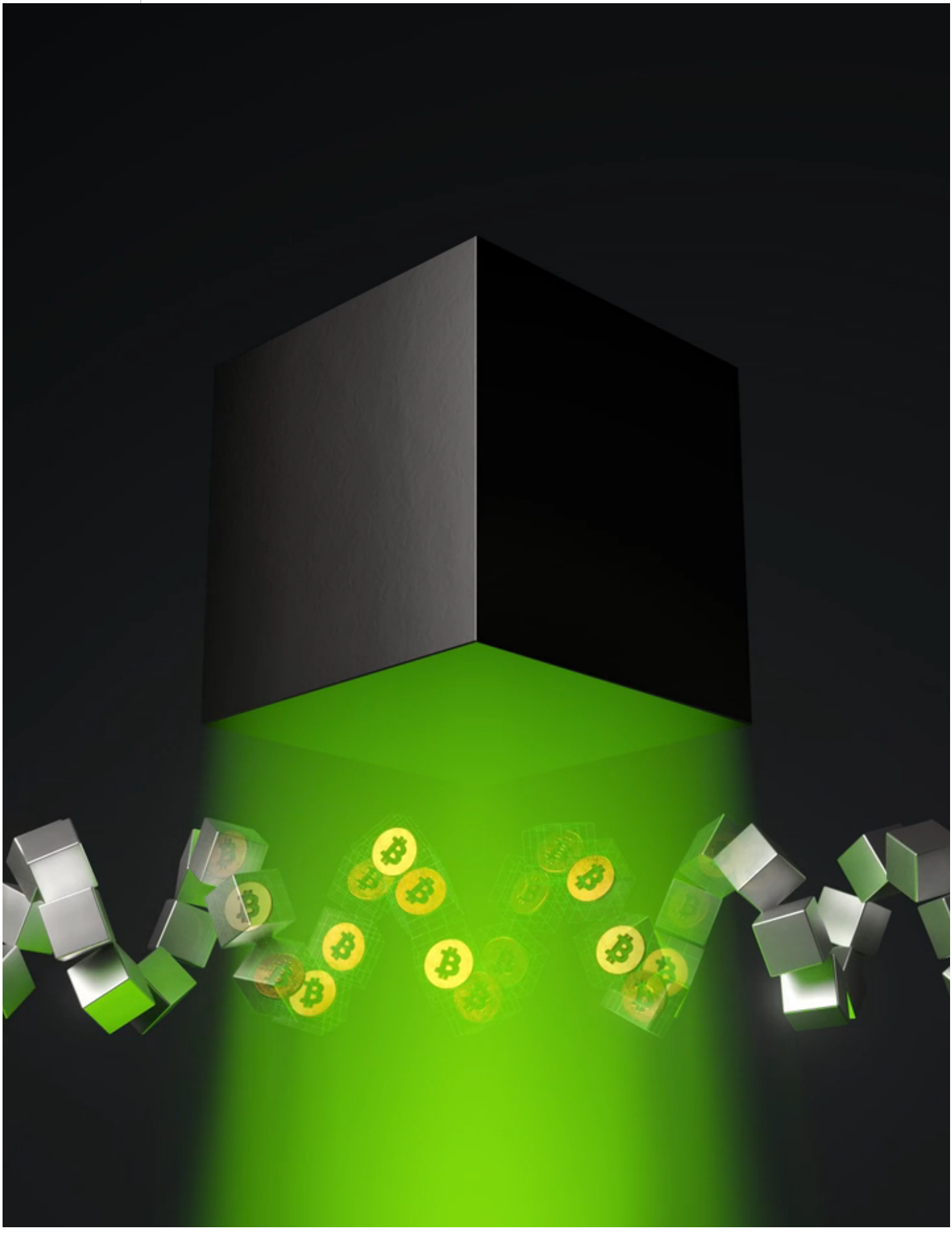


How 'Trustless' Is Bitcoin, Really?

In myth, the cryptocurrency is egalitarian, decentralized and all but anonymous. The reality is very different, scientists have found.

By Siobhan Roberts

Published June 6, 2022 Updated June 7, 2022, 8:51 a.m. ET

Alyssa Blackburn, a data scientist at Rice University and Baylor College of Medicine in Houston, has spent several years performing digital detective work with her trusty lab assistant, Hail Mary, a shiny black computer with orange trim. She has been collecting and analyzing leaks from the Bitcoin blockchain, the immutable public ledger that has recorded all transactions since the cryptocurrency's launch in January 2009.

