1989).

24 See Bruce Weber, "Kim Peek, Inspiration for *Rain Man*, Dies at 58," *The New York Times*, December 26, 2009, https://www.nytimes.com/2009/12/27/us/27peek.html.

These various capabilities remind me of the digital systems we are embedded in. An e-book reader or a smartphone can easily store 12,000 books. A dataset with all US cities including various data dimensions is less than 1 megabyte in size. Digital data has capabilities very similar to the skills described. But, as Borges writes, Funes's new skill came at a cost: "He was, let us not forget, almost incapable of general, platonic ideas."[25]

The digital impulse of eternal data, of the human urge to "save for later," to record history to the fullest might replace the virtue of knowledge with masses of data. Planetary-scale networked computation is based on the principle of storing everything, but already Friedrich Nietzsche commented on the power of forgetting:

> Imagine the most extreme example, a human being who does not possess the power to forget, who is damned to see becoming everywhere; such a human being would no longer believe in his own being, would no longer believe in himself, would see everything flow apart in turbulent particles, and would lose himself in this stream of becoming [...]. All action requires forgetting, just as the existence of all organic things requires not only light, but darkness as well.[26]

In *"On the Utility and Liability of History for Life"* (1874), Nietzsche describes too much history as a danger to life: Instead of focusing on experiencing the present, we are too busy focusing on how we are going to reflect on events once they have passed.[27]—Recording takes over living. The question of how one remembers annihilates the experience of the moment.

From a very different perspective, Théodule Ribot writes in *Les maladies de la mémoire* (1881) that "the paradoxical result [is] that one condition of remembering is that we should forget. Without totally forgetting a prodigious number of states of consciousness, and momentarily forgetting a large number, we could not remember at all."[28]

Human memory works in a fundamentally different way than the data memory of computation: In the human brain, sensory memories only last a fraction of a second. The things we devote our attention to become short-term memories, lasting only for a few seconds. The hippocampus plays a key role in abstracting the world, merging details into long-term memories of the past.[29] These various layers of abstraction are missing from how digital data is conceptualized: data does not abstract to synthesize concepts. As Borges writes: "To think is to forget a difference, to generalize, to abstract. In the overly replete world of Funes there were nothing but details, almost contiguous details."[30]

25 Borges, "Funes," 114.

26 Friedrich Nietzsche, "On the Utility and Liability of History for Life," in *The Complete Works of Friedrich Nietzsche*, vol. 2, *Unfashionable Observations*, ed. and trans. Richard T. (Stanford: Stanford University Press, 1995), 83–168 (quotation on 89).

27 Nietzsche writes: *"Before the war is even over, it has already been transformed into a hundred thousand pages of printed paper, it has already been served up as the latest delicacy to the exhausted palates of the history-hungry." Nietzsche, "On the Utility," 116.*

28 Théodule Ribot, *Les maladies de la mémoire* (Paris: Librairie Germer Baillière et Cie, 1881), 46, quoted in William James, *The Principles of Psychology*, (New York: Henry Holt and Company, 1890), 1:681.

29 See Rodrigo Q. Quiroga, *Borges and Memory: Encounters with the Human Brain*, trans. Juan P. Fernández (Cambridge, MA: MIT Press, 2012), chap. 4, EPUB.

30 Borges, "Funes," 115.

Forgetting and abstracting are two key components that digital storage has not grasped. While humans store sensory memories only for a fraction of a second, this level of gathering demonstrates the concept of how data storage works long term. Recently, I requested Google to send me a data dump of all the information the company has about me. The dataset contained a vast list of my search history, spanning almost the last two decades. Google does not forget what I searched for on May 17, 2010, at 8:53 a.m. Data is not abstracted, nor forgotten over time. Just like Funes in Borges's story, digital data is conceptualized as a hoarding of details. The natural and societal consequences of such a misconception of the differences between memory and data are vast.

