

Hito Steyerl

A Sea of Data: Apophenia and Pattern (Mis-)Recognition

01/14

e-flux journal #72 — april 2016 Hito Steyerl
A Sea of Data: Apophenia and Pattern (Mis-)Recognition

This is an image from the Snowden files. It is labeled “secret.”¹ Yet one cannot see anything on it.

This is exactly why it is symptomatic.

Not seeing anything intelligible is the new normal. Information is passed on as a set of signals that cannot be picked up by human senses. Contemporary perception is machinic to large degrees. The spectrum of human vision only covers a tiny part of it. Electric charges, radio waves, light pulses encoded by machines for machines are zipping by at slightly subluminal speed. Seeing is superseded by calculating probabilities. Vision loses importance and is replaced by filtering, decrypting, and pattern recognition. Snowden’s image of noise could stand in for a more general human inability to perceive technical signals unless they are processed and translated accordingly.

But noise is not nothing. On the contrary, noise is a huge issue, not only for the NSA but for machinic modes of perception as a whole.

Signal v. Noise was the title of a column on the internal NSA website running from 2011 to 2012. It succinctly frames the NSA’s main problem: how to extract “information from the truckloads of data”:

It’s not about the data or even access to the data. It’s about getting information from the truckloads of data ... Developers, please help! We’re drowning (not waving) in a sea of data – with data, data everywhere, but not a drop of information.²

Analysts are choking on intercepted communication. They need to unscramble, filter, decrypt, refine, and process “truckloads of data.” The focus moves from acquisition to discerning, from scarcity to overabundance, from adding on to filtering, from research to pattern recognition. This problem is not restricted to secret services. Even WikiLeaks Julian Assange states: “We are drowning in material.”³

Apophenia

But let’s return to the initial image. The noise on it was actually decrypted by GCHQ technicians to reveal a picture of clouds in the sky. British analysts have been hacking video feeds from Israeli drones at least since 2008, a period which includes the recent IDF aerial campaigns against Gaza.⁴ But no images of these attacks exist in Snowden’s archive. Instead, there are all sorts of abstract renderings of intercepted broadcasts. Noise. Lines. Color patterns.⁵ According to leaked training manuals, one needs to apply all sorts of massively secret operations to produce these kinds of images.⁶

But let me tell you something. I will decrypt this image for you without any secret algorithm. I will use a secret ninja technique instead. And I will even teach you how to do it for free. Please focus very strongly on this image right now.

Doesn't it look like a shimmering surface of water in the evening sun? Is this perhaps the "sea of data" itself? An overwhelming body of water, which one could drown in? Can you see the waves moving ever so slightly?

I am using a good old method called apophenia.

Apophenia is defined as the perception of patterns within random data.⁷ The most common examples are people seeing faces in clouds or on the moon. Apophenia is about "drawing connections and conclusions from sources with no direct connection other than their indissoluble perceptual simultaneity," as Benjamin Bratton recently argued.⁸

One has to assume that sometimes, analysts also use apophenia.

Someone must have seen the face of Amani al-Nasara in a cloud. The forty-three-year-old was blinded by an aerial strike in Gaza in 2012 in front of her TV:

"We were in the house watching the news

02/14

on TV. My husband said he wanted to go to sleep, but I wanted to stay up and watch Al Jazeera to see if there was any news of a ceasefire. The last thing I remember, my husband asked if I changed the channel and I said yes. I didn't feel anything when the bomb hit – I was unconscious. I didn't wake up again until I was in the ambulance." Amani suffered second degree burns and was largely blinded.⁹

What kind of "signal" was extracted from what kind of "noise" to suggest that al-Nasara was a legitimate target? Which faces appear on which screens, and why? Or to put it differently: Who is "signal," and who disposable "noise"?

Pattern Recognition

Jacques Rancière tells a mythical story about how the separation of signal and noise might have been accomplished in Ancient Greece. Sounds produced by affluent male locals were defined as speech, whereas women, children, slaves, and foreigners were assumed to produce garbled noise.¹⁰ The distinction between speech and noise served as a kind of political spam filter. Those identified as speaking were labeled citizens and the rest as irrelevant, irrational, and



This image from the Snowden files was captioned: "A single frame of scrambled video imagery."

potentially dangerous nuisances. Similarly, today, the question of separating signal and noise has a fundamental political dimension. Pattern recognition resonates with the wider question of political recognition. Who is recognized on a political level and as what? As a subject? A person? A legitimate category of the population? Or perhaps as “dirty data”?