Bitcoin represents a techno-utopian dream. Satoshi Nakamoto, its pseudonymous inventor, proposed that the world run not on centralized financial institutions but on an egalitarian, math-based electronic money system distributed through a computer network. And the system would be "trustless" — that is, it would not rely on a trusted party, such as a bank or government, to arbitrate transactions. Rather, as Satoshi Nakamoto wrote in a 2008 white paper, the system would be anchored in "cryptographic proof instead of trust." Or, as T-shirts proclaim: "In Code We Trust."

The practicalities have proved complicated. Price turbulence is enough to induce the Bitcoin bends, and the system is environmentally destructive, since the computational network uses exorbitant amounts of electricity.

Ms. Blackburn said her project was agnostic to Bitcoin's pros and cons. Her goal was to pierce the scrim of anonymity, track the transaction flow from Day 1 and study how the world's largest cryptoeconomy emerged.

Satoshi Nakamoto had presented the currency as anonymous: For Bitcoin transactions (buying, selling, sending, receiving et cetera), users employ pseudonyms, or addresses — alphanumeric cloaks that hide their real identities. And there was apparent confidence in the anonymity; in 2011, WikiLeaks announced that it would accept donations via Bitcoin. But over time, research revealed data leakage; the identity protections weren't so watertight after all.

"Drip-by-drip, information leakage erodes the once-impenetrable blocks, carving out a new landscape of socioeconomic data," Ms. Blackburn and her collaborators report in their new paper, which has not yet been published in a peer-reviewed journal.

Aggregating multiple leakages, Ms. Blackburn consolidated many Bitcoin addresses, which might have seemed to represent many miners, into few. She pieced together a catalog of agents and concluded that, in those first two years, 64 key players — some of whom were the community's "founders," as the researchers called them — mined most of the Bitcoin that existed at the time.

"What they figured out, just how concentrated early mining and use of Bitcoin was, that's a scientific discovery," said Eric Budish, an economist at the University of Chicago. Dr. Budish, who has conducted research in this realm, received a two-hour video preview with the authors. Once he came to understand what they had done, he thought, "Wow, this is cool detective work," he said. Referring to those early key players, Dr. Budish suggested that the paper be titled "The Bitcoin 64."

The computer scientist Jaron Lanier, an early reader of the paper, called the investigation “important and significant” in its ambitions and social implications. “The nerd in me is interested in the math,” said Mr. Lanier, who is based in Berkeley, Calif. “The techniques used to extract information are interesting.”

The demonstration of blockchain leakage, he noted, will be surprising to some, not to others. “This thing isn’t hermetically sealed,” Mr. Lanier said. He added: “I don’t think it’s the end of the story. I think there’s further innovation that will take place, extracting information from these types of systems.”

One of Ms. Blackburn’s tactics was simple perseverance. “I kicked it till it broke,” she said, recalling how the principal investigator, Erez Lieberman Aiden, an applied mathematician, computer scientist and geneticist at Baylor College of Medicine and Rice University, characterized her method.

More precisely, Ms. Blackburn developed hacks for the period of time that was of particular interest: from the cryptocurrency’s start to when Bitcoin achieved parity with the U.S. dollar in February 2011, which coincided with the establishment of the Silk Road, a Bitcoin-based black market. She leveraged human lapses such as insecure user behavior; she exploited operational features inherent to Bitcoin’s software; she deployed established techniques for linking the pseudonymous addresses; and she developed new techniques. Ms. Blackburn was particularly interested in miners, the agents who verify transactions by engaging in an elaborate computational tournament — a puzzle hunt, of sorts, guessing and checking random numbers against a target, in search of a lucky number. When a miner wins, they earn Bitcoin income.

Whether 64 seems like a small or large number of key miners depends on one’s proximity to the crypto undertow. Scholars have questioned whether Bitcoin is truly a decentralized currency. From Dr. Lieberman Aiden’s perspective, the population under investigation was “even more concentrated than it seems.” Although the analysis showed that the big players numbered 64 over two years, at any given moment, according to the researchers’ modeling, the effective size of that population was only five or six. And on many occasions, just one or two people held most of the mining power.

As Ms. Blackburn described it, there were very few people “wearing the crown,” functioning as arbiters of the network — “which is not the ethos of decentralized trustless crypto,” she said.

Finding treasures in the data

Alyssa Blackburn, left, a data scientist, and Erez Lieberman Aiden, a geneticist and computer scientist, tested Bitcoin's identity protections and claims of decentralization. Annie Mulligan for The New York Times

For Ms. Blackburn and Dr. Lieberman Aiden, Bitcoin's data — 324 or so gigabytes archived in the blockchain — presented a cache of temptation. Dr. Lieberman Aiden's lab does biological physics and widely applied mathematics; one focus is three-dimensional genome mapping. But as a scholar, he is also intrigued by the use of new kinds of data to explore complex phenomena. In 2011, he published a quantitative cultural analysis using more than five million digitized books from 1800 to 2000, with Google Books and collaborators. "Culturomics," he called it. For instance, the team introduced the Google Ngram Viewer, which lets users type in a word or phrase and observe its usage plotted over the centuries.