In his book *Borges and Memory,* Rodrigo Quian Quiroga relates his studies in neurology to the story of Funes as it matched his research findings. Thinking, for Quiroga, is only possible by generating meaning and concepts through the act of abstracting and forgetting. Both factors are highly related to one another as abstracting implies neglecting and forgetting details. Quiroga makes analogies to the technical apparatus we use to capture the world: "Wedo not process images in our brains in the same way a camera does; on the contrary, we extract a meaning and leave aside a multitude of details."[31]

Furthermore, Quiroga calls the 21st-century reality a *"Funes kind of world"*[32] in which we are constantly bombarded with digital information, from social media, and emails, to 24-hour news channels. From my perspective, the very design of how memory is conceptualized as data, as the ever-lasting never-forgetting lies at the core of this problem. Memory, for Quiroga, is a creative process: "To think is to forget differences, to generalize, to abstract."[33] Thus, remembering is not, like data, a form of endless repetition of the same but rather something in flux, an ever-changing interplay between context and meaning.

A New Kind of Datum?

While the topic of digital decay is far from being a well-researched field, it is also untrue that there have been no attempts in this direction: the subject is being explored in various ways, and sometimes decay is a by-product, sometimes it is carried out on purpose. In this last section, I will introduce various approaches and suggest how data decay might be conceptualized further.

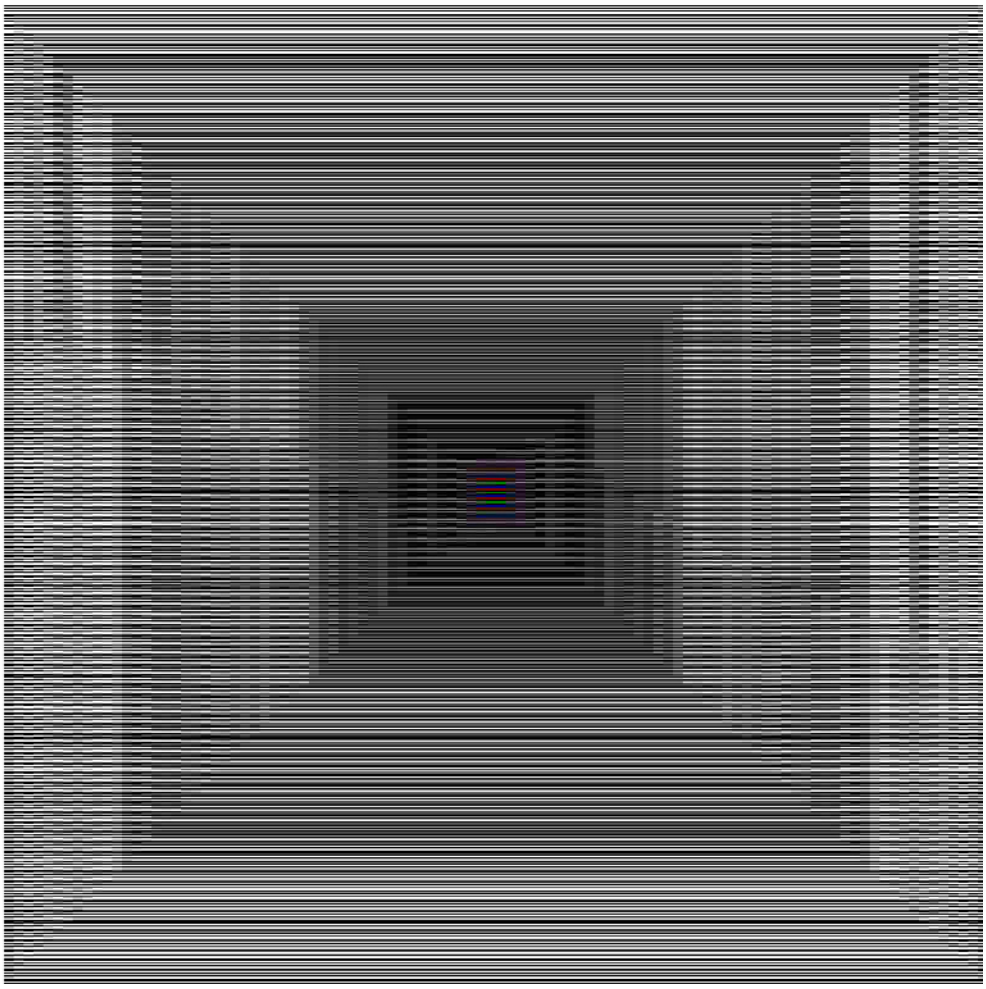
Lossy compression describes a set of techniques to reduce data in size through inexact approximations and partial data discarding. Especially the early internet with its low bandwidth gave rise to such formats. Well-known examples are the image format JPEG, the audio format MP3 as well as various MPEG video formats. The signal processing technique underlying lossy compression is called *discrete cosine transform*.[34] In JPEG images, for example, blocks of eight by eight pixels are transformed into a cosine wave function, which requires less storage compared to the original pixel blocks. If a file is compressed to a high degree, the