What are dirty data? Here is one example:

Sullivan, from Booz Allen, gave the example the time his team was analyzing demographic information about customers for a luxury hotel chain and came across data showing that teens from a wealthy Middle Eastern country were frequent guests.

“There were a whole group of 17 year-olds staying at the properties worldwide,” Sullivan said. “We thought, ‘That can’t be true.’”¹¹

The demographic finding was dismissed as dirty data – a messed up and worthless set of information – before someone found out that, actually, it was true.

03/14

Brown teenagers, in this worldview, are likely to exist. Dead brown teenagers? Why not? But rich brown teenagers? This is so improbable that they must be dirty data and cleansed from your system! The pattern emerging from this operation to separate noise and signal is not very different from Rancière’s political noise filter for allocating citizenship, rationality, and privilege. Affluent brown teenagers seem just as unlikely as speaking slaves and women in the Greek polis.

On the other hand, dirty data are also something like a cache of surreptitious refusal; they express a refusal to be counted and measured:

A study of more than 2,400 UK consumers by research company Verve found that 60% intentionally provided wrong information when submitting personal details online. Almost one quarter (23 percent) said they sometimes gave out incorrect dates of birth, for example, while 9 percent said they did this most of the time and 5 percent always did it.¹²

Dirty data is where all of our refusals to fill out the constant onslaught of online forms accumulate. Everyone is lying all the time,



Rose Mary Woods, Nixon's lifelong secretary, demonstrates the “Rose Mary Stretch,” a gesticulation that purportedly led to the erasure of a section of the Watergate tapes. The quality of noise in this section of the tapes has been thoroughly analyzed to understand if the omission was intentional. Photo: Wikimedia commons.

whenever possible, or at least cutting corners. Not surprisingly, the “dirtiest” area of data collection is consistently pointed out to be the health sector, especially in the US. Doctors and nurses are singled out for filling out forms incorrectly. It seems that health professionals are just as unenthusiastic about filling out forms for systems designed to replace them, as consumers are about performing clerical work for corporations that will spam them in return.

In his book *The Utopia of Rules*, David Graeber gives a profoundly moving example of the forced extraction of data. After his mom suffered a stroke, he went through the ordeal of having to apply for Medicaid on her behalf:

I had to spend over a month ... dealing with the ramifying consequences of the act of whatever anonymous functionary in the New York Department of Motor Vehicles had inscribed my given name as “Daid,” not to mention the Verizon clerk who spelled my surname “Grueber.” Bureaucracies public and private appear – for whatever historical reasons – to be organized in such a way as to guarantee that a significant proportion of actors will not be able to perform their tasks as expected.¹³

04/14

Graeber goes on to call this an example of utopian thinking. Bureaucracy is based on utopian thinking because it assumes people to be perfect from its own point of view. Graeber’s mother died before she was accepted into the program.

The endless labor of filling out completely meaningless forms is a new kind of domestic labor in the sense that it is not considered labor at all and assumed to be provided “voluntarily” or performed by underpaid so-called data janitors.¹⁴ Yet all the seemingly swift and invisible action of algorithms, their elegant optimization of everything, their recognition of patterns and anomalies – this is based on the endless and utterly senseless labor of providing or fixing messy data.

Dirty data is simply real data in the sense that it documents the struggle of real people with a bureaucracy that exploits the uneven distribution and implementation of digital technology.¹⁵ Consider the situation at LaGeSo (the Health and Social Affairs Office) in Berlin, where refugees are risking their health on a daily basis by standing in line outdoors in severe winter weather for hours or even days just to have their data registered and get access to



This photograph from June 6, 2012 shows a student pilot and sensor operator manning the controls of a MQ-9 Reaper in a ground-based cockpit during a training mission flown from Hancock Field Air National Guard Base, Syracuse, New York. Photo: AP Photo.

services to which they are entitled (for example, money to buy food).¹⁶ These people are perceived as anomalies because, in addition to having the audacity to arrive in the first place, they ask that their rights be respected. There is a similar political algorithm at work: people are blanked out. They cannot even get to the stage to be recognized as claimants. They are not taken into account.

