

Dangerous Protocols

Nadia Asparouhova



Summer of Protocols | 2023

© 2023 Ethereum Foundation. All contributions are the property of their respective authors and are used by Ethereum Foundation under license. All contributions are licensed by their respective authors under CC BY-NC 4.0. After 2026-12-13, all contributions will be licensed by their respective authors under CC BY 4.0. Learn more at: summerofprotocols.com/ccplus-license-2023

Summer of Protocols :: summerofprotocols.com

ISBN-13: 978-1-962872-11-9 print
ISBN-13: 978-1-962872-37-9 epub
Printed in the United States of America
Printing history: March 2024

Inquiries :: hello@summerofprotocols.com

Cover illustration by Midjourney :: prompt engineering by Josh Davis
«warning symbol, protocols, exploded technical diagram»

Protocols are touted as a liberating alternative to the walled technological gardens that govern our lives: an elusive utopia—centered around user freedoms instead of commercial interests—that we, poor lost sinners, have been trying to find our way back to since the internet’s genesis.¹

A force of liberation and control

Certainly, protocols have the *potential* to liberate us, offering a level of interoperability and customization that platforms don’t. It doesn’t feel good to be controlled, especially by technology, and platforms show no shame in flexing their might over us. Twitter’s most tyrannical moments following its most recent transfer of ownership—censoring Substack links, throttling reading and messaging features for non-paying users—drove waves of devoted fans to seek shelter elsewhere, whether Mastodon, Bluesky, or Threads. Protocols are the infallible mistress we imagine ourselves escaping to, offering a richer, freer, and more exciting life than the dreary, drudgerous, doomscrolling serfdom that we’re confined to today.

I think this is fantasy. A historical look at protocols suggests that their purpose has always been to *simplify* coordination and communication. With enough time, protocols converge upon conformity. They do not liberate us, but rather seek to control us completely.

Protocols demand not just our compliance, but our *loyalty* in relinquishing our decision-making power to a formless entity. And because protocols are not owned or mediated by a central authority, if you

don’t like the protocol you’re in, escape is not as “easy” as switching platforms (the latter already, as we’ve seen, is seemingly impossible).

Protocols are not *necessarily* bad, however. Sometimes, relinquishing control enables us to spend precious brain power on other, more pressing decisions. We need protocols to help us manage complexity. Protocols are better understood as neither uniformly harmful nor liberating, but—as writer and computer programmer Alexander Galloway once put it—“dangerous.”² Because they make our lives simpler, it is always tempting to outsource more decisions to protocol. Like the sorcerer’s apprentice in Disney’s *Fantasia*, though, when left unchecked, protocols threaten to subsume human agency.

It is particularly important to acknowledge the dangerous nature of protocols today, as the social protocols that govern our behavior are becoming increasingly *implicit*, or embedded into our consciousness, and therefore harder to detect or resist. In the software world, when we think of protocols, we think of technical examples: TCP/IP, SMTP, Bitcoin, Ethereum. But “protocolization” pervades our social worlds even more strongly.

These technical protocols are comparatively easy to identify and discuss, so we might overlook other protocol-like things we find ourselves embedded in. But protocols have a much longer history than software, spanning many applications and industries. Although they may seem different from the (computer-mediated) protocols we’re familiar with, all share similar characteristics that make them distinctly *protocological*.

1. As I explore later, I don’t actually think platforms should be classified as an alternative to protocols; rather, they provide the hard constraints that enable protocols to emerge.

2. “... it’s not that protocol is bad but that protocol is *dangerous*.” Alexander Galloway, *Protocol: How Control Exists after Decentralization* (Cambridge, Mass: MIT Press, 2004), p. 16.

Protocols are primarily social, not technological, in nature

Protocolized thinking boomed with the onset of industrialization, or what historian James Beniger called the *control revolution*, a need to regain control over the sudden explosion of information created by new manufacturing processes:

For centuries most goods had moved with the speed of draft animals down roadway and canal. . . . This infrastructure, controlled by small organizations of only a few hierarchical levels, supported even national economies. Suddenly—owing to the harnessing of steam power—goods could be moved at the full speed of industrial production . . . To [move these goods], however, required an increasingly complex system of manufacturers and distributors, central and branch offices, transportation lines and terminals, containers and cars.³ [emphasis mine]

Protocols give us a predetermined way to manage large amounts of unorganized information, reducing the need for redundant decision making. Standardized forms and TCP/IP, for example, are protocols for collecting and handling data. Management theory and organizational governance—whether hierarchy or holocracy—are protocols for coordinating people. International diplomacy relies on protocols to maintain a standard of etiquette and respect. Aptitude testing is a protocol for measuring the quality of a candidate for school or work.

I define protocols as *procedural systems of social control that simplify communication between actors*:

- *Protocols reduce the number of decisions that need to be made*: Protocols dissuade participants from reason, which

is outsourced to the protocol. Instead of extensive research on each candidate's academic background, for example, university admissions officers use grades and SAT scores as a proxy.

