Data & Policy (2023), 5: e1
doi:10.1017/dap.2022.38

CAMBRIDGE UNIVERSITY PRESS

COMMENTARY

The forgotten preconditions for a well-functioning internet

Geoffrey Goodell

Department of Computer Science, University College London, London, United Kingdom
E-mail: G.Goodell@cs.ucl.ac.uk

Received: 30 May 2022; **Revised:** 02 September 2022; **Accepted:** 24 November 2022

Key words: Computational Governance; Internet Architecture; Internet Transport; Decentralised Systems; Network Neutrality

Abstract

For decades, proponents of the Internet have promised that it would one day provide a seamless way for everyone in the world to communicate with each other, without introducing new boundaries, gatekeepers, or power structures. What happened? This article explores the system-level characteristics of a key design feature of the Internet that helped it to achieve widespread adoption, as well as the system-level implications of certain patterns of use that have emerged over the years as a result of that feature. Such patterns include the system-level acceptance of particular authorities, mechanisms that promote and enforce the concentration of power, and network effects that implicitly penalize those who do not comply with decisions taken by privileged actors. We provide examples of these patterns and offer some key observations, toward the development of a general theory of why they emerged despite our best efforts, and we conclude with some suggestions on how we might mitigate the worst outcomes and avoid similar experiences in the future.

Policy Significance Statement

Although the Internet was evidently designed to be robust and decentralized, some of its design characteristics, combined with the human tendency to seek protection and convenience, have given rise to paternalism, exposing control points that enable certain groups to play outsized roles. Such groups have managed to effectively corner the market for four essential aspects of human communication: access, naming, trust, and reputation. Registries for names and addresses, certification authorities, and web infrastructure providers have emerged as de facto gatekeepers, to the extent that using the Internet in practice implies accepting the primacy of these actors. Without a global alternative to the Internet, there is no way to ensure that the Internet serves the interests of its users.

1. Introduction

A preponderance of articles and books have emerged in recent years to decry pernicious abuse of the infrastructure of the Internet. Much of the argument is about *surveillance capitalism.* We know that companies such as Facebook undermine human autonomy by creating indexable records of individual persons and generate tremendous revenue along the way (Zuboff, 2015). We also know that companies such as Microsoft provide Internet services, such as Microsoft Teams, to draw ordinary individuals and businesses into walled gardens that make it possible for the companies to observe their habits and activities (Waters, 2021). At the same time, we know that some other businesses have come to depend upon the data-harvesting ecosystem, and we worry that addressing the harms of surveillance capitalism might necessarily entail collateral damage (Chen, 2022). Economic incentives and entrenched power

dynamics have certainly given rise to the patterns of application service provision via the software-as-a-service revenue model, and network effects have buttressed the concentration of resources in the "cloud" and data centers. Arguments about business motivations and economic imperatives are powerful and moving, and we might hope to mitigate their negative externalities with the right set of changes to law and regulation, applied carefully.

This view is half right. But the design of the Internet has a fundamental property that is less often discussed: *The Internet is designed to expose network-level metadata about its end-users, both to each other and to network carriers.* There are important reasons underpinning this design choice, most notably the fact that Internet users expect the network to carry a packet to its destination without being told explicitly how to get there. The guiding principle has been that the network should "just work" with minimal expectations of knowledge about the structure of the network on the part of end users and their devices. This design has many benefits, including both resilience and dynamism. If part of the network fails, then the network can correct itself without the involvement or concern of end-users. New users, services, and entire networks can join or leave with minimal impact or cost to the set of existing participants. Arguably, this feature was instrumental in facilitating the early adoption of the Internet.

However, there is a dark side to this design choice as well, which might lead us to recognize it as a flaw. As its value in drawing new users waned with the passage of time, its value as a mechanism of control became apparent. Exposing network-level metadata about how end-users are connected to the Internet to carriers and to other end-users allows for the concentration and abuse of power, in the form of discrimination, surveillance, and coercion.

The impact of this design choice may have been felt more acutely in recent years, but it is certainly not new. The underlying protocols and assumptions have been in place for decades, giving rise to control points related to *access*, *naming*, *trust*, and *reputation.* Within each of these categories, the control points have enabled majoritarian decisions about the locus of authority and penalties for non-compliance, which have in turn undermined the potential of the Internet to serve everyone. Although the design flaw is structural, and even technical, it offers insight into a foundational problem that is about humanity as much as it is about technology.

2. Access: One Internet for all?

We might say that the Internet is a global institution. However, this is misleading. The Internet is not really institutional, and we might say that the Internet is successful precisely because it is *not* institutional. There are no particular rules beyond the protocol specifications and no global institutions with the power to enforce the few rules that do exist, which indeed are often broken. Unlike many global systems, including decentralized systems such as most cryptocurrency networks, the Internet does not require global consensus to function, beyond agreement on its foundational protocols. The Internet has never been globally consistent, and achieving consistency would be theoretically impossible. There is no authority that mandates changes to the protocol, and as of 2021, the world is currently in the process of migrating the data plane of the Internet from a version originally developed in 1981 to a version originally developed in 1995. As a "best effort" communication medium, the Internet has no mechanisms to ensure quality of service, or indeed any quality of service at all. Nonetheless, I shall argue that all of these characteristics are features, not bugs, in the design of the Internet, and are essential contributing factors to the success of its design.