On the other hand, technology also promises to separate different categories of refugees. IBM's Watson AI system was experimentally programmed to potentially identify terrorists posing as refugees:

IBM hoped to show that the i2 EIA could separate the sheep from the wolves: that is, the masses of harmless asylum-seekers from the few who might be connected to jihadism or who were simply lying about their identities ...

IBM created a hypothetical scenario, bringing together several data sources to match against a fictional list of passport-carrying refugees. Perhaps the most important dataset was a list of names of

05/14

casualties from the conflict gleaned from open press reports and other sources. Some of the material came from the Dark Web, data related to the black market for passports; IBM says that they anonymized or obscured personally identifiable information in this set ...

Borene said the system could provide a score to indicate the likelihood that a hypothetical asylum seeker was who they said they were, and do it fast enough to be useful to a border guard or policeman walking a beat.¹⁷

The cross-referencing of unofficial databases, including dark web sources, is used to produce a "score," which calculates the probability that a refugee might be a terrorist. The hope is for a pattern to emerge across different datasets, without actually checking how or if they correspond to any empirical reality. This example is actually part of a much larger subset of "scores": credit scores, academic ranking scores, scores ranking interaction on online forums etc., which classify people according to financial interactions, online behavior, market



The Russian TV station Zvezda claimed this flock of birds over New York City appeared to form the shape of President Vladimir Putin's face. YouTube video screenshot.



A plate of spaghetti meatballs return our gaze, courtesy of Google inceptionism. Source: Mary-Ann Russon, "Google DeepDream robot: 10 weirdest images produced by AI 'inceptionism' and users online," International Business Times, July 6, 2015 <http://www.ibtimes.co.uk/google-deepdream-robot-10-weirdest-images-produced-by-ai-inceptionism-users-online-1509518>

data, and other sources. A variety of inputs are boiled down to a single number – a superpattern – which may be a “threat” score or a “social sincerity score,” as planned by Chinese authorities for every single citizen within the next decade. But the input parameters are far from being transparent or verifiable. And while it may be seriously desirable to identify Daesh moles posing as refugees, a similar system seems to have worrying flaws.

The NSA’s SKYNET program was trained to find terrorists in Pakistan by sifting through cell phone customer metadata. But experts criticize the NSA’s methodologies. “There are *very few* ‘known terrorists’ to use to train *and test* the model,” explained Patrick Ball, a data scientist and director of the Human Rights Data Analysis Group, to *Ars Technica*. “If they are using the same records to train the model as they are using to test the model, their assessment of the fit is completely bullshit.”¹⁸

Human Rights Data Analysis Group estimates that around 99,000 Pakistanis might have ended up wrongly classified as terrorists by SKYNET, a statistical margin of error that might have had deadly consequences given the fact that the US is waging a drone war on suspected militants in the country and between 2500 and four thousand people are estimated to have been killed since 2004: “In the years that have

08/14

followed, thousands of innocent people in Pakistan may have been mislabelled as terrorists by that ‘scientifically unsound’ algorithm, possibly resulting in their untimely demise.”¹⁹

One needs to emphasize strongly that SKYNET’s operations cannot be objectively assessed, since it is not known how it’s results were utilized. It was most certainly not the only factor in determining drone targets.²⁰ But the example of SKYNET demonstrates just as strongly that a “signal” extracted by assessing correlations and probabilities is not the same as an actual fact, but determined by the inputs the software uses to learn, and the parameters for filtering, correlating, and “identifying.” The old engineer wisdom “crap in – crap out” seems to still apply. In all of these cases – as completely different as they are technologically, geographically, and also ethically – some version of pattern recognition was used to classify groups of people according to political and social parameters. Sometimes it is as simple as, *we try to avoid registering refugees*. Sometimes there is more mathematical mumbo jumbo involved. But many methods used are opaque, partly biased, exclusive, and – as one expert points out – sometimes also “ridiculously optimistic.”²¹

Corporate Animism

How to recognize something in sheer noise? A



Previously unknown archaeological monuments have been revealed as of September 2015 by the Stonehenge Hidden Landscapes Project. The findings include new information about the world’s largest “super henge” and include ritual monuments such as the mortuary building pictured above in a 3-D reconstruction. Copyright: LBI ArchPro, Joachim Brandtner

striking visual example of pure and conscious apophenia was recently demonstrated by research labs at Google:²²

