Ethics and Information Technology (2024) 26:70
https://doi.org/10.1007/s10676-024-09810-5

ORIGINAL PAPER

Digital sovereignty and artificial intelligence: a normative approach

Huw Roberts[1]

Abstract
Digital sovereignty is a term increasingly used by academics and policymakers to describe efforts by states, private companies, and citizen groups to assert *control* over digital technologies. This descriptive conception of digital sovereignty is normatively deficient as it centres discussion on how power is being asserted rather than evaluating whether actions are legitimate. In this article, I argue that digital sovereignty should be understood as a normative concept that centres on *authority* (i.e., legitimate control). A normative approach to digital sovereignty is beneficial as it supports critical discourse about the desirability of actors' assertions of control. It is also more closely aligned with traditional definitions of sovereignty that are grounded in ideas of sovereign authority. To operationalise this normative approach to digital sovereignty and demonstrate the deficiencies of a descriptive approach, the role that "Big Tech" companies are playing in controlling artificial intelligence is considered from both perspectives. Through this case study, it is highlighted that Big Tech companies assert a high degree of control (i.e., descriptive digital sovereignty), but that they lack strong input legitimacy and have a questionable amount of output legitimacy. For this reason, it is argued that Big Tech companies should only be considered quasi-sovereigns over AI.

Keywords Digital sovereignty · Ethics · Legitimacy · Artificial intelligence · Big Tech

Introduction

Following the outbreak of the Covid-19 pandemic in early 2020, governments around the world looked to digital contact tracing to curb the spread of the virus. Early efforts to develop mobile applications capable of tracking infections and informing individuals of close contacts faced difficulties. Chief among them was platform fragmentation between mobile devices running on Google's Android and Apple's IOS which prevented effective interoperability, and platform policies which stopped Bluetooth signals from being sent when a mobile application was closed (Tretter, 2022).

In May 2020, Apple and Google released an API that enabled effective interoperability between contact tracing applications from public health authorities on IOS and Android devices, before subsequently enabling Bluetooth contact tracing in their underlying platforms. However, Apple/Google designed their contract tracing API in a decentralised and privacy preserving manner, meaning public health authorities would not be able to build centralised databases of infection data. After facing difficulties developing effective contact tracing applications, the United Kingdom, Germany, and several other jurisdictions agreed to adopt Apple/Google's approach (Sharon, 2020). Digital sovereignty was sacrificed – or rather shared – with the private sector because the benefits of working within terms set by these companies were perceived to outweigh the costs. In France, restrictions on data collection and use were perceived to be too constraining for the country's health policy choices, leading them to reject Apple/Google's model in favour of an autonomous application (Pizzul & Veneziano, 2023).

The case of contact tracing applications is just one example of the "fights for digital sovereignty" that are playing out across jurisdictions and sectors (Floridi, 2020). While inconsistency remains in how academics and policymakers use the term "digital sovereignty", including to whom it should be applied (Couture & Toupin, 2019; Hummel et al., 2021; Roberts et al., 2021), most definitions are grounded in some idea of *control* over digital technologies (Falkner et

✉ Huw Roberts
huw.roberts@oii.ox.ac.uk

[1] Oxford Internet Institute, University of Oxford, 1 St Giles', Oxford OX1 3JS, UK

Published online: 18 October 2024

al., 2022).[1] Understood in these terms, governments, private sector actors, and citizen groups are variously cooperating and competing to assert control over technologies. Some of these partnerships are relatively transient, as turned out to be the case for contact tracing applications, while others are longer-lasting, like the US's reliance on "Big Tech" companies for cloud services that underpin the functioning of the state.

In this article, I argue that conceptualising digital sovereignty solely in terms of control is normatively deficient. Put simply, this approach provides rich descriptive insights into the way actors are cooperating and competing for control over digital technologies, but it offers little for evaluating these actions against normative standards. In the case of contact tracing outlined above, descriptive approaches to digital sovereignty document the way in which states and Big Tech companies asserted control, but are agnostic with respect to the legitimacy of actors' actions. This is detrimental for ensuring that digital technologies are controlled in a way that is societally beneficial. It is also misaligned with traditional definitions of sovereignty which centre on the *authority* (legitimate control) of sovereign actors (Philpott, 2016).