- *Protocols gain legitimacy from participation*: Like other socially entrenched norms, protocols persist because they have persisted; the more we use them, the stronger they get. We then find ways to justify their legitimacy because we participate in them. COVID masking protocols were first adopted because they were a simple path to feeling safe; over time, some participants felt it was not just their duty, but a *virtue* to follow protocol.
- *Protocols resist central management*: Protocols may have an original creator or claim to be managed by an organization, but as a protocol's legitimacy grows, it becomes increasingly difficult to influence.⁴ The Ethereum Foundation and Ethereum's inventor, Vitalik Buterin, for example, make it clear that they don't "run" Ethereum, but are just one among many actors supporting its development.

Open-source projects too, such as Python and its creator, Guido van Rossum, may start with

a "benevolent dictator for life," but transition to community ownership as they become more powerful.

- *Protocols accomplish a function, but not a purpose*: While protocols might have been designed with certain goals in mind, as they mature it becomes increasingly difficult to identify an agenda that drives their persistence. Protocols are a way to accomplish a goal, which may deviate from the explicit goal itself. Paperwork may follow protocol, for example, but it is not always an efficient path to accomplishing the participant's intended goal.

Protocolization 2.0 is defined not by the protocolization of data, but of ideas.

3. James R. Beniger, *The Control Revolution: Technological and Economic Origins of the Information Society* (Cambridge, Mass.: Harvard University Press, 1986), p. 12.

4. At its worst, this is akin to Jo Freeman's "tyranny of structurelessness." www.jofreeman.com/joreen/tyranny.htm

Protocols are frequently seen as a tool that we, the fully agentic human, can design, wield, or take away if needed. Some blockchain governance enthusiasts like to point out the ways in which technical protocols can improve social coordination, suggesting perhaps that governance could be enhanced by the inclusion of technology such as immutability and verifiability. I think it is the other way around. We do not control protocols; but with enough time and participation, protocols control us.

Protocols are ultimately enforced by the self

What compels participants to follow protocol? Protocols derive their initial authority from external mechanisms, but as they become more entrenched, are unconsciously internalized by the self.

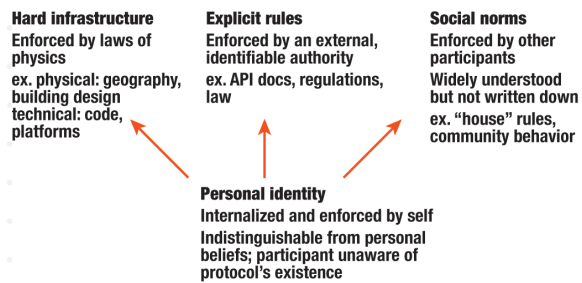
External forces—whether hard infrastructure, explicit rules, or social norms—provide the scaffolding for protocols to emerge. Just as a deck of playing cards provides the hard constraints for many games, protocolization “creates new ‘water’” in which we swim.⁵

But as protocols become more entrenched, they seep into our identity layer. Over time, participants don’t comply with protocols because they fear consequences from an outside authority, nor being shunned by their peers. Instead, participants believe that following protocol is a form of *self-expression*, and the protocol’s legitimacy becomes indistinguishable from their own beliefs.

When someone takes a Myers-Briggs personality test for work or SATs for admission to university, they don’t see themselves as being *told* who they are, but rather a reflection of their *true* selves. “I was born to be a leader, that’s why I came out as an ENTJ.” “Of course I’m good at math, that’s why I my math SAT score was 800.”

5. Venkatesh Rao, “Magic, Mundanity and Deep Protocolization: The next world-transformation process is here,” July 1, 2023. studio.ribbonfarm.com/p/magic-mundanity-and-deep-protocolization

What drives protocolization?



Driving force	Type of expression	Type of authority	Enforcement mechanism
Hard infrastructure	Explicit	Centralized	Laws of physics
Explicit rules	Explicit	Centralized	Fear of punishment from authority “It is my duty to comply”
Social expectations	Explicit (somewhat)	Decentralized	Social sanctions, enforced among participants “We ought to comply”
Protocol as identity	Implicit	Internalized	Internal drive, personal identity “It is a virtue to comply”

Waiting in line for a highly anticipated movie, festival, or product release is another example of how protocols become internalized. Participants wait in line because, frankly, they have to. But camping out for the midnight release of the latest *Star Wars* movie, dressed in costume with friends, and vying to be first in the door eventually becomes part of one’s identity, rather than rote compliance with the rules of the movie theater.

Explicit vs. implicit protocols

Because protocols abstract away complexity, they’re most effective when we don’t think about them. In earlier stages of entrenchment, however, protocols are still *explicit*. All participants know what the protocol is and willingly enter under its governance. Explicit protocols frequently arise in connection with social institutions—school, family, work. They are comparatively easier to identify and manipulate, because they are

stated and enforced by physical constraints or a central authority.⁶

But just as antibacterial products can lead to the evolution of resistant bacteria, so do protocols learn to avoid our collective immune system, emerging in a more powerful form that are harder to discern or avoid: *implicit* protocols.