We train an artificial neural network by showing it millions of training examples and gradually adjusting the network parameters until it gives the classifications we want. The network typically consists of 10–30 stacked layers of artificial neurons. Each image is fed into the input layer, which then talks to the next layer, until eventually the “output” layer is reached. The network’s “answer” comes from this final output layer.²³

Neural networks were trained to discern edges, shapes, and a number of objects and animals and then applied to pure noise. They ended up “recognizing” a rainbow-colored mess of disembodied fractal eyes, mostly without lids, incessantly surveilling their audience in a strident display of conscious pattern overidentification.

Google researchers call the act of creating a pattern or an image from nothing but noise “inceptionism” or “deep dreaming.” But these entities are far from mere hallucinations. If they are dreams, those dreams can be interpreted as condensations or displacements of the current technological disposition. They reveal the networked operations of computational image creation, certain presets of machinic vision, its hardwired ideologies and preferences.

One way to visualize what goes on is to turn the network upside down and ask it to enhance an input image in such a way as to elicit a particular interpretation. Say you want to know what sort of image would result in “Banana.” Start with an image full of random noise, then gradually tweak the image towards what the neural net considers a banana. By itself, that doesn’t work very well, but it does if we impose a prior constraint that the image should have similar statistics to natural images, such as neighboring pixels needing to be correlated.²⁴

In a feat of genius, inceptionism manages to visualize the unconscious of prosumer networks: images surveilling users, constantly registering their eye movements, behavior, preferences, aesthetically helplessly adrift between *Hundertwasser* mug knockoffs and Art Deco friezes gone ballistic. Walter Benjamin’s “optical unconscious” has been upgraded to the unconscious of computational image divination.²⁵

09/14

e-flux journal #72 — april 2016 Hito Steyerl
A Sea of Data: Apophenia and Pattern (Mis-)Recognition

By “recognizing” things and patterns that were not given, inceptionist neural networks eventually end up effectively identifying a new totality of aesthetic and social relations. Presets and stereotypes are applied, regardless of whether they “apply” or not: “The results are intriguing – even a relatively simple neural network can be used to over-interpret an image, just like as children we enjoyed watching clouds and interpreting the random shapes.”²⁶

But inceptionism is not just a digital hallucination. It is a document of an era that trains smartphones to identify kittens, thus hardwiring truly terrifying jargons of cutesy into the means of production.²⁷ It demonstrates a version of corporate animism in which commodities are not only fetishes but morph into franchised chimeras.

Yet these are deeply realist representations. According to György Lukacs, “classical realism” creates “typical characters,” insofar as they represent the objective social (and in this case technological) forces of our times.²⁸

Inceptionism does that and more. It also gives those forces a face – or more precisely, innumerable eyes. The creature that stares at you from your plate of spaghetti and meatballs is not an amphibian beagle. It is the ubiquitous surveillance of networked image production, a form of memetically modified intelligence that watches you in the shape of the lunch that you will Instagram in a second if it doesn’t attack you first. Imagine a world of enslaved objects remorsefully scrutinizing you. Your car, your yacht, your art collection observes you with a gloomy and utterly desperate expression. *You may own us, they seem to say, but we are going to inform on you. And guess what kind of creature we are going to recognize in you!*²⁹

Data Neolithic

But what are we going to make of automated apophenia?³⁰ Are we to assume that machinic perception has entered its own phase of magical thinking? Is this what commodity enchantment means nowadays: hallucinating products? It might be more accurate to assume that humanity has entered yet another new phase of magical thinking. The vocabulary deployed for separating signal and noise is surprisingly pastoral: data “farming” and “harvesting,” “mining” and “extraction” are embraced as if we lived through another massive neolithic revolution³¹ with its own kind of magic formulas.

All sorts of agricultural and mining technologies – that were developed during the neolithic – are reinvented to apply to data. The stones and ores of the past are replaced by silicone and rare earth minerals, while a Minecraft paradigm of extraction describes the

processing of minerals into elements of information architecture.³²

Pattern recognition was an important asset of neolithic technologies too. It marked the transition between magic and more empirical modes of thinking. The development of the calendar by observing patterns in time enabled more efficient irrigation and agricultural scheduling. Storage of cereals created the idea of property. This period also kick-started institutionalized religion and bureaucracy, as well as managerial techniques including laws and registers. All these innovations also impacted society: hunter and gatherer bands were replaced by farmer kings and slaveholders. The neolithic revolution was not only technological but also had major social consequences.