Every device connected to the Internet speaks Internet Protocol, either the 1981 version or the 1995 version, and has an address (an *IP address*). The fundamental unit of aggregation in the control plane of the Internet is the *autonomous system*, or AS. Autonomous systems communicate with each other via a protocol called Border Gateway Protocol, or BGP, the first version of which was developed in 1989. The purpose of BGP is to exchange reachability information about autonomous systems (Lougheed and Rechter, 1989). There are over 400,000 autonomous systems, and every IP address that is globally addressable can be recognized as part of a block (or *prefix*) of addresses, which in turn is advertised via BGP by an autonomous system. The operators of autonomous systems determine, as a matter of local policy, which advertisements to accept, which not to accept, and which to prioritize over others

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

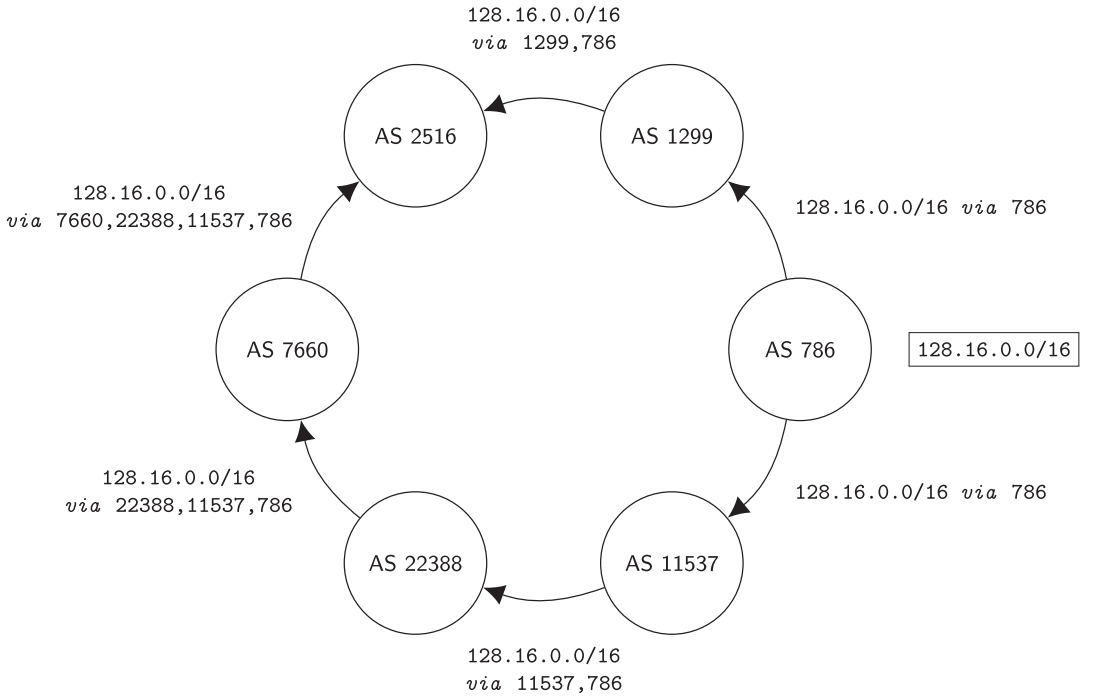

Figure 1. An example of BGP route advertisements. *AS 786 advertises prefix 128.16.0.0/16 to its neighbors, who propagate the advertisement by appending their respective AS numbers to the path. AS 2516 has a policy decision to make: Should it send packets destined for 128.16.0.0/16 to AS 1299 or to AS 7660?*

(see Figure 1). Reconciliation is pairwise: if the operators of an AS receive an advertisement from a neighboring AS that does not look right, or if their customers have a problem reaching certain addresses advertised by a neighboring AS, then they can raise the matter with their peers.

There are five regional information registries that manage the allocation of specific IP address prefixes and autonomous system numbers. However, these registries are not really authorities. Their role is convenience: to prevent collisions among the numbers that different parties are using, not to adjudicate the allocation of some scarce resource. Specific numbers do not confer any particular advantage, and address prefixes, while limited, are generally not the factor that limits the ability of an autonomous system to send and receive traffic. Furthermore, registries have no ability to enforce the usage of the numbers in their registries; ultimately, autonomous systems will route traffic as they please, and allow for the creation of arbitrary patterns of connection. So far, so good.

However, to use the Internet, everyone must implicitly agree upon the authority of this particular set of registrars, since collectively they determine who can have IP addresses and which addresses they can have. We all agree, by majority decision, upon a single globally-meaningful mapping of users to IP addresses. At the same time, there is no mechanism to ensure that all IP addresses are equal, and some are more important than others. In practice, IP addresses are controlled by gatekeepers. Control is generally delegated, creating a hierarchical structure of accountability that can be characterized as a kind of federation. A carrier can assign an arbitrary address to a device connected to a local network, and use *network address translation* to route global traffic to and from local devices without requiring the local devices to have globally meaningful names. As a result, IP addresses have mostly become the concern of Internet carriers and service providers rather than end-users. The ability to differentiate customers based on whether they can be reachable for new connections allows carriers to offer their services at different price points, with businesses that maintain servers subsidizing ordinary individuals with personal computers and mobile devices.