Implicit protocols are fully internalized by participants who are unaware that they are operating by protocol, just as we're unaware of the gut bacteria that inhabit our bodies and influence our minds. There is no visible authority; instead, participants believe they are expressing themselves as individuals. When someone broadcasts their views on social media about a controversial topic, for example, they may genuinely believe these views to be their own, rather than a reflection of protocolized thinking.

Whereas Michel Foucault's *discipline societies* or Emile Durkheim's *social facts* (enforced respectively by the omnipresent panopticon or by institutions such as religion or family) highlight the *external* forces that control individual behavior, the protocolization of identity represents an *internalized* drive to conform, which the individual sees as a form of self-expression rather than compliance. It is this final stage—the protocolization of identity—that presents a new and dangerous chapter of protocol history today.

From bureaucratic to crisis mindset: Protocolization 2.0

As Beniger underscored in *The Control Revolution*, the onset of industrialization created a flood of information that needed to be managed. From the late 1800s through the mid-20th century, a growing and widespread obsession with organizing information reflected an internalized *bureaucratic*

mindset that pervaded every aspect of society and which led to the rise of protocolized thinking.

But even Beniger, writing in 1986, could not have fully anticipated how mass adoption of the internet, a decade later, would create an even bigger tidal wave. With the creation of the social web (“Web 2.0”) in the 2000s came another deluge of information that continues to grow exponentially.

The bureaucratic Protocolization 1.0 mindset, applied today, would emphasize the manipulation and management of social data itself: privacy, data ownership, surveillance. If protocols are primarily social, not technological, phenomena, however, it becomes more important to focus on what has followed not from the data itself—packets moving through cyberspace—but instead by what the data contains: *opinions*. Protocolization 2.0 is defined not by the protocolization of data, but of *ideas*.

We now have unlimited, infinite access to off-the-shelf takes on everything from reusable shopping bags to European geopolitics, at a magnitude we've never experienced before. The prevalence of freely available ideas, for most people, cheapens the value of original thinking. Instead of having to reason about our own beliefs, we can simply borrow others' ideas and share them as our own, aided by recommender systems that automate curation. The “takes” of others have become social currency that we use to barter with others.

This is, in my view, rational behavior. In many situations, the protocolization of ideas is efficient and therefore beneficial, enabling us to process abstractions quickly. The cost of developing an original idea is expensive compared to having free and reasonably vetted ideas at one's fingertips. Just as, thanks to industrialization and economies of scale, most people don't own and raise cows for meat and dairy, most people no longer can be expected to do the hard work of cultivating opinions of their own. As historian and philosopher Justin E. H. Smith writes:

6. Peer enforcement is less explicit than these two forces, but because participants are the ones spreading the protocol to others, they are still able to articulate what it is and why it's important to the group.

It [seems] to me recently that this present moment must be to language something like what the Industrial Revolution was to textiles. A writer who works on the old system of production can spend days crafting a sentence [...] only to find, once finished, that the internet has already produced countless sentences that are more or less just like it...⁷

In situations where creativity and novelty are needed to make progress on thorny problems, however, the protocolization of ideas can be limiting. Writer and researcher Aaron Z. Lewis calls this phenomenon “solved conversations,”⁸ a term derived from so-called solved games like chess—where all possible moves are known or knowable—versus games like poker, where imperfect information makes it computationally impossible to solve the game:

Solved conversations [are] like the perfect players in a computer model, following an algorithm neatly mapped out for them by their culture.

Conversation about the weather or a crazy weekend party may be boring, but they don’t seem to do any real harm. *What’s*

dangerous are solved conversations about things that really matter. When people identify deeply with a political party or a particular philosophy, it’s not difficult to predict the outcome of any given conversation. *[emphasis mine]*

Just as the Industrial Revolution created a tidal wave of unorganized information, followed by an attempt to regain control, so has the birth of the social web caused a tidal wave of unorganized ideas. Social platforms, combined with explicit rules around content and behavior, created the conditions from which a new set of social protocols emerged. In the 2010s, the strength of memetic

culture meant that the spread of, and compliance with, certain ideas was aggressively enforced by peers.

Conversational AI can be thought of as the productization of these solved conversations, promising us less work in exchange for our reason. It is the aggregation of millions of ideas, packaged in the form of a recommendation. It’s self-learning, beyond the direct control of humans, which means that as it becomes more powerful, nobody will know exactly what goes into it. “Nobody is responsible for the protocol.”

For a while, it was easy enough to see when an idea signaled an affinity with one’s preferred tribe—an external social force. But today, protocolized ideas have become part of our personal identities.⁹ No one talks about filter bubbles anymore. We don’t perceive ourselves as living in a filtered world; instead, each of us are in the truest and most important world. Because these ideas are so personally cherished, yet

spread so quickly, they feel urgent. Instead of Protocolization 1.0’s bureaucratic mindset, we are now governed by Protocolization 2.0’s

crisis mindset.