Today, expressions of life as reflected in data trails become a farmable, harvestable, minable resource managed by informational biopolitics.³³

And if you doubt that this is another age of magical thinking, just look at the NSA training manual for unscrambling hacked drone intercepts. As you can see, you need to bewitch the files with a magic wand. (Image Magick is a free image converter):

10/14

The supposedly new forms of governance emerging from these technologies look partly archaic and partly superstitious. What kind of corporate/state entities are based on data storage, image unscrambling, high-frequency trading, and Daesh Forex gaming? What are the contemporary equivalents of farmer kings and slaveholders, and how are existing social hierarchies radicalized through examples as vastly different as tech-related gentrification and jihadi online forum gamification? How does the world of pattern recognition and big-data divination relate to the contemporary jumble of oligocracies, troll farms, mercenary hackers, and data robber barons supporting and enabling bot governance, Khelifah clickbait and polymorphous proxy warfare? Is the state in the age of Deep Mind, Deep Learning, and Deep Dreaming a Deep State™? One in which there is no appeal nor due process against algorithmic decrees and divination?

But there is another difference between the original and the current type of “neolithic,” and it harks back to pattern recognition. In ancient astronomy, star constellations were imagined by projecting animal shapes into the skies. After cosmic rhythms and trajectories had been recorded on clay tablets, patterns of movement





Could this image be a representation of the neo-neolithic? Source: Mary-Ann Russon, "Google DeepDream robot: 10 weirdest images produced by AI 'inceptionism' and users online," International Business Times, July 6, 2015 <http://www.ibtimes.co.uk/google-deepdream-robot-10-weirdest-images-produced-by-ai-inceptionism-users-online-1509518>.

started to emerge. As additional points of orientation, some star groups were likened to animals and heavenly beings. However, progress in astronomy and mathematics happened not because people kept believing there were animals or gods in space, but on the contrary, because they accepted that constellations were expressions of a physical logic. The patterns were projections, not reality. While today statisticians and other experts routinely acknowledge that their findings are mostly probabilistic projections, policymakers of all sorts conveniently ignore this message. In practice you become coextensive with the data-constellation you project. Social scores of all different kinds – credit scores, academic scores, threat scores – as well as commercial and military pattern-of-life observations impact the real lives of real people, both reformatting and radicalizing social hierarchies by ranking, filtering, and classifying.

images of this attack, at least as far as I know of. All we know is what she told Human Rights Watch. This is what she said: “I can’t see – ever since the bombing, I can only see shadows.”³⁴

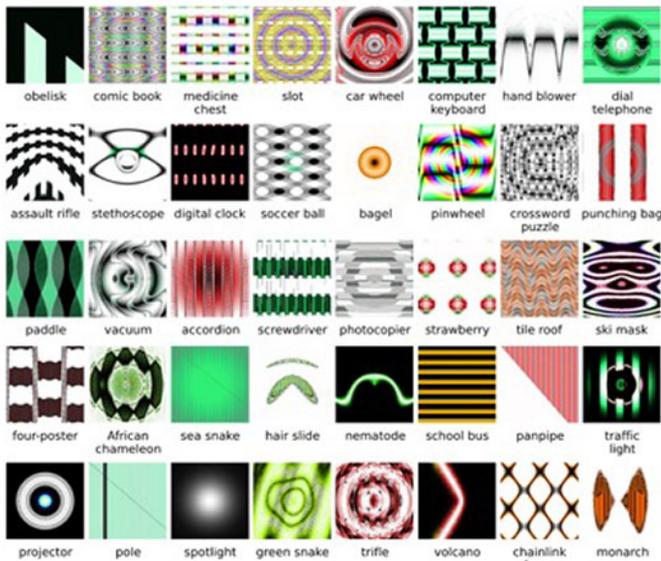
So there is one more way to decode this image. It’s plain for everyone to see. We see what Amani *cannot* see.

In this case, the noise must be a “document” of what she “sees” now: “the shadows.”