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

Thus, the first manifestation of the structural flaw of the Internet is exposed: On the basis of reachability alone, the Internet has already been divided into first-class and second-class citizens. Unsurprisingly, a primary motivating factor for the threat to end-to-end system design is revenue for carriers through price discrimination, not the scarcity of unique addresses, available bandwidth, or any technical resource in particular (Odlyzko, 2004). This state of affairs might be seen as an unintended consequence of a primary incentive for carriers to provide service, and in fact, many ordinary users with mobile and broadband connections are indeed assigned globally unique addresses that are purposefully blocked by their carriers from receiving new connections.

The incentive to create price discrimination, and its concomitant stratification, exists irrespective of material costs related to infrastructure development. The opportunity is available to network carriers as a direct consequence of the metadata that expose different patterns of use of the network. Put another way, network neutrality is foolhardy: Carriers that do not take advantage of the metadata are leaving revenue on the table and might even cause their investors to question whether they have disregarded their fiduciary responsibilities.

For those customers with the good fortune to have Internet devices that are reachable, the fact that they generally want a way for others to know how to reach them introduces another risk of unintended consequences, which we shall explore next.

3. Naming: Context is Everything

Although the management of AS numbers and IP address prefixes is relatively peaceful, the management of human-meaningful names for Internet hosts is much more contentious. Human-meaningful names are problematic when there is an expectation that everyone respects the same naming convention. Who gets to decide what names we use, and what makes a decision legitimate? A commonly recognized principle called Zooko's triangle holds that the names we use cannot be universally recognized and human-meaningful without being managed by a common authority (Wilcox-O'Hearn, 2001). Similar arguments have been made throughout human history, with memorable parables ranging from the Tower of Babel to Humpty Dumpty. Notwithstanding the validity of these longstanding arguments, the Domain Name System, or DNS (Mockapetris, 1983), represents yet another effort to achieve exactly this kind of global agreement and becomes a second manifestation of the structural flaw of the Internet. Perhaps the authors of the original specification in 1983 had not anticipated the number of Internet devices that would eventually require names and how important those names would become.

DNS is hierarchically managed, with a set of globally agreed *root servers* that delegate authority to a set of *top-level domains*, which in turn delegate authority to registrars and others who facilitate registration of names by individuals, institutions, businesses, and other organizations, usually for a fee. As with the management of IP addresses, there is a kind of federation. However, we all agree, by majority decision, upon the root of the hierarchy, and all decisions about the binding of names to end-users flow from that agreement. The names themselves refer to IP addresses and other metadata that are used to access Internet services. A name is said to be *fully qualified* if it references a specific path of delegation back to the root; an example of a fully qualified name is "www.ntt.com.", which indicates delegation from the root authority to the registrar of "com." to the registered authority for "ntt.com.". But what organization has the privilege to call itself ntt.com? Short, pithy, and prestigious names are scarce, and in practice they demand a hefty premium in the global marketplace for names, if they are traded at all.

To make matters worse, the use of DNS generally involves the active involvement of another gatekeeper, the client-side resolver. The client-side resolver can observe the hostnames requested by the client, potentially exposing the client to profiling and discrimination. Over the years, various approaches such as DNSSEC (Eastlake and Kaufman, 1997) and DNS over TLS (Hu et al., 2016) have been developed to reduce the risk that a client's carrier might manipulate the names that a client receives. However, these approaches are not really private. While they protect the data inside the DNS requests from eavesdropping by client-side carriers, they shift the locus of eavesdropping to global-scale platform

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

service providers instead, such as Google or Cloudflare. Such platform service providers amass much more data than most local carriers ever could, benefiting their own profiling activities and potentially attracting attacks. At the same time, the local network can still determine the sites clients visit, since the clients generally send traffic to the addresses they resolve.

Ostensibly in response to criticism, Cloudflare recently partnered with Apple and Fastly to propose "Oblivious DNS over HTTPS" (ODoH), which adds end-to-end encryption and a proxy service to stand between DNS clients and its DNS resolvers (Singanamalla et al., 2021). However, since the proxy service and, presumably, its network carriers learn both the IP address of the client as well as the exact timing and sequence information of its contact with the resolver, this single-hop proxy does not offer much privacy. In particular, Cloudflare acknowledges relationships with "leading proxy partners, including PCCW Global, SURF, and Equinix" (Cloudflare Research, 2022), and both parties could easily collaborate to de-anonymize a client, if not as part of routine business, then certainly when compelled by an authority to do so. Even if it were expected that clients would mainly communicate with the resolver via independent anonymizing proxy networks, operators of large-scale DNS resolvers would still be positioned to exercise outsize authority over Internet names.

Inexorably, the globally recognized allocation of names to registrants introduces a power dynamic and raises the question of fairness. After all, why are some parties privileged to the exclusion of others? There are various initiatives, such as Namecoin (2021), that seek to reject the authority of DNS entirely and replace it with a transparent marketplace, but such attempts ultimately introduce more problems than they solve. For example, can any system that is first-come, first-served can ever be fair? Should the best name always go to the highest bidder? It would seem that both the paternalistic hand of a trusted authority and the invisible hand of the marketplace fall short of furnishing a solution that works for everyone, precisely because any solution that demands global consensus from a process must also assume global and durable acceptance of the outcome of that process.