I’ve previously written about doomer industries as a byproduct of the social web: climate change, AI safety, misinformation, and population decline being a few examples of networks that formed around the threat of an apocalyptic event.¹⁰ While moral panics, especially in the realm of technology, are nothing new, a difference in recent years is how much these fears have been legitimized and codified into industries. Rather than a short-term, emotional

Conversational AI can be thought of as the productization of these solved conversations, promising us less work in exchange for our reason.

7. Justin E. H. Smith, “It’s All Over,” *The Point*, January 3, 2019. thepointmag.com/examined-life/its-all-over
8. Aaron Z. Lewis, “Solved Conversations,” June 1, 2014. aaronzlewis.com/blog/2014/06/01/solved-conversations

9. Smith, “It’s All Over”: “Who has not found themselves thrust into the uncomfortable position . . . of being told that what we thought were our considered beliefs are in fact something else entirely?”
10. Nadia Asparouhova, “Mapping out the tribes of climate. Addendum: ‘Doomer industries’ and the search for meaningful work,” November 30, 2022. nadia.xyz/climate-tribes#addendum-doomer-industries-and-the-search-for-meaningful-work

cycle of collective panic that burns itself out, today's crisis mindset is paradoxically stable. Just as managing a big corporation was once the big career dream, working dutifully towards the prevention of a major crisis—real or imagined—is seen as a laudable choice. Rather than intense, focused public attention on a single crisis, multiple crises coexist alongside one another.

Accepting the inevitability of a major crisis establishes a frame of control, enabling social protocols to arise within each ecosystem. (*Recycle! Reduce energy! Break up Big Tech! Make more babies!*) Rather than stepping up to *do one's duty*, as the

Protocolization 1.0 bureaucratic mindset dictated in social institutions such as the military, patriotism, school, or family, the Protocolization 2.0

crisis mindset is a form of self-expression (*doing meaningful work*).

Similarly to the 20th century's bureaucratic mindset, today's crisis mindset is an efficient, protocolized way to organize people toward directed outcomes. It is not that we need to do away with the crisis mindset—it is probably here to stay for a while. Rather, by bringing awareness to its existence and to the extent to which it governs our lives, we might consider how to use this way of thinking to accomplish big social goals of our *own* choosing, rather than those determined by passive participation—just as technocrats wielded the bureaucratic mindset to accomplish big things in the mid-20th century.

Subverting protocols

While a degree of protocolization is unavoidable and even beneficial to our lives, once we become aware of how protocols can govern us in undesirable ways, the natural question of how to escape them arises.

Subverting protocols is difficult without exiting the network the protocol belongs to. Exiting the protocols followed by members of the Church of Jesus Christ of Latter-day Saints, for example, requires leaving the church which is socially expensive because it also means leaving behind one's friends and family. We can consider this the nuclear option, chosen only when no other solutions exist.

If one remains within the network, explicit resistance to protocols tends to be ineffective, because there is no clear authority to overthrow.¹¹ Attempting to gain control over implicit protocols feels

like shadowboxing. Logic doesn't work against other participants, nor ourselves, as we are the ones vehemently defending our captors.

Galloway in *Protocol* asks us to imagine a town that needs to solve a speeding problem. They could either install speed bumps or they could pass laws to reduce the legal speed limit and increase police radar surveillance. While the latter may seem more “protocol-like” because it uses technology and explicit protocols, Galloway asserts that the former is more protocoological. Signage and police compel drivers to slow down with “what amounts to nothing more than a polite request.” With speed bumps, on the other hand, the responsibility of compliance rests upon the driver (participants), rather than an external authority (the police):

With bumps, the driver *wants* to drive more slowly. With bumps, it becomes a virtue to drive slowly. But with police presence, driving slowly can never be more than coerced behavior.¹²

What's more, participants are the ones who advocate for speed bumps—even

11. The presence of a visible authority is why centrally governed platforms make for more convenient villains than protocols, despite the latter being more controlling.

12. Galloway, *Protocol*, p. 241.

though there is evidence to suggest that they *increase* overall lives lost by hindering emergency vehicles transporting patients in crisis.¹³ And yet, speed bumps persist as a desirable traffic protocol—not because anyone tells us they’re good, but because they’re earnestly *believed* to be good.¹⁴

Protocol tai chi

Instead, the art of protocol subversion looks more commonly like tai chi, where participants follow protocol, but in a way that better suits their needs. As Galloway proposed, “it is *through* protocol that one must guide one’s efforts, not against it.”¹⁵

There is a traffic light by my house that almost never turns green. Unfortunately, I also need to turn left at that light nearly every time I need to go somewhere. Breaking protocol would mean running the red light and turning left. Working *through* the protocol means finding a way to bend, but not break the rules. I can turn right on red first, then make a legal U-turn to continue in my desired direction. In doing so, I am still bound by traffic protocol, while also satisfying my needs.