In line traditional conceptions of sovereignty, I propose that digital sovereignty should be understood as a normative concept centred on the idea of *authority*. This normative approach foregrounds ethical evaluations of the legitimacy of actors' actions in "fights" for digital sovereignty, with public consent a prerequisite for being considered a sovereign actor. To make this argument, the remainder of the paper is structured as follows. Section 2 offers a brief mapping of debates conceptualising digital sovereignty. Section 3 outlines the normative approach to digital sovereignty proposed in this article. Section 4 operationalises this normative approach by assessing the role that Big Tech companies are playing in controlling artificial intelligence (AI), with a particular focus on whether they possess sufficient legitimacy to be considered sovereign actors.

Debates over digital sovereignty

Since the early 1990s, scholars and policymakers have evoked some variant of digital sovereignty, including *cyber sovereignty*, *internet sovereignty*, and *information sovereignty*, to conceptualise a notion of control over (certain) digital technologies (Falkner et al., 2022; Hummel et al., 2021). However, beyond this, there has been disagreement as to how digital sovereignty and related terms should be understood and delineated, and to whom the concepts should be applied (Couture & Toupin, 2019; Baezner & Robin, 2018). Here, digital sovereignty is understood as an umbrella term encompassing the various sub-variants mentioned above. Uses of the term digital sovereignty can generally be split into two overarching categories: *deep but narrow* and *shallow but broad.*

Deep but narrow conceptions are continuous with traditional international relations understandings of the term "sovereignty". At its core, this entails grounding digital sovereignty in the idea of a state's supreme authority within and over a territory, as well as its resources and citizens (Krasner, 2007; Morgenthau, 1948; Philpott, 2016). Conceptions of digital sovereignty that fall into this first category are "deep", in the sense of being "deeply rooted" in traditional literature on sovereignty, but narrow because the nation state is perceived as the only "supreme" actor that can hold sovereignty.

Deep but narrow conceptions of digital sovereignty emerged in response to the new challenges and complexities introduced by digital technologies and the internet, with respect to a state's supreme authority over its territory (Mueller, 2019). By design, the internet facilitates the free flow of information and compresses time-space, making connectivity ubiquitous and communication instantaneous in a manner that does not respect traditional state borders (though the internet of course relies on a considerable amount of physical infrastructure on land and undersea) (Ganz et al., 2024). This connectivity challenges the authority of the state in several ways: national security is threatened by malicious foreign actors who can conduct cyber-attacks on a state without entering its territory (Meyer, 2020; Taddeo, 2017); the free flow of data between borders limits a state's ability to control the information reaching its citizenry, leaving them subject to foreign influences and thus undermining control (Deibert et al., 2008); and state capacity to regulate economic activity is undermined on account of the difficulty of controlling companies located in different territories (Perritt, 1998).

Some deep but narrow scholarship focuses on applying the concept of sovereignty as it exists in international law to the digital. For instance, both the NATO-funded Tallinn Manual (2.0.) and the *United Nations Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace* concluded that some notion of sovereignty should be applied to cyberspace. However, neither definitively concluded what this entails and what would constitute a violation of sovereignty (Schmitt, 2017; Henriksen, 2019). On account of these difficulties, and arguably the inadequacy of international law for dealing with challenges to sovereignty that the internet poses (Franzese, 2009), a more substantive literature has emerged focusing on how states

[1] While some scholars use the term "authority" in respect to digital sovereignty, they often do so without maintaining its normative connotations.

can unilaterally assert digital sovereignty. Some scholars argue that states should build virtual borders to defend themselves from cyber threats (Demchak, 2016; Demchak & Dombrowski, 2014). Others consider how information flows can be controlled or data localised to protect citizens' fundamental rights, to promote economic growth, or to censor content for authoritarian purposes (Basu et al., 2019; Fraser, 2016; Sargsyan, 2016; Savelyev, 2016). Many of these provisions are already being used by states globally, such as China's use of the so-called "Great Firewall" which is designed to restrict citizens' access to sensitive information, and the EU's data adequacy agreements which seek to ensure personal data is protected in foreign jurisdictions.