A better approach to naming would be to recognize that in human settings, names are not bestowed by authorities except for specific contexts in which the authorities exercise their authority. In general, individual persons and the local communities in which they are embedded decide which names to use to identify ordinary people, groups, objects, concepts, and so on. Perhaps we might encourage users to do what humans have always done, and assign *their own* names for the parties and services with whom they interact. An approach described as "Petnames" (Stiegler, 2005) can facilitate this. A major selling point of hyperlinks is the idea that users can navigate without explicitly concerning themselves with the official universal resource identifiers (URIs) of the resources they seek. Perhaps users should assign their own names using tools akin to browser bookmarks, and the sites themselves could even suggest appropriate names to use. Freeing universally recognizable URIs from the requirement to be human-meaningful allows for the possibility that they might be made secure.

4. Trust: Anchors or Millstones?

Since Internet routing is fundamentally decentralized, how can one be sure that the other party to a conversation, as identified by its IP address or domain name, is authentic? This is a question of security, and we know that we can use cryptography not only to protect our conversations from eavesdropping but also to verify the authenticity of a conversation endpoint using digital certificates (Housley et al., 1999). This is straightforward if we can directly and securely exchange public keys with everyone that we might want to engage in a conversation via the Internet. In practice, however, people use the Internet to engage with many different kinds of actors and seldom exchange certificates directly. When was the last time your bank advised you to verify the fingerprint of its public key? Thus, we have identified a third structural flaw of the Internet: By providing for the establishment of regional or even global authorities, its design has bypassed local institutions and authorities as mechanisms for creating trust, and a popular shortcut, which we describe in this section, has undermined those mechanisms and institutions. Although decentralized governance is recognized as promoting "greater accountability and better services," it also tends to deliver less "technical and governance capacity" (Smoke, 2015). Individuals and local communities are in no

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

position to negotiate with global authorities, but has the Internet raised the bar for technical and governance capacity beyond what local authorities can manage?

Most modern operating systems and web browsers come pre-installed with a set of so-called *trust anchors* that serve as trusted third parties to verify the public keys associated with particular IP addresses or domain names. In principle, users could add or remove trust anchors from this list according to their personal preferences, but almost nobody actually does. Moreover, since Internet services must present their certificates to web browsers and other client software, the question facing the operators of those services is: Which trust anchors are the web browsers and other applications running on end-user devices likely to accept? The operators then seek signatures from the trust anchors that are commonly shipped with end-user software and present those signatures in the form of certificates. Since obtaining, storing, and presenting certificates from certificate authorities carry operational costs (and sometimes economic costs, although the emergence of free alternatives Let's Encrypt, 2021 has changed this), the operators are strongly motivated to be parsimonious. Thus, we have an implicit agreement between site operators and software distributors about the correct set of trust anchors to use, and as a result, those trust anchors become powerful gatekeepers. Just as with access and naming, and consistently with the pattern of majoritarianism, end-users have been forced to accept what has already been agreed.

What happens when a trust anchor fails? Routine breaches of privileged information such as private keys take place from time to time and do not constitute a theoretical question. For example, consider the well-publicized compromise of Comodo in March 2011 (Mills, 2011). Because certificate authorities are trusted by default by widespread client software, the stakes of a breach are high. The compromise of private keys of certificate authorities enabled the Stuxnet worm, which was discovered in 2010 and had been used in attacks against an Iranian nuclear facility (Anderson, 2012), and a subsequent, successful attack on DigiNotar, another certificate authority, allowed the interception of Internet services provided by Google (Adkins, 2011). If vesting so much power in a small number of globally trusted institutions and businesses seems like a dangerous idea, that's because it is.

One initiative, Certificate Transparency (Laurie et al., 2021), aims to address the security problem by creating audit logs to publicize known certificates issued by certification authorities. However, because its primary effect is to shift the locus of control from the certification authorities themselves to the maintainers of the audit logs, it is not really a satisfying answer to the question: Quis custodiet ipsos custodes*?* By implicitly anticipating that all certificates would be logged publicly, it also raises the question of whether it should be possible for legitimate trust relationships to be established and maintained outside the public view. If the Tower of Babel teaches us that the public cannot agree on the authority of one set of trust anchors, then why would the public believe in the ultimate authority of one public logging facility and its ecosystem of maintainers and users?

5. Reputation: Credit or Extortion?

DNS is a formal system with formal rules and institutionally trusted authorities, and the secure verification of Internet services also carries the weight of trusted institutions. However, just as not all power is institutional, not all structural shortcomings of the Internet can be characterized as concerns about the weaknesses and illegitimacy of trusted authorities. A mafia family might be expected to serve the interests of a local community by providing protection, and whether such protection is legitimate might be a matter of perspective, perhaps even ethically debatable. Arguably, such behavior is endemic to the Internet, particularly given its common role as a venue for anti-social and threatening behavior. The mechanisms that enable such behavior to emerge and flourish represent a fourth structural flaw of the Internet.