Acts of protocol subversion are everywhere. Teenagers who are forbidden to pass notes in class instead use Google Docs’ comment features.¹⁶ The unforgiving nature of online dating apps and their filters have led some single people to publish “Date Me” docs or offer cash bounties for a successful match.¹⁷ Disney’s introduction of the

Fastpass system, intended to reduce wait times for rides at theme parks, instead led enthusiasts to collect and hoard Fastpasses earlier in the day for later use.¹⁸ In each of these cases, participants don’t exit the underlying protocol, but rather find a way to work through it.

The evolution of the social web is another example of protocol tai chi. Exiting the protocol would require quitting social media entirely, but “Delete Facebook” and other public campaigns are faded trends. Recent Twitter exoduses drove users to other services like Mastodon, but didn’t help them quit the underlying social *protocol* of externalizing one’s thoughts to a wider audience. Instead, these initial rebellions gave way to quieter forms of subversion, such as group chats, newsletters, private messaging, and semi-private social media apps like BeReal and Retro. Anonymous accounts enable social media users to say what they really think.

With protocol tai chi, we participate in social media, but we make ourselves less detectable. The spiciest conversations have been pushed down a few layers into the cozy web, or

the private, gatekeeper-bounded spaces of the internet we have all retreated to over the last few years.¹⁹

Mimetic culture is giving way to “vibe culture,” as Peter Limberg eloquently expressed in his “Meme to Vibe: A Philosophical Report”:

Memes are cultural information the mind perceives, coupled with an urge to replicate them. *Vibes* are cultural “exformation” the

-
13. Katarzyna Kosakowska, “Evaluation of the impact of speed bumps on the safety of residents—selected aspects,” *Transportation Research Procedia* 60 (2022), pp. 418-423. doi.org/10.1016/j.trpro.2021.12.054
14. *Tricksters*—think Loki or Prometheus from mythology or the boy who cried that “the emperor has no clothes”—can help participants get “unstuck” from a protocol. As outside observers, tricksters bring awareness to protocols.
15. Galloway, *Protocol*, p. 17.
16. Taylor Lorenz, “The Hottest Chat App for Teens Is . . . Google Docs,” *The Atlantic*, March 14, 2019. www.theatlantic.com/technology/archive/2019/03/hottest-chat-app-teens-google-docs/584857
17. Jenny Gross and Livia Albeck-Ripka, “Tired of Dating Apps, Some Turn to ‘Date-Me Docs,’” *New York Times*,

-
- August 2, 2023. www.nytimes.com/2023/08/02/style/date-me-docs.html
18. Cory Doctorow, “Now you’ve got two problems (Part III): Amusement parks, crowd control and load-balancing,” July 25, 2021. doctorow.medium.com/now-youve-got-two-problems-part-iii-45e1328c5ae1
19. Maggie Appleton, “The Dark Forest and the Cozy Web: An illustrated diagram exposing the inner layers of the dark and cozy web,” 2020. maggieappleton.com/cozy-web

body receives, coupled with a choice to experience them.²⁰ [emphasis mine]

What are vibes? They're hard to describe: that's the point. Vibes govern our public behavior (e.g., "good vibes," "the vibes are off"), but participants feel they are the ones curating and discerning culture, rather than passively consuming or spreading it.

In contrast to meme culture's doom-scrolling and pornification (or "the phenomena of everything being commodified and shamelessly put on display"), vibe culture is about "experiencing, not replicating," a way of "protecting one's mind while online." If the (largely overrated) battle of explicit protocols is between *social media platforms vs. [technical] protocols*, the hidden battle of implicit protocols is between *memetic vs. vibe culture*.

There are many efforts to create a Twitter alternative—Bluesky, Farcaster, Substack Notes, Mastodon, Nostr, Are.na, Threads. However, without sufficient innovation at the "hard infrastructure" layer, we should expect the same behaviors to inevitably emerge at scale, no matter how new or different each option seems to be.

Modifying and creating new protocols

Creating new protocols is difficult for reasons that are probably intuitive. Like creating a new standard or religion, creation is easy, but adoption is hard. Because implicit protocols are emergent properties of external forces, they cannot be directly modified. Modifying these underlying forces, however, can create downstream effects that eventually result in new protocols.

There are many efforts to create a Twitter alternative—Bluesky, Farcaster, Substack Notes, Mastodon, Nostr, Are.na, Threads. However, without sufficient innovation at the "hard infrastructure" layer, we should expect the same behaviors to inevitably emerge at scale, no matter how new or different each option seems to be. Otherwise, it might be more accurate to characterize such efforts as competing *platforms* that

serve as new interfaces for the same social protocols.

To modify these underlying social behaviors, we need to radically modify the infrastructure itself. Substack newsletters and group chats took off as incumbent social platforms became more hostile, because participants wanted a place to be themselves without judgment or fear of repercussions. But these alternatives work best by not trying to mimic their predecessors at all.