Deep but narrow scholarship captures the threats that digital technologies pose to state control, but it often fails to acknowledge the limitations of traditional conceptions of sovereignty in the digital age. There is acute difficulty conceptualising digital sovereignty in state-centric terms when the threats to national security are increasingly coming from non-state actors (Timmers, 2019). And the threat to state sovereignty from non-state actors in the realm of cyberconflict only represents the tip of the iceberg.

International relations scholarship has traditionally stressed that the primary purpose of the state is in the realm of defence and that competition between states was chiefly over territory. However, technological developments and market forces have shifted the rationale of the state. Control over territory is no longer the most effective way to accumulate wealth, with affluence now dependent on market share (Strange, 1996). When read in line with the development of nuclear weapons, which drastically increase the stakes of war, the incentives for interstate war become less appealing. As a consequence, competition between states has predominantly shifted from kinetic warfare towards achieving market share within the global economy (Strange, 1995). Contemporary power is less about controlling territory and instead about command over the "nature, location and manner of production and distribution of goods" (Strange, 1996, p. 45). States are by no means displaced wholesale by corporations as sovereign actors in this regard, but governments have a shrinking capacity to provide goods for their own populaces. Numerous public functions are now either outsourced or fully privatised, including prison administration, elements of policing, migration, and asylum (Acharya, 2013).

In recent years, the dependency of governments on the private sector to provide public services can be especially seen in the growing role that large technology firms are playing. Through providing data-driven rather than subject-specific expertise, a small handful of companies are increasingly relied upon across many "spheres" of society (Sharon, 2020), including in the enforcement of statutory laws (Lazar, 2024). This produces a "profound rebalancing of power … privileging corporations with large-scale data power", thus creating a cycle of dependency on them (Magalhães & Couldry, 2021, p.354). For conceptualising sovereignty in a way that is useful for the 21st century, it is essential to recognise that the rationale and capacities of the state itself have shifted.

Broad but shallow conceptions of digital sovereignty recognise the limitations of applying the term solely to the state and as such, seek to encompass a wider array of stakeholders. This leaves the concepts *broad* but also *shallow*, insofar as they are weakly tied to traditional ideas of "sovereignty". This is not to say that broad but shallow conceptions fail to engage with traditional literature on sovereignty; rather, this scholarship is shallow in the sense that it breaks from deeply established norms that are present in the state-centric view of sovereignty. This leaves fewer points of consensus among broad but shallow scholars, such as which actors should be regarded as sovereigns.

Private companies are most commonly presented as a new type of sovereign actor, as despite not holding anything close to *de jure* sovereignty over any territory, they nonetheless increasingly perform *de facto* several of the roles usually reserved for the state (Lehdonvirta, 2022). Some scholars also argue that citizens (Hummel et al., 2018), indigenous communities (Kukutai & Taylor, 2016), civil society groups (Haché, 2014), and cyberspace itself (Barlow, 1996) can or should be considered sovereign. However, the limited influence of these groups in controlling digital technologies means such arguments have gained less traction.

Applying the concept of digital sovereignty beyond the state is helpful and more reflective of the reality of contemporary sovereignty as something that is not just held by the state (Connolly, 2004). In line with this, a growing literature acknowledges the relational nature of digital sovereignty as something that is shared between various stakeholders who are asserting control over different aspects of the digital (Floridi, 2020; Tretter, 2022). Typically, there is no absolute sovereign over a particular technology, with cooperation and competition for control of different aspects instead simultaneously taking place "horizontally" between states and "vertically" between states and non-state actors (Bradford, 2023).