Scale is a powerful tool for creating change in the world. However, only a handful of actors can wield it, and those who can usually have an incentive not to change the environment that enabled them to thrive. The counterexamples, however, illustrate both the opportunities and the danger.

For context, consider how a powerful organization, Google, used its scale to change the Internet. In 2012, a team at Google sought to optimize the performance of communication between its servers, which were increasingly carrying multimedia content, and personal computers running its Chrome web browser;

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

the team designed and implemented a new transport protocol called QUIC (Roskind, 2012). Because Google controlled both a significant volume of web servers (through its online services) and a significant share of web browsers (through Chrome), this protocol quickly reached broad adoption and experienced a relatively swift journey through the standardization process (Hamilton et al., 2016; Iyengar and Thomson, 2021). The protocol is widely used today. Adoption of the protocol was generally voluntary, and so this use of power can be considered benign.

When scale is combined with the problem of assessing creditworthiness or reputation, the results are somewhat less benign. Consider the following case that again involves Google and relates to e-mail spam. No one really likes to receive unsolicited messages from dodgy actors, which is a system-level consequence of the common practice of using the same e-mail address across different contexts, a topic for a separate article. A technical approach to mitigating spam is to require senders to prove that they are the rightful owners of the domain names they use to represent themselves; this approach forms the essence of the Sender Policy Framework, or SPF (Wong and Schlitt, 2006), which was first proposed in 2000 (DMARCian, 2019), as well as DomainKeys Identified Mail, or DKIM (Allman et al., 2007), which was introduced in 2007. Notice that this approach further entrenches the authority of DNS system operators. One might imagine that it would be impossible to compel all of the mail servers on the Internet to stop sending mail without valid SPF or DKIM signatures, and indeed both the specification for SPF and the specification for DKIM advise mail servers not to reject messages "solely" based on an inability to authenticate the sender (Wong and Schlitt, 2006; Allman et al., 2007). However, one mail server operator, Google, implemented a policy that stretched the limits of this recommendation by routinely classifying mail sent to users of its popular Gmail service that did not include a valid SPF or DKIM header as spam (Google, 2017). As a result, many mail servers were forced to implement SPF or DKIM because their users wanted to be able to send mail to Gmail users. And so, by leveraging its scale, Google had successfully managed to twist the arms of mail server operators around the world.

A decidedly less benign example involves the creation of blacklists of IP addresses based on their reputation, ostensibly for the purpose of mitigating attacks on servers. The idea of using routing information, such as IP address, as a convenient way to judge the reputation of a sender is not new (Goodell and Syverson, 2007), and the ability to wield disproportionate power is not always limited to large, wealthy corporations. Consider the case of SORBS, an Australian organization that aims to reduce the preponderance of spam through the publication of a list of IP addresses associated with spam or abuse. Although this site does not directly engage in filtering traffic, many e-mail server operators configure their servers to consult this list as part of their routine operation and flag messages as spam or reject them outright if they are received from a mail server whose IP address appears on the SORBS list. SORBS generally requires the operators of servers with blacklisted IP addresses to explicitly request expungement, subject to the discretion of SORBS staff (SORBS, 2021), and for years, such operators were required to give money to charity as well (SORBS, 2021). Similarly, services such as Spamhaus maintain lists of IP addresses that are apparently "directly involved in malicious behaviour" or have "a bad reputation associated with them" (Spamhaus Project, 2021). Carriers are incentivized to maintain discipline among their subscribers, perhaps even to threaten them with termination, if such services list their IP addresses.

A related example involves the practice of using similar blacklists to restrict access to websites. For example, the popular web infrastructure provider Cloudflare (2021) offers its customers an option to block IP addresses associated with public VPNs and anonymizing proxy networks such as Tor (Dingledine et al., 2004). In 2015, Akamai Technologies released a report claiming that in an experiment conducted by its researchers, "Tor exit nodes were much more likely to contain malicious requests" and recommended that traffic from Tor should be either heavily scrutinized… or completely blocked" using tools such as its proprietary security module "Akamai Kona Site Defender" (Akamai Technologies, 2015). Kona Site Defender includes a service called "Client Reputation" that categorizes IP addresses as being malicious or not on the basis of their interaction with Akamai services around the world (Akamai Cloud Security Solutions, 2018; Akamai Technologies, 2021), which its customers can use to block access to their websites.

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

A direct system-level consequence of this pattern of blacklisting is that Internet users who share IP addresses with each other, and especially Internet users who rely upon anonymity networks for their privacy, are denied access to much of the Internet. A second-order system-level consequence is that Internet users are forced to relinquish their privacy and must instead consider maintaining favorable reputations in the eyes of powerful gatekeepers, as well as strong relationships with their Internet carriers, as a prerequisite to using the Internet.

In all such cases, extortionate behavior is enabled by the separation of the function of maintaining the blacklist from the choice to use the blacklist, the perceived benefit of the blacklist on the part of its users, and the lack of accountability with respect to policy on the part of server operators that use the blacklist to the users of their own services. Are these the stewards that we wanted? Certainly they are the stewards that we have.

6. Negative Externalities

All of our distributed systems, however good they are, are susceptible to abuse of the power that has been vested in the systems themselves. In particular, systems based on the Internet reflect, at minimum, the unfairness and power dynamics that are intrinsic to the Internet itself.