Another successful example of creating new protocols is the fragmentation of climate action. For the last few decades, the climate crisis response was based on energy *scarcity*: the belief that reducing energy usage was needed to bring us to "sustainable" levels. In recent years,

some people, such as policy researchers Eli Dourado and Austin Vernon, began to ask: *What if we don't assume energy scarcity, but instead seek energy abundance?*²¹ Reframing the climate narrative this way required breaking from the existing network and forming a new one—where participants tacitly identify as "energy" rather than "climate" advocates—that enabled a different set of behaviors to develop, such as revival of the once-defunct nuclear movement and investment into geothermal energy.

What do good protocols look like?

With a bit of reflection, each person could likely make their own list of protocols that they'd consider good or bad. It's attempting to convert these into heuristics that makes

20. Peter Limberg, "Meme to Vibe: A Philosophical Report," March 06, 2023. lessfoolish.substack.com/p/meme-to-vibe-a-philosophical-report

21. Austin Vernon and Eli Dourado, "Energy Superabundance: How Cheap, Abundant Energy Will Shape Our Future," June 30, 2022. www.thecgo.org/research/energy-superabundance

it clear how hard it is to objectively evaluate the “goodness” of protocols.

Purpose: Fulfilling the participant’s desires

Whether a protocol is dangerous depends firstly on a prospective participant’s goals, and which parts of their brain they want to automate away. Most people would probably agree that they’d prefer not to fill out unnecessary paperwork. But things get fuzzier as we move up the hierarchy of needs. Some people are happy to give up their “thinking” agency and parrot secondhand opinions, while others prefer to retain their autonomy. Some people are happy running macOS, while others prefer the autonomy of Linux.

Bryan Johnson’s Blueprint Protocol is a set of nutrition, fitness, and measurement practices he’s chosen to reduce the number of health-related decisions made daily.²² Johnson’s father, who lost 25 pounds in three months following the plan, explained that “removing me from myself has been the best thing I’ve done in decades.” While not everyone would choose to opt into Blueprint if they value personal choice in diet and fitness, it could be beneficial to those who choose to follow the plan.

One way to evaluate the “goodness” of a protocol, then, is to compare its outcomes to the problem that participants believe it solves. For example, we might intuitively state that historical resistance to handwashing and other sanitation practices was a “bad” protocol, whereas the practice of handwashing today is a “good” protocol. But in both cases, the protocol was perfectly “good” in terms of function. Both resistance and adoption were widespread, and defended by individuals who believed they were acting in the public interest. We only know that the latter is “good” because it accomplished the public health goal of keeping more people safe and healthy. Similarly, fierce debates over the “goodness”

of COVID masking protocols derive from evaluating the protocol against two conflicting goals: personal choice versus public health.

Function: Bad protocol design (and the Kafka Index)

Next, we might consider the “goodness” of a protocol’s design, regardless of participants’ goals. Bad protocols create more complexity, instead of abstracting complexity away. Does the machine accomplish its function or do its buttons fail to respond, its cogs spinning out of place?

Note that “good vs. bad” in this context refers to the elegance and fluidity of a protocol’s design, rather than a subjective or normative evaluation. Solved conversations are a “good” protocol in the sense that they operate so smoothly that we don’t notice them, which is different from the ethical question of whether it’s “good” that our conversations have become so automated.

This brings us to the *Kafka Index*: a set of evaluative criteria for identifying bad protocols (see table next page).²³

Stable, yet suboptimal protocols

Because nobody is responsible for bad protocols, participants can be trapped indefinitely in suboptimal outcomes. Eliezer Yudkowsky called these “inadequate equilibria” in his book of the same name²⁴—“a sticky, stable equilibrium of *everyone* acting insane in a way that’s secretly a sane response to everyone else acting insane”—where civilizations get stuck at a local maximum:

[T]he *frustrating* parts of civilization are the times when you’re stuck in a Nash equilibrium that’s Pareto-inferior to *other Nash equilibria*. . . . What makes an

22. Bryan Johnson, “Blueprint Protocol.” protocol.bryanjohnson.com

24. Eliezer Yudkowsky, *Inadequate Equilibria: Where and How Civilizations Get Stuck* (Berkeley, Calif.: Machine Intelligence Research Institute, 2017). equilibriabook.com

Kafka Index

Evaluative criteria for identifying bad protocols

- ❑ **No (or hidden) feedback loop**
 - ❑ Lack of consequences for failed outcomes
 - ❑ Outcomes aren't visible to participants
 - ❑ No evaluative metrics, or wrong metrics prioritized
- ❑ **Too many edge cases addressed at once**
 - ❑ Binary success response; participant required to pass through all use cases sequentially
 - ❑ No branching or forking of use cases
- ❑ **No happy path to follow**
 - ❑ Protocol increases the number of decisions that participant must make
 - ❑ User error is possible (multiple ways to "plug it in")
- ❑ **Success outcomes are randomized or ambiguously defined**
 - ❑ Outcomes succeed or fail inexplicably, even when all inputs appear to be the same
 - ❑ Outcomes can't be debugged or explained retrospectively by participants
- ❑ **Multiple protocols exist that attempt to solve the same problem**
 - ❑ Redundant protocols create conflict and confusion regarding the desired outcome
- ❑ **Recursive, nested protocols**
 - ❑ Protocol's complexity is sprawling, with multiple dead ends
 - ❑ Participants can get trapped in endless loops or "whirlpools" with no resolution
- ❑ **No market or alternatives exist**
 - ❑ High cost to participate, with no other options available
 - ❑ Significant costs incurred if participants defect