Adopting a relational view of digital sovereignty is appropriate, but not without its problems. Most notably, efforts to expand the concept of digital sovereignty to include non-state actors raise questions over which of these actors possess legitimate authority. In traditional conceptions of sovereignty and by extension, deep but narrow understandings of digital sovereignty, the state derives legitimacy from a social contract with citizens. Sovereign authority of the state is absolute, meaning the authority to control digital

technologies "within" its territory is legitimised as one part of the wider social contract with citizens.[2] When this idea of a state's absolute authority is removed, who has the legitimacy to control digital technologies is less clear. Typically, broad but shallow conceptions of digital sovereignty have paid little attention to this normative question that emerges from expanding the concept and have instead focused solely on the descriptive question of who is asserting control over digital technologies.

A normative approach to digital sovereignty

A normative approach to digital sovereignty which establishes standards that any potential sovereign needs to meet to justify their control over digital technologies is preferable to taking a descriptive approach focused solely on the control itself. Such an account foregrounds evaluating the actions of those who control the technologies and in doing so, practically supports the development of policy interventions that leave digital technologies more closely aligned with citizens' interests. Following a normative approach is also consistent with traditional conceptions of sovereignty, which include justification of a sovereign's control over their territory, rather than merely describing it.

To develop a normative approach to digital sovereignty,[3] this article takes the commonly recognised features of sovereignty – (i) *authority*, (ii) *territoriality*, and (iii) *supremacy* (Philpott, 2016) – and adapts them for the realities of digital governance the 21st century. In doing so, this normative approach is designed to include a broader range of stakeholders as potential sovereigns, in line with a *broad but shallow* approach, while preserving the ethical considerations present in *deep but narrow* conceptions of digital sovereignty.

The first feature, *authority*, is understood here continuously with traditional definitions of sovereignty as *legitimate control* (Philpott, 2016). Control is an integral element of sovereign authority, as it enables states or others agents to enact their sovereignty in practice (Krasner, 2001). With respect to control, I follow Floridi's (2020, p. 371) definition of control as "the ability to influence something (e.g. its occurrence, creation, or destruction) and its dynamics (e.g. its behaviour, development, operations, interactions), including the ability to check and correct for any deviation from such influence." Control can be asserted through coercion, assent, seduction, co-optation, and so on (Agnew, 2005).

To possess authority, control must be legitimate rather than just an assertion of coercive power. However, defining legitimacy in the context of digital technologies is difficult. There is a rich political philosophy literature that discusses the sources that legitimate a state's authority, including descriptive accounts based on the acceptance of authority by a populus (C. Taylor, 1985) and normative accounts which establish a justification or benchmark for wielding political power (Simmons, 1999). Yet, these debates have been developed to legitimise a state's exclusive right to coercive force over citizens (Stone & Mittelstadt, 2024) rather than control over a specific digital technology. While the potential for technology companies to infringe on individual liberties through their control of platforms and digital products is growing, it still pales in comparison to the powers wielded by the state, which can make demands on citizens like conscription (Taylor, 2021). Discussions of legitimacy in political philosophy also focus on the state as the sole holder of this authority, which is of limited use for assessing whether efforts to control digital technologies by multiple different actors are legitimate. Thus, debates in political philosophy on sources of state legitimacy are instructive but cannot be directly applied.

Against this backdrop, legitimacy is understood in broader terms as the criteria that determine whether control over digital technologies is morally acceptable. For control to be legitimate, those impacted by digital technologies must provide consent (Raz, 1995),[4] which can be manifested in different ways (Beetham, 1991; Stoker, 1998). One source of consent is *input* into the decision-making process that underpins an institution or rule, for instance through democratic decision-making procedures or robust transparency and accountability mechanisms (Christiano, 1996). An alternative is the *output* or effectiveness of a rule or institution in effectively addressing public needs, such as through improving individual (Arneson, 2003) or collective (Wellman, 1996) fundamental rights.

Unlike political philosophy literature which seeks a single answer to the question of what legitimises a state's authority, no equivalent silver bullet is sought here. Technologies are controlled by different actors, with the sources drawn on for the legitimation of authority varied, and the

[2] This is not to say the state is considered infallible in a deep but narrow approach. The legitimacy of the state (i.e., rather than its legitimacy specifically in relation to digital technologies), how a state uses technology (i.e., rather than its authority to control them), and the degree to which citizens are obligated to accept a state's uses could be questioned.