In the same spirit, he wondered what treasures might be submersed in Bitcoin's data lake. "We literally have a record of every single transaction," he said. "These are remarkable economic and sociological data sets. Clearly, there's a lot of information in there, if you can get at it."

Getting at it proved nontrivial. Ms. Blackburn was barred from the university's supercomputing cluster — with her file folder labeled "Bitcoin," she was suspected of mining the cryptocurrency. "I objected," she said. She said she tried to convince an administrator that she was conducting research, but "they were completely unmoved."

A key tactic of Ms. Blackburn's was to trace patterns in plots of numbers that in theory should have been random and meaningless. In one case, she was chasing the "extranonce," one piece of the mining puzzle: a short field of 0s and 1s tucked within a longer string that encodes each block, or bundle, of transactions. The extranonce leaked information about a computer's activity. This led Ms. Blackburn to reconstruct the miners' behavior: when they were mining, when they stopped and when they started up again. She speculates that the extranonce's leaky behavior was tolerated because it allowed Bitcoin's creator to keep an eye on miners; the source code was modified to plug this leak shortly before Satoshi Nakamoto disappeared from the public Bitcoin community in December 2010.

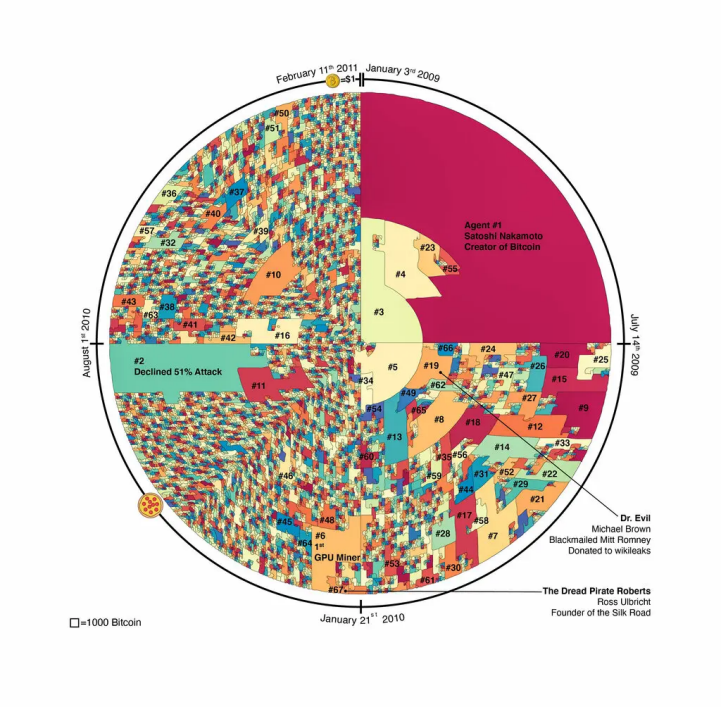
Once Ms. Blackburn had put various toeholds to use — allowing her to erode the identity-masking protections — she began merging addresses, linking nodes on a graph, consolidating the effective population of mining agents. Then she cross-referenced and validated the results with information scraped from Bitcoin discussion forums and blogs. Initially, the catalog of agents who mined most of the Bitcoin tallied a couple of thousand; then it hovered for a while around 200. Ultimately, Hail Mary spit out 64. (Eventually, Hail Mary's brains were incorporated into the lab's computer cluster, Voltron.)

The study's purpose was not to name names; it's the job of the F.B.I. and the I.R.S. to bust Bitcoin criminals. But the researchers pinpointed the identities of a couple of the top players who were publicly known Bitcoin criminals: Agent No. 19 is Michael Mancil Brown, a.k.a. "Dr. Evil," who was found guilty of a 2012 fraud and extortion scheme involving Mitt Romney, then a candidate for president. Agent No. 67 is associated with Ross Ulbricht, a.k.a. "DreadPirateRoberts," creator of the Silk Road. Naturally, Agent No. 1 is Satoshi Nakamoto — whose true identity the researchers did not try to determine.

Mark Gerstein, a professor of bioinformatics at Yale University, found in the research implications for data privacy. He recently stored a genome on a private blockchain, which allowed for a secure and tamperproof record. But he noted that in a public setting, as with Bitcoin's blockchain, a data set's size and subtle patterns made it susceptible to breaches, even as the data remained immutable. (Ms. Blackburn wasn't tampering with the Bitcoin blockchain's records.)