Applying the Kafka Index

Interactive voice response (IVR) systems for customer service
too many edge cases addressed at once;
no happy path; recursive nested protocols

Overcriminalization, or excessive laws that criminalize civilian behavior, even when no criminal intent exists
no feedback loop; ambiguous success outcome;
redundant and conflicting protocols; no alternatives

Airport security
no feedback loop; randomized success outcome;
conflicting protocols; no alternatives

3 archetypes of protocol "goodness"

Kafka — failure archetype
"I can't find my way round in this darkness."

Protocol holds too much power.
Participant is trapped in a maze that they can't understand and also can't escape.

Bartleby — failure archetype
"I would prefer not to."

Participant holds too much power.
Maintaining a high level of agency limits their ability to manage increasingly complex tasks.

Whitehead — success archetype
"Civilization advances by extending the number of important operations which we can perform without thinking about them."

Balanced power between protocol and participant. By relinquishing some agency, participants are able to accomplish much more than they could alone.

Thanks to Rafa Fernandez and Eric Alston for coining this term. The Kafka Index was a collaborative effort, initially developed with Rafa, Seth Killian, Drew Austin, and others at the July 2023 Summer of Protocols retreat.

equilibrium *inadequate*, a fruit that seems to hang tantalizingly low and yet somehow our civilization isn't plucking, is when there's a better *stable* state and we haven't reached it.

Being trapped in protocols dictated by a functional-yet-suboptimal system feels eerily calm, yet unsatisfying. Everything works, sort of, but participants feel a curious lack of fulfillment. (Remember that protocols are designed to accomplish a function, but not a purpose.)

Examples include the United States medical system, which Yudkowsky calls "the most broken system *that still works* ever recorded in human history," and academic science, which is stuck in "a Nash equilibrium that it wandered into, which includes statistical methods that were invented in the first half of the 20th century and editors not demanding that people cite replications." No single entity *caused* these protocols to end up where they are today, but no one participant can influence or defy it, except at personal cost.

What should we expect from our protocols?

At their best, protocols help us reduce complexity and unnecessary decision-making, reduce transaction and coordination costs, and give participants the same context quickly and easily, as many emergency response and diplomatic protocols do, for example. Protocols can liberate us from decisions we don't need or want to make, or that we lack the authority to make confidently, capturing and codifying consensus in human knowledge.

Protocols can be *dangerous* when they become so deeply ingrained that we fail to recognize they are there, defending their virtues as if they were our own. It's hard to evaluate protocols that run silently in the background, because that is also how they function best.

But when protocols enter the identity layer, they can threaten successful outcomes

that *require* creativity and agency over automation (e.g., knowledge work, policy design, debate).

Protocols are dangerous precisely because they control us so well. Though it may seem contradictory, the more powerful a protocol is, the harder it is to understand or explain it to others. Unfortunately, this quality is present in both good and bad protocols.

Protocols are *outright bad* when they no longer fulfill a participant's goals, but the participant is also unable to exit.

In this situation, participants typically suffer from a "fog of war" problem, where feedback loops no longer exist to evaluate a protocol's ongoing benefit. Instead, they remain trapped in a functional, yet suboptimal system. Thus we might consider three archetypes of protocol "goodness"²⁵:

- Kafka — undesirable — "*I can't find my way round in this darkness.*"²⁶ Protocol holds too much power. Participant is trapped in a maze that they can't understand and also can't escape.
- Bartleby — undesirable — "*I would prefer not to.*"²⁷ Participant holds too much power. Maintaining a high level of agency limits their ability to manage increasingly complex tasks.
- Whitehead — desirable — "*Civilization advances by extending the number of important operations which we can perform without thinking about them.*"²⁸ Balanced power between protocol and participant. By relinquishing some agency,

25. Thanks to Venkatesh Rao for prompting this train of thought.

26. Franz Kafka's writing often features protagonists who are trapped in surreal predicaments that defy comprehension such as in *The Trial*, where the main character is arrested, tried, convicted, and executed for an unknown crime.

27. Name derives from Herman Melville's short story, "Bartleby, the Scrivener." Bartleby is a copywriter who refuses to conform to the expectations of his job, preferring instead to gaze out the window at a wall. Eventually, he starves and dies for reasons his peers cannot understand.

28. Alfred North Whitehead was a mathematician who described reality primarily in terms of processes, rather than stable objects.

participants are able to accomplish much more than they could alone.

The controlling nature of protocols is consistently omitted from rosy imagined futures about their merits. It is difficult to notice how much they control us, because control is only perceived when our desires conflict with those of protocol.

Their influence is especially hard to detect today, as the social web created a new era of protocolization—aimed at simplifying the complexity of ideas, rather than just information—and we are just beginning to transition from Protocolization 1.0’s *bureaucratic mindset* to Protocolization 2.0’s *crisis mindset*.