31 Quiroga, *Borges and Memory*, chap. 12.

32 Ibid.

33 Ibid.

34 For a good explanation of the process see Cinephile, "JPEG DCT, Discrete Cosine Transform (JPEG Pt2)," *YouTube*, May 22, 2015, https://www.youtube.com/watch?v=Q2aEzeMDHMA.



Data decay of three vertical colored lines

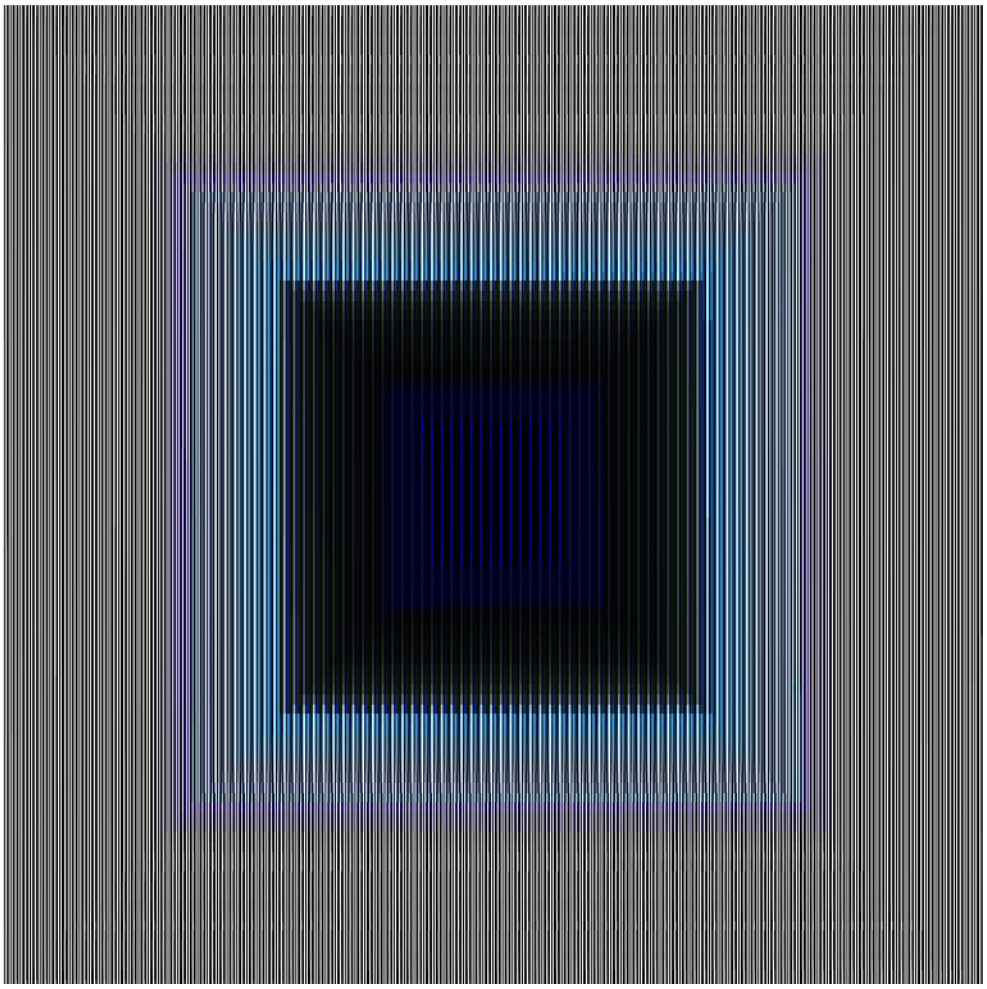
wave functions become visible within the image. Lossy compression affects each new copy of the data, this means a JPEG saved over and over again will lose information over time. But this kind of compression was never meant to forget, nor to abstract or summarize. The motive behind this compression is to send files faster through the slow internet of the 1990s.