This is a tragedy of the commons. The same control points that are called upon to attract users by offering convenience are inevitably exploited to control them. The features that allow key users and service providers to privately benefit introduce significant public costs. It is easy to see how those with an inclination to ignore the effects of externalities in mechanism design might be attracted to the features of the Internet, wherein there are ample opportunities to profit by imposing diffuse, often imperceptible costs upon a broad set of users.

An important lesson of the Internet is that users mostly do not understand what can be taken away from them until after it is gone. As metadata leakage furnishes control to powerful gatekeepers, the rest of the world shrugs, having no recourse. Having achieved tremendous power and influence, platforms defy efforts to resist their locus of control. It might seem that individuals have a choice to ignore the power of platforms, but they do not: The value of a platform is defined by its network effects (Baron, 2022). This is true not only for popular applications and social media platforms, but also for the implicit structures under the surface that define the rules for routing, naming, public access, and public trust.

7. Conclusion: Protecting our Infrastructure from Ourselves

It is worth considering the original design of the Internet as a reminder that it is possible to build a system that avoids vesting too much power in in its operators. This concept is generally described as the *end-to-end principle* and achieved widespread appreciation well before the protocols described earlier in this article were developed (Saltzer et al., 1984). We find that the Internet is not majoritarian per se, but weaknesses in its design have allowed certain actors to develop and exercise power over others. But, how do we undo the dangerous power structures that have become part of the de facto architecture of the Internet? Moreover, how do we ensure that the systems we design in the future do not become centralized in the first place?

We note that global consensus and, more generally, the interest in establishing an authoritative global view of something is exactly where we see the greatest risk of centralization in practice. Part of the interest in an authoritative global view derives from the convenience of not having to worry about whether we share context with our neighbors, although it is perhaps when there is a perception of risk that the tendency to embrace an authoritative global view is most pronounced. Under such circumstances, we collectively retreat to the safety of big, dominant actors that can exercise their control and, in so doing, help us avoid the responsibility of decision-making and the potential blame that might follow (Jackall, 2009). Our observation should be a warning sign for those who seek to build systems that rely upon global consensus. Do systems that require global consensus grind to a halt when anyone disagrees, or do they implicitly

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

ignore the wishes of the minority? Put another way, who or what is the arbiter, and what is to prevent that arbiter from being compromised by those who disagree from us?

With respect to the Internet, there is no easy solution, although we can imagine its characteristics. Identity must be non-hierarchical and separate from routing, so that reputations are not earned and lost based upon how users are connected to the network. Internet routers must not be able to infer who is speaking to whom, or what kind of information is being communicated, from the traffic they carry. This would seem to imply a requirement for a second, *oblivious* layer of routing between the existing network layer that relies upon convenient connectivity and the end-to-end protocols that users see. Perhaps this could be achieved by encouraging and expecting Internet users to use onion routing (Syverson et al., 1997; Dingledine et al., 2004) as a way to shield end-to-end conversations from the pastoral gaze of network operators, and to use self-certifying names, such as those used by Tor onion services, to prevent the accumulation and exercise of illegitimate authority by third-party actors. Self-certifying names should be opaque, to reduce the likelihood that people would confuse them with meaningful names or concepts. It is not far-fetched to imagine inserting another layer to the Internet, between the layer at which routing operates and the layer that manages end-to-end conversations, to mitigate the harm associated with authorities occupying positions of control within the network.

Finally, it is in the interest of Internet users to avoid being pinned to a single identity, either over time or across their many relationships, and they should have tools that help them assume different identities for their different relationships, to mitigate the risk of both spam and blacklisting at the same time.

Above all, we should eliminate the assumption that metadata exposed to the system will not be misused, as well as the assumption of paternalistic goodwill on the part of powerful actors and service providers. It would behoove us to consider the fundamental trade-off intrinsic to all global systems and pursue a principle of parsimony: How can we provide the system with enough power to do what we need it to do, without providing it with enough power to profit at our expense? We must take a clear-eyed view to decisions that were made long ago in the interest of marshalling attention and serving the narrow interests of actors with a role in developing key infrastructure, but which have since outlasted their usefulness. It is time to rethink the relationship between the network and metadata: The Internet is not fully mature until the edges can communicate without the permission of the center.

Acknowledgments. We express gratitude for the feedback of anonymous reviewers and attendees of the CoGMa Workshop, as well as the continued support of Professor Tomaso Aste, the Centre for Blockchain Technologies at University College London, and the Systemic Risk Centre at the London School of Economics. A pre-print of this article is also available online (Goodell, n.d.).

Funding Statement. The author acknowledges the support of TODAQ Financial and the Peer Social Foundation.

Author Contributions. Writing—review & editing: G.G.

Data Availability Statement. Data availability is not applicable to this article as no new data were created or analyzed in this study.

Competing Interest. The author declares no competing interests exist.

References

Adkins H (2011) "An Update on Attempted Man-in-the-Middle Attacks." Security Blog, Google, 2011-08-29 [online]. Available at https://security.googleblog.com/2011/08/update-on-attempted-man-in-middle.html (accessed 22 September 2021).