The bureaucratic mindset now feels like a vaguely dated relic from the industrial era, but looking forward, we must now learn to build our immunity towards—and make the most of—our collective crisis mindset.

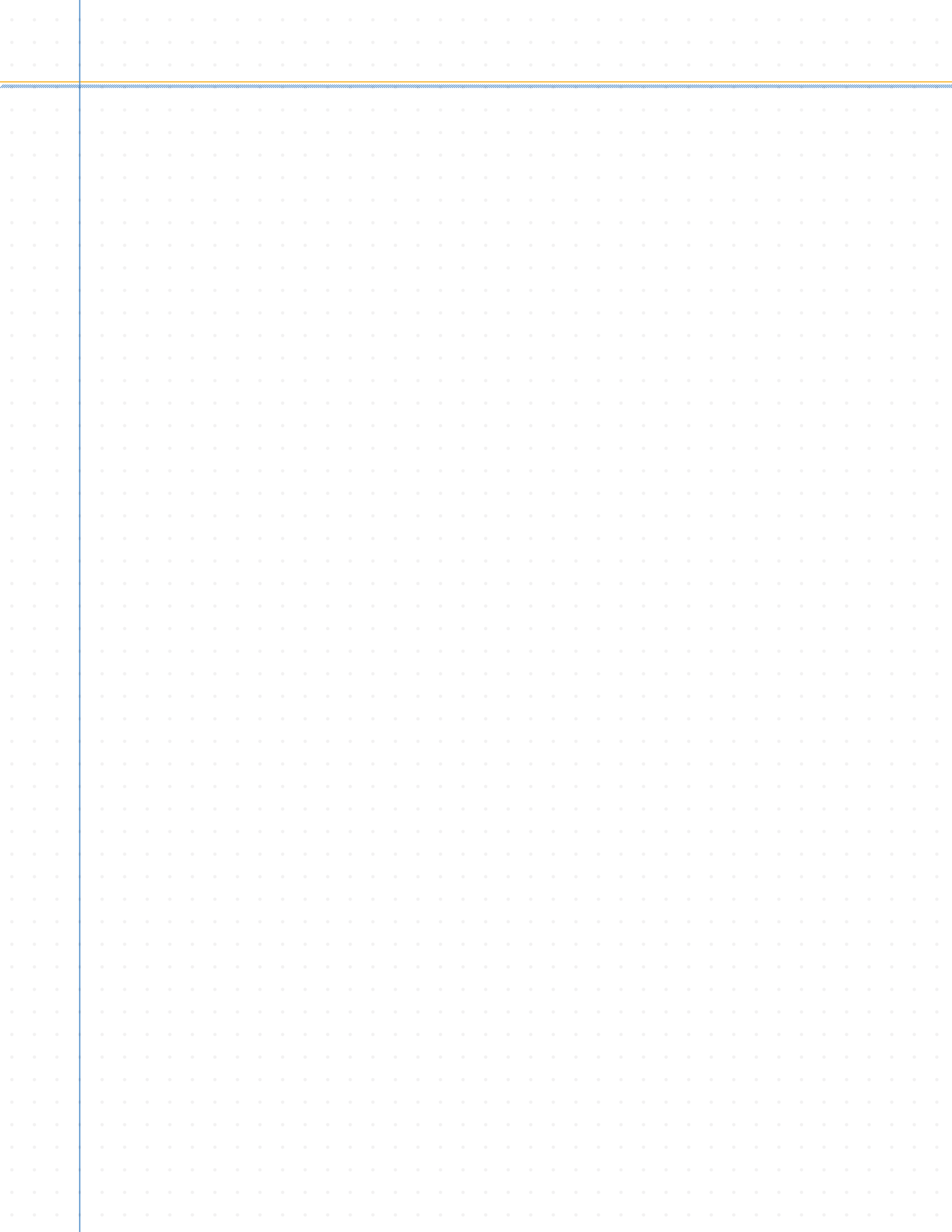
While we can exit specific protocols, the role of protocols itself is an inescapable part

of managing complexity in a modern world. Instead of resisting protocols, we ought to bring a greater awareness to their overall influence, so that we can make better decisions about which protocols we wish to be a part of and how we might live peaceably under their reign. ▲

Thanks to Venkatesh Rao, Rafa Fernandez, Seth Killian, Eric Alston, Scott Moore, Tim Beiko, and Brett Fujioka for their feedback and contributions.

NADIA ASPAROUHOVA is a writer and researcher. She is the author of *Working in Public: The Making and Maintenance of Open Source Software* (Stripe Press) and “Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure” (Ford Foundation). Her work has been supported by Protocol Labs, Emergent Ventures, and Schmidt Futures.

nadia.xyz



Protocol*Kit*

summerofprotocols.com
hello@summerofprotocols.com

ISBN 978-1-962872-11-9 9 0 0 0 0 >



Protocol*Kit*



RETROSPECTUS



NEWSLETTER