Is this a document of the drone war’s optical unconscious? Of it’s dubious and classified methods of “pattern recognition”? And if so, is there a way to ever “unscramble” the “shadows” Amani has been left with?

×

Acknowledgments: The initial version of this text was written at the request of Laura Poitras, who most generously allowed access to some unclassified documents from the Snowden archive, and a short version was presented during the opening of her show “Astro Noise” at the Whitney Museum. Further thanks to Henrik Moltke for facilitating access to the documents, to Brenda and other members of Laura’s studio, to Linda Stupart for introducing me to the term “apophenia,” and to Ben Bratton for fleshing it out for me.



Source: Anh Nguyen, Jason Yosinski, and Jeff Clune, “Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images,” [cv-foundation.org](http://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Nguyen_Deep_Neural_Networks_2015_CVPR_paper.pdf), 2015 http://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Nguyen_Deep_Neural_Networks_2015_CVPR_paper.pdf.

Gestalt Realism

But let’s assume we are actually dealing with projections. Once one accepts that the patterns derived from machinic sensing are *not* the same as reality, information definitely becomes available with a certain degree of veracity.

Let’s come back to Amani al-Nasara, the woman blinded by an aerial attack in Gaza. We know: the abstract images recorded as intercepts of IDF drones by British spies do *not* show the aerial strike in Gaza that blinded her in 2012. The dates *don’t* match. There is *no* evidence in Snowden’s archive. There are *no*

Hito Steyerl is a filmmaker and writer who lives in Berlin.

1
See
<https://www.documentcloud.org/documents/2699846-Anarchist-Training-mod5-Redacted-Com pat.html>

2
"The SIGINT World Is Flat," *Signal v. Noise* column, December 22, 2011.

3
Michael Sontheimer, "SPIEGEL Interview with Julian Assange: 'We Are Drowning in Material,'" *Spiegel Online*, July 20, 2015 <http://www.spiegel.de/international/world/spiegel-interview-with-wikileaks-head-julian-assange-a-1044399.html>

4
Cora Currier and Henrik Moltke, "Spies in the Sky: Israeli Drone Feeds Hacked By British and American Intelligence," *The Intercept*, January 28, 2016 <https://theintercept.com/2016/01/28/israeli-drone-feeds-hacked-by-british-and-american-intelligence/>

5
Ibid. Many of these images are currently part of Laura Poitras's excellent show "Astro Noise" at the Whitney Museum in New York.

6
In the training manual on how to decode these feeds, analysts proudly declared they used open source software developed by the University of Cambridge to hack Sky TV. See <https://www.documentcloud.org/documents/2699846-Anarchist-Training-mod5-Redacted-Com pat.html>

7
See <https://en.wikipedia.org/wiki/Apophenia>

8
Benjamin H. Bratton, "Some Trace Effects of the Post-Anthropocene: On Accelerationist Geopolitical Aesthetics," *e-flux journal* 46 (June 2013) www.e-flux.com/journal/some-trace-effects-of-the-post-anthropocene-on-accelerationist-geopolitical-aesthetics/

9
"Israel: Gaza Airstrikes Violated Laws of War," *hrw.org*, February 12, 2013 <https://www.hrw.org/news/2013/02/12/israel-gaza-airstrikes-violated-laws-war>

10
Jacques Rancière, "Ten Theses on Politics." *Theory & Event*, Vol. 5, No. 3 (2001). "In order to refuse the title of political subjects to a category – workers, women, etc. – it has traditionally been sufficient to assert that they belong to a 'domestic' space, to a space separated from public life; one from which only groans or cries expressing suffering, hunger, or anger could emerge, but not actual speeches demonstrating

a shared aisthesis. And the politics of these categories ... has consisted in making what was unseen visible; in getting what was only audible as noise to be heard as speech."

11
Verne Kopytoff, "Big data's dirty problem," *Fortune*, June 30, 2014 <http://fortune.com/2014/06/30/big-data-dirty-problem/>

12
Larisa Bedgood, "A Halloween Special: Tales from the Dirty Data Crypt," *relevategroup.com*, October 30 2015 <http://www.relevategroup.com/a-halloween-special-tales-from-the-dirty-data-crypt/#sthash.BWrxNRaE.dpuf>. The article continues: "In late June and early July 1991, twelve million people across the country (mostly Baltimore, Washington, Pittsburgh, San Francisco, and Los Angeles) lost phone service due to a typographical error in the software that controls signals regulating telephone traffic. One employee typed a '6' instead of a 'D.' The phone companies essentially lost all control of their networks."