While file compression functions on the data level of digital files, various apps with data self-destruction capabilities have been released over the past years. Gmail, for example, has the capability to send emails in a confidential mode.[35] This allows users to set an expiration date on their mails. Similarly, the messenger app Telegram allows for self-destructing messages.[36] When Snapchat was first launched, the app became known for its disappearing photos and videos. These examples demonstrate that digital ephemerality is not unthinkable. However, the systems are all based on a binary concept of forgetting. There is, again, no level of generalization or abstraction.

Rather than following the hype around cryptocurrencies in which remembering every transaction is a key component of the system, cryptography can also be used in reverse, i.e., to forget. The project *Vanish*

35 See "Protect Gmail Messages with Confidential Mode," *Google Workspace Admin Help*, Google, accessed November 1, 2022, https://support.google.com/a/answer/7684332?hl=en.

36 See "Auto-Delete, Widgets and Expiring Invite Links," Telegram, accessed November 1, 2022, https://telegram.org/blog/autodelete-inv2/de?setln=en.



Data decay of horizontal blue lines

by the University of Washington is an example of this: the system makes data unreadable after a user-specified period time; the result is self-destructing data.[37] Already in 2007, Viktor Mayer-Schönberger made a similar argument about self-deleting data from a legal perspective.[38] But the examples observed thus far are all based on the concept of deletion as the only mechanism of forgetting.

Research in the field of artificial neural networks, however, might allow novel conceptualizations of digital memory that include generalization and abstraction. These techniques are inspired by biological neural networks that constitute animal brains, and movements towards abstraction are built into these systems.[39] Each of the hidden convolutional layers of a deep learning architecture abstracts the input layer towards the output layer. That said,

37 See Roxana et al., "New Directions for Self-Destructing Data Systems," (technical report, University of Washington, 2011).

38 See Viktor Mayer-Schönberger, "Useful Void: The Art of Forgetting in the Age of Ubiquitous Computing," KSG Working Paper no. RWP07-022, April 2007, http://dx.doi.org/10.2139/ssrn.976541.

39 See Lorenza Saitta and Jean-Daniel Zucker, "Abstraction in Machine Learning," in *Abstraction in Artificial Intelligence and Complex Systems* (New York: Springer, 2013), 273–327, https://doi.org/10.1007/978-1-4614-7052-6_9; and David Abel, "A Theory of Abstraction in Reinforcement Learning," (PhD diss., Brown University, 2020), arXiv, March 1, 2022, https://doi.org/10.48550/arxiv.2203.00397.

the current machine learning workflows are far from the new kind of data I am imagining: machine learning is based on hoarding massive amounts of data, which are so large and computationally resource-intensive that only large corporations have the capacity to train these models.

The various examples given show that data decay is possible, but also that it is far from the levels of abstraction occurring in a human brain. What I am imagining is a new form of digital storage format in which data decays and abstracts over time.—Rather than having a list of everything I have ever searched for on Google, that data should self-abstract and generalize over time. The further back in time, the list items should merge into one another, so that there is not a record of every search from ten years ago, but rather of the general topics of interest from that year.

Similar to the current conceptualizations of degrowth[40] questioning the notion of development in political, economic, and social terms, as a society, we need to redesign our notion of data. Within human memory, sensory memories turn into short-term memories in a fraction of seconds, and these then accumulate into long-term memories. The hippocampus plays a key role in this transformation. For a sustainable and socially-minded collective future, we will need a hippocampus for data: a clearing mechanism for database structures to abstract, compress, and forget.—Rather than keeping 50 photographs of one moment, these images would turn into a memory by slowly decaying into each other.

I hope that the beginning of the 21st century will go down in history as a time of data hoarding, as an excessive time of storing the unneeded, while in a near future we will move towards a more sustainable data ecology in which data, like everything else, decays, merges, compresses, abstracts, and simplifies over time.

40 See Matthias Schmelzer, Aaron Vansintjan, and Andrea Vetter, *The Future Is Degrowth: A Five Book Plan* (London/New York: Verso Books, 2022).