Akamai Cloud Security Solutions (2018) "Kona Site Defender: Protect Your Websites and Web Applications from Downtime and Data theft." Product Brief, April 2018 [online] Available at https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RWBnqk (accessed 29 August 2022).

Akamai Technologies (2015) "Akamai's State of the Internet/Security: Q2 2015 Report." August 2015 [online]. Available at https://web.archive.org/web/20170317110115/; https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/2015-q2-cloud-security-report.pdf (accessed 29 September 2021).

Akamai Technologies (2021) "Client Reputation" [online] Available at https://learn.akamai.com/en-us/products/cloud_security/client_reputation.html (accessed 22 September 2021).

Allman E, **Callas J**, **Delany M**, **Libbey M**, **Fenton J and Thomas M** (2007) "DomainKeys Identified Mail (DKIM) Signatures." Internet Engineering Task Force RFC 4871, May 2007 [online]. Available at https://datatracker.ietf.org/doc/html/rfc4871 (accessed 22 September 2021).

Anderson N (2012) "Confirmed: US and Israel Created Stuxnet, Lost Control of It." Ars Technica, 2012-06-01 [online]. Available at https://arstechnica.com/tech-policy/2012/06/confirmed-us-israel-created-stuxnet-lost-control-of-it/ (accessed 22 September 2021).

Baron M (2022) "Elon Musk is Right. Web3 is BS." Medium, 2022-01-09 [online]. Available at https://maciekbaron.medium.com/elon-musk-is-right-web3-is-bs-1cdafc3f96f7 (accessed 21 February 2022).

Chen B (2022) "The Battle for Digital Privacy is Reshaping the Internet." *The New York Times*, 2021-09-16 [online]. Available at https://www.nytimes.com/2021/09/16/technology/digital-privacy.html (accessed 22 September 2021).

Cloudflare (2021) "Understanding Cloudflare Tor Support and Onion Routing." 2021-02-16 [online]. Available at https://support.cloudflare.com/hc/en-us/articles/203306930-Understanding-Cloudflare-Tor-support-and-Onion-Routing (accessed 22 September 2021).

Cloudflare Research (2022) "Oblivious DNS over HTTPS" [online]. Available at https://research.cloudflare.com/projects/network-privacy/odns/ (accessed 29 August 2022).

Dingledine R, **Mathewson N and Syverson P** (2004) "Tor: The second-generation onion router." In *Proceedings of the 13th USENIX Security Symposium*. San Diego, California: USENIX, pp. 303–320 [online]. Available at https://svn.torproject.org/svn/projects/design-paper/tor-design.pdf (accessed 22 September 2021).

DMARCian (2019) "The History of SPF" [online]. Available at https://dmarcian.com/history-of-spf/ (accessed 18 March 2019).

Eastlake D and Kaufman C (1997) "Domain Name System Security Extensions." Internet Engineering Task Force RFC 2065, January 1997 [online]. Available at https://datatracker.ietf.org/doc/html/rfc2065 (accessed 21 February 2022).

Goodell G (2022). "The Forgotten Pre-Conditions for a Well-Functioning Internet" [online]. Available at https://arxiv.org/abs/2109.12955; https://ssrn.com/abstract=4205600.

Goodell G and Syverson P (2007) The right place at the right time. *Communications of the ACM 50*(5), pp. 113–117. https://doi.org/10.1145/1230819.1241689

Google (2017) "Bulk Senders Guidelines." Gmail Help, Archived 2017-06-16 [online]. Available at https://web.archive.org/web/20170616095129/https://support.google.com/mail/answer/81126 (accessed 22 September 2021).

Hamilton R, **Iyengar J**, **Swett I and Wilk A** (2016) "QUIC: A UDP-Based Secure and Reliable Transport for HTTP/2." Internet Engineering Task Force Internet-Draft draft-tsvwg-quic-protocol-02, 2016-01-13 [online]. Available at https://datatracker.ietf.org/doc/draft-tsvwg-quic-protocol/ (accessed 29 August 2022).

Housley R, **Ford W**, **Polk W and Solo D** (1999) "Internet X.509 Public Key Infrastructure Certificate and CRL Profile" Internet Engineering Task Force RFC 2459, January 1999 [online]. Available at https://datatracker.ietf.org/doc/html/rfc2459 (accessed 22 September 2021).

Hu Z, **Zhu L**, **Heidemann J**, **Mankin A**, **Wessels D and Hoffman P** (2016) "Specification for DNS over Transport Layer Security (TLS)." Internet Engineering Task Force RFC 7858, May 2016 [online]. Available at https://datatracker.ietf.org/doc/html/rfc7858 (accessed 21 February 2022).

Iyengar J and Thomson M (2021) QUIC: A UDP-Based Multiplexed and Secure Transport." Internet Engineering Task Force RFC 9000, May 2021 [online]. Available at https://datatracker.ietf.org/doc/html/rfc9000 (accessed 29 August 2022).

Jackall R (2009) *Moral Mazes: The World of Corporate Managers*. Oxford: Oxford University Press. ISBN: 978–0199729883.

Laurie B, **Messeri E and Stradling R** (2021) "Certificate Transparency Version 2.0." Internet Engineering Task Force RFC 9162, December 2021 [online]. Available at https://datatracker.ietf.org/doc/html/rfc9162 (accessed 26 August 2022).