13
David Graeber, *The Utopia of Rules: On Technology, Stupidity and the Secret Joys of Bureaucracy* (Brooklyn: Melville House, 2015), 48.

14
Steve Lohr, "For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights," *New York Times*, August 17, 2014 http://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-janitor-work.html?_r=0

15
See "E-Verify: The Disparate Impact of Automated Matching Programs," chap. 2 in the report *Civil Rights, Big Data, and Our Algorithmic Future*, *bigdata.fairness.io*, September 2014 <https://bigdata.fairness.io/e-verify/>

16
See Melissa Eddy and Katarina Johannsen, "Migrants Arriving in Germany Face a Chaotic Reception in Berlin," *New York Times*, November 26, 2015 http://www.nytimes.com/2015/11/27/world/europe/germany-berlin-migrants-refugees.html?_r=0. A young boy disappeared among the chaos and was later found murdered.

17
Patrick Tucker, "Refugee or Terrorist? IBM Thinks Its Software Has the Answer," *Defense One*, January 27, 2016 <http://www.defenseone.com/technology/2016/01/refugee-or-terrorist-ibm-thinks-its-software-has-answer/125484/>. This example was mentioned by Kate Crawford in her brilliant lecture "Surviving Surveillance," delivered as part of the panel discussion "Surviving Total

Surveillance,” Whitney Museum, February 29, 2016.

18
Christian Grothoff and J. M. Porup, “The NSA’s SKYNET program may be killing thousands of innocent people,” *Ars Technica*, February 16, 2016, italics in original <http://arstechnica.co.uk/SECURITY/2016/02/the-nsas-skyne-t-program-may-be-killing-tho-usands-of-innocent-people/>. An additional bug of the system was that the person who seemed to pose the biggest threat of all according to this program was actually the head of the local Al Jazeera office, because he obviously traveled a lot for professional reasons. A similar misassessment also happened to Laura Poitras, who was rated four hundred out of a possible four hundred points on a US Homeland Security threat scale. As Poitras was filming material for her documentary *My Country, My Country* in Iraq – later nominated for an Academy Award – she ended up filming in the vicinity of an insurgent attack in Baghdad. This coincidence may led to a six-year ordeal that involved her being interrogated, surveilled, searched, etc., every time she reentered the United States from abroad.

19
Ibid.

20
See Michael V. Hayden, “To Keep America Safe, Embrace Drone Warfare,” *New York Times*, February 19, 2016 <http://www.nytimes.com/2016/02/21/opinion/sunday/drone-warfare-precise-effective-imperfect.html>. The director of the CIA from 2006–09, Hayden asserts that human intelligence was another factor in determining targets, while admitting that the program did indeed kill people in error: “In one strike, the grandson of the target was sleeping near him on a cot outside, trying to keep cool in the summer heat. The Hellfire missiles were directed so that their energy and fragments splayed away from him and toward his grandfather. They did, but not enough.”

21
Grothoff and Porup, “The NSA’s SKYNET program.”

22
Thank you to Ben Bratton for pointing this out.

23
“Inceptionism: Going Deeper into Neural Networks,” Google Research Blog, June 17, 2015 <http://googleresearch.blogspot.de/2015/06/inceptionism-going-deeper-into-neural.html>

24
Ibid.

25
Walter Benjamin, “A Short History of Photography,”

available at monoskop.org http://monoskop.org/images/7/79/Benjamin_Walter_1931_1972_A_Short_History_of_Photography.pdf

26
“Inceptionism.”

27
See *ibid.*

28
Farhad B. Idris, “Realism,” in *Encyclopedia of Literature and Politics: Censorship, Revolution, and Writing, Volume II: H–R*, ed. M. Keith Booker (Westport, CT: Greenwood), 601.