"That's the amazing thing about big data," Dr. Gerstein said. "If you have a big enough data set, it starts to leak information in unexpected ways." Even more so when data from different sources are connected, he said: "When you combine one data set with another to make a bigger data set, nonobvious linkages can arise."

'Decentralization theater'



A map of the bitcoin blockchain constructed by Ms. Blackburn and Dr. Lieberman Aiden using data leakages. "Each agent corresponds to a single map tile, whose area is proportional to the quantity of bitcoin mined by the agent," they noted in their recent paper. Alyssa Blackburn and Erez Lieberman Aiden

Once Ms. Blackburn had assembled the catalog of agents, she analyzed the income they had reaped from mining. She found that within a few months of the cryptocurrency's introduction — and contrary to Bitcoin's egalitarian promise — a classic distribution of income inequality emerged: A small fraction of the miners held most of the wealth and power.

(Mining income demonstrated what is called a Pareto distribution, after Vilfredo Pareto, a 19th-century economist.)

The lab unintentionally replicated this dynamic when they invented "CO2 coin," a cryptocurrency that could be used to buy snacks from a student-run store. In due course, some CO2 miners became more successful than others, and the store marked up snack prices catering to the tastes of the rich.

"The people who had a lot of crypto resources had very strong control over what the store would acquire, which other people didn't feel great about," Dr. Lieberman Aiden recalled. The economy collapsed — that is, there was a revolt — when the shop began charging in CO2 to use the coffee machine.

In the formal study, Ms. Blackburn also observed that the concentration of resources threatened the network's security, with a miner's computational resources being directly proportionate to his or her mining income. On several occasions, individual miners wielded more than 50 percent of the computational power and, as a result, could have taken over like a tyrant using what's called a "51 percent attack." For instance, they could have cheated the system and repeatedly spent the same Bitcoins on different transactions.

Sarah Meiklejohn, a cryptographer at University College London, said that the investigation's findings, assuming they were error-free, provide empirical confirmation of an "intuition that has been floating around in this space for a while." (Dr. Meiklejohn developed some address-linking techniques used in the investigation and recently devised a technique for tracking a type of transaction flow called a peel chain.)

"We all kind of knew that mining was fairly centralized," she said. "There aren't that many miners. This is true even today, of course, and it was even more true at the beginning." As for what should be done about it, "we do need to really examine that question," she said. "How do we make mining more decentralized?" She thought the results of this investigation might encourage the field to take the issue more seriously.

But to add a twist, Ms. Blackburn found that while some miners had the power to execute 51 percent attacks, they repeatedly chose not to. Rather, they acted altruistically — preserving the cryptocurrency's integrity, even though the decentralization-based fraud-prevention mechanism had been compromised.

In parsing this finding, Ms. Blackburn's team turned to the tools of experimental economics. They gathered human subjects online to participate in game-theory scenarios that modeled the "social dilemma" faced by the founders — that is, how people behave when they find themselves as the trustee of an appreciating good.

"In scenarios like this, it appears that people don't like to kill the golden goose — they don't like to spoil it for the group," Dr. Lieberman Aiden observed. Whatever you believe about the motivations of the "Bitcoin 64," he said, the fact that the network was vulnerable to individual decision makers changes the understanding of its security.

"Sure, decentralization protects the blockchain," he said. "But even on occasions when the mining pool became centralized, the dominant miners declined to attack it. That is a very different picture than the idealized model people have for why these cryptocurrencies are secure."

As the authors concluded in the paper: "Although Bitcoin was designed to rely on a decentralized, trustless network of anonymous agents, its early success rested instead on cooperation among a small group of altruistic founders."

For Glen Weyl, an economist at Microsoft Research who was consulted on the research, this finding demonstrates how decentralization played a rhetorical rather than substantive role. "And that rhetorical role was very powerful — it bound together this community, much as other myths have bound together other communities, like nations," Dr. Weyl said; he and Mr. Lanier wrote about this research for CoinDesk. But the myth and the promise, he said, were in tension with the reality that emerged. "It's just fascinatingly ironic, and also predictable, repeating the historical patterns it aspires to erase."

Mr. Lanier called it "decentralization theater." Cryptocurrencies create an illusion: "'Now we're in utopia. Everything's decentralized. Everybody's equal.' There's this notion of democracy without annoyance."

But, he said, these systems end up hiding a new elite, which is probably just an old elite in a new arena. And the technology cuts both ways. "Whatever you think you can achieve using new algorithms, or big data, or whatever, can also be used against you," Mr. Lanier said. "The same algorithms can be used by scientists to interrogate and investigate these castles

that are put up by the new elite."

One moral of the story, Ms. Blackburn said, is simply: "You have to be careful." There is a limited timeline for encryption, "a horizon beyond which it will no longer be useful. When you are encrypting private data and making it public, you cannot assume that it'll be private forever."