Let's Encrypt (2021) [online]. Available at https://letsencrypt.org/ (accessed 22 September 2021).

Lougheed K and Rechter Y 1989 "A Border Gateway Protocol (BGP)." Internet Engineering Task Force RFC 1105, June 1989 [online]. Available at https://datatracker.ietf.org/doc/html/rfc1105 (accessed 22 September 2021).

Mills E (2011) "Comodo: Web Attack Broader than Initially Thought." CNet, 2011-03-30 [online]. Available at https://www.cnet.com/tech/services-and-software/comodo-web-attack-broader-than-initially-thought/ (accessed 22 September 2021).

Mockapetris P (1983) "Domain Names – Implementation and Specification." Internet Engineering Task Force RFC 883, November 1983 [online]. Available at https://datatracker.ietf.org/doc/html/rfc883 (accessed 22 September 2021).

Namecoin (2021). [online]. Available at https://www.namecoin.org/ (accessed 22 September 2021).

Odlyzko A (2004) Pricing and architecture of the internet: Historical perspectives from telecommunications and transportation. In *Proceedings of TPRC 2004* [online]. Available at http://www.dtc.umn.edu/~odlyzko/doc/pricing.architecture.pdf (accessed 29 August 2022).

Roskind J (2012) "QUIC: Design Document and Specification Rationale." April 2012 [online]. Available at https://docs.google.com/document/d/1RNHkx_VvKWyWg6Lr8SZ-saqsQx7rFV-ev2jRFUoVD34/edit (accessed 29 August 2022).

Saltzer J, **Reed D and Clark D** (1984) End-to-end arguments in system design. *ACM Transactions in Computer Systems 2*(4), 277–288. Available at https://web.mit.edu/Saltzer/www/publications/endtoend/endtoend.pdf (accessed 20 February 1999).

Singanamalla S, **Chunhapanya S**, **Hoyland J**, **Vavruša M**, **Verma T**, **Wu P**, **Fayed M**, **Heimerl K**, **Sullivan N and Wood C** (2021) Oblivious DNS over HTTPS (ODoH): A practical privacy enhancement to DNS. *Proceedings on Privacy Enhancing Technologies 4*, 575–592. Available at https://files.research.cloudflare.com/publication/Singanamalla2021.pdf (accessed 29 August 2022).

Smoke P (2015) Accountability and service delivery in decentralising environments: Understanding context and strategically advancing reform. In *A Governance Practitioner's Notebook: Alternative Ideas and Approaches*. Paris: Organisation for Economic

https://doi.org/10.1017/dap.2022.38 Published online by Cambridge University Press

Co-operation and Development, pp. 219–232 [online]. Available at https://web.archive.org/web/20220119021649/; https://www.oecd.org/dac/accountable-effective-institutions/Governance%20Notebook%202.6%20Smoke.pdf (accessed 19 January 2022).

SORBS (2021) "About SORBS" [online]. Available at http://www.sorbs.net/general/about-SORBS.shtml (accessed 22 September 2021).

SORBS (2021) "SpamDB FAQ." Archived 2010-07-25 [online]. Available at https://web.archive.org/web/20100725031143/; http://www.dnsbl.au.sorbs.net/faq/spamdb.shtml (accessed 22 September 2021).

Spamhaus Project (2021) "What is a Blocklist?" [online]. Available at https://web.archive.org/web/20210925072715/; https://check.spamhaus.org/faqs/?id=what_is_a_blocklist (accessed 25 September 2021).

Stiegler M (2005) "An Introduction to Petname Systems." February 2005 [online]. Available at http://www.skyhunter.com/marcs/petnames/IntroPetNames.html (accessed 11 May 2018).

Syverson P, Goldschlag D and Reed M (1997) Anonymous connections and onion routing. In *Proceedings of the 18th Annual Symposium on Security and Privacy*. Oakland, CA: IEEE, CS Press, pp. 44–54 [online]. Available at http://www.onion-router.net/Publications/JSAC-1998.pdf (accessed 23 August 2021).

Waters R (2021) "Microsoft Looks to Make 2021 the Year of Teams." *Financial Times*, 2021-01-05 [online]. Available at https://www.ft.com/content/1bbe1b15-dde6-4a3b-9728-8991818b6c92 (accessed 20 September 2021).

Wilcox-O'Hearn Z (2001) "Names: Decentralized, Secure, Human-Meaningful: Choose Two" [online]. Available at https://web.archive.org/web/20011020191610/; http://zooko.com/distnames.html (accessed 24 April 2018).

Wong M and Schlitt W (2006) "Sender Policy Framework (SPF) for Authorizing Use of Domains in E-Mail, Version 1." Internet Engineering Task Force RFC 4408, April 2006 [online]. Available at https://datatracker.ietf.org/doc/html/rfc4408 (accessed 22 September 2021).

Zuboff S (2015) Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology 30*, 75–89 [online]. Available at http://www.palgrave-journals.com/jit/journal/v30/n1/pdf/jit20155a.pdf (accessed 17 September 2018).

Cite this article: Goodell G (2023). The forgotten preconditions for a well-functioning internet. *Data & Policy*, 5: e1. doi:10.1017/dap.2022.38