29
Is apophenia a new form of paranoia? In 1989, Frederic Jameson declared paranoia to be one of the main cultural patterns of postmodern narrative, pervading the political unconscious. According to Jameson, the totality of social relations could not be culturally represented within the Cold War imagination – and the blanks were filled in by delusions, conjecture, and whacky plots featuring Freemason logos. But after Snowden’s leaks, one thing became clear: all conspiracy theories were actually true. Worse, they were outdone by reality. Paranoia is anxiety caused by an absence of information, by missing links and allegedly covered-up evidence. Today, the contrary applies. Jameson’s totality has taken on a different form. It is not absent. On the contrary: it is rampant. Totality – or maybe a correlated version thereof – has returned with a vengeance in the form of oceanic “truckloads of data.” Social relations are distilled as contact metadata, relational graphs, or infection spread maps. Totality is a tsunami of spam, atrocity porn, and gadget handshakes. This quantified version of social relations is just as readily deployed for police operations as for targeted advertising, for personalized clickbait, eyeball tracking, neurocurating, and the financialization of affect. It works both as social profiling and commodity form. Klout Score-based A-lists and presidential kill lists are equally based on obscure proprietary operations. Today, totality comes as probabilistic notation that includes your fuckability score as well as your disposability ratings. It catalogs affiliation, association, addiction; it converts patterns of life into death by aerial strike.

This type of totality is actually something else, namely a “singularity.” Singularity – the pet myth of California ideology – describes, among other things, a time when artificial intelligences take over. The NSA’s SKYNET program was named after one of the most famous Hollywood singularities, an AI robot gunning for world domination in Schwarzenegger’s cyborg vehicle *Terminator*.

14/14

e-flux journal #72 — april 2016 Hito Steyerl
A Sea of Data: Apophenia and Pattern (Mis-)Recognition

But singularity also means something different: general rules no longer apply to this entity, especially not the rule of law. It’s case-by-case instead, or rather, every-case-for-itself. Accordingly, there seem to be competing singularities. Another, even more powerful singularity of our times is most obviously the semi-divine mythical entity called “the markets,” a set of organizations regarded as both autonomous and superintelligent, of such providence that reason has to bow to its vast superiority. If the myth of the market today can be said to have replaced the myth of socialism, then the actually existing singularities today – government surveillance and market domination, each dominated by singular *laissez-faire* ideas – plus the vast and confusing bureaucracies, oligarchies, quasi-states, informal dictatorships, dark net start-ups, econometric SWAT teams, and unclassifiable paramonopolies sustaining their operations have replaced the actually existing socialisms of the twentieth century: ideological entities spread out across junk space data centers, high-frequency trading exchanges, and vast zones of impunity and violence that create a completely biased and unsustainable distribution of both information and resources.

30
More recent, extremely fascinating examples include Christian Szegedy et. al, “Intriguing properties of neural networks,” arxiv.org, February 19, 2014 <http://arxiv.org/pdf/1312.6199v4.pdf>; and Anh Nguyen, Jason Yosinski, and Jeff Clune, “Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images,” [cv-foundation.org](http://www.cv-foundation.org), 2015 http://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Nguyen_Deep_Neural_Networks_2015_CVPR_paper.pdf. The first paper discusses how the addition of a couple of pixels – a change imperceptible to the human eye – causes a neural network to misidentify a car, an Aztec pyramid, and a pair of loudspeakers for an ostrich. The second paper discusses how entirely abstract shapes are identified as penguins, guitars, and baseballs by neural networks.

31
“Do We Need a Bigger SIGINT Truck?” *Signal v. Noise* column, January 23, 2012.

32
See Jussi Parikka, “The Geology of Media,” *The Atlantic*, October 11, 2013 <http://www.theatlantic.com/technology/archive/2013/10/the-geology-of-media/280523/>

33
Contemporary soothsayers are

reading patterns into data as if they were the entrails of sacrificial animals. They are successors of the more traditional augurs that Walter Benjamin described as photographers *avant la lettre*: “Is not every spot of our cities the scene of a crime? Every passerby a perpetrator? Does not the photographer – descendent of augurers and haruspices – uncover guilt in his pictures?”

There is a decisive distinction, though, between the twentieth-century photographer and the filterers and decoders of the twenty-first. Contemporary pattern extractors are not mainly supposed to recognize the guilty after the fact. They are expected to predict the perpetrator as well as the crime before it ever happens – and to preempt it. Every spot of our cities is mapped out as a probable crime site, fully decked out with gender- and age-based targeted advertising – surveilled by animated commodities, divinatory cellphone cameras, and aerial views from tapped drones.

34
“Israel: Gaza Airstrikes Violated Laws of War.”