[3] Note, the normative definition of digital sovereignty developed in this article builds on previous work from Floridi (2020) and Roberts et al. (2021).

[4] Consent accounts of legitimacy have been criticised as, in reality, citizens do not directly provide morally valid consent and proxies, like voting, are performed under coercion (Wellman, 1996). However, as mentioned above, the focus here is on consent to control certain aspects of the digital rather than to legitimise a state's supreme right to use coercive force, with morally valid consent for the former far more achievable than the latter.

appropriateness of justifications determined by the specific context (Walzer, 2008). Democratic states may point to democratic input and electoral promises related to technology as a source of input legitimacy, while authoritarian regimes like China and "rentier" Gulf states could premise their legitimacy to control technologies on wider social contracts that are underpinned by output performance and positive rights (Angle, 2002; Beblawi, 1987). In the same vein, private actors may point to their governance of technologies producing societally beneficial outputs through enhancing freedom of speech (e.g., through social media platforms) and economic rights (e.g., through remote work). Each of these sources of legitimacy will have drawbacks and potential criticisms. For example, introducing a democratic procedure specifically for governing a digital technology may contribute to justifying control as legitimate, but it may also come with efficiency costs or sub-optimal outcomes (Himmelreich, 2023). Meanwhile, relying on output as a source of legitimacy is controversial on account of the difficulty of measuring output legitimacy and scope for manipulation (Piattoni, 2010), as well as questions over whether capability to effectively manage justifies control (Stone & Mittelstadt, 2024).[5] The idea of legitimacy presented here is meant to serve as a discursive framework for comparing claims in different contexts, rather than offering an overarching formula that can be applied for identifying a single legitimate actor in every context. Accordingly, such drawbacks are not an inherent weakness; instead, they are simply factors to be considered when assessing competing claims of legitimate control.

Second, considering *territoriality*, traditional concepts of sovereignty tie a state's authority to a specific territory. This is because territory provides states with a functional space in which they can assert legitimate control (Storey, 2017). Until recently, this was intuitive as the provision of public goods had a built-in "territorial bias", but with globalisation and the shifting rationale of the state, authority is increasingly tied to other types of space. In particular, authority is functioning in networks (Agnew, 2005), like the internet, where state and non-state actors variously hold authority over different parts of this network (Pohle & Thiel, 2020). The nature of this network, which cuts through territorial boundaries, means states cannot meaningfully assert ultimate legitimate control (Roberts et al., 2021). While this reformulation of territorial authority may seem radical, there have historically been many types of polities where authority is not tied to strict territorial boundaries (e.g., hunter-gather tribes, seaborne empires) (Agnew, 2005).

In place of territory, the concept of digital sovereignty is concerned with digital technologies and the spaces they occupy. Here, digital technologies can be understood in line with Sheikh's (2022) "stack" of layers that computation depends on at a planetary scale. This includes the resource layer (e.g., rare minerals), the chips layer (e.g., semiconductors), the network layer (e.g., telecommunications infrastructure), the cloud layer (e.g., computation), the intelligence layer (e.g., AI algorithms), the applications layer (e.g., the mobile apps), and the connected device layer (e.g., a mobile phone). This encompassing understanding of digital sovereignty is inclusive of the central materials, hardware, and software, that make up common digital technologies. Some aspects of this stack, like territorially-bound datacentres, lend themselves more to state governance than, for instance, rules for platform governance which is less territorially bounded (Roberts et al., 2021).

Considering the final feature, *supremacy*, the state has traditionally been conceptualised as holding absolute authority over their territory. If understood in constitutional terms (i.e., the legal recognition of a state), then supreme sovereignty is still held by the state. This is because supremacy in these terms is a binary: a state is either recognised as a sovereign entity by its citizenry and other states or it is not (James, 1999). However, this constitutional understanding of sovereignty is not useful for assessing where states are being challenged by other entities asserting control or authority over digital technologies. Accordingly, supremacy here is assessed in terms of the "scope of matters over which the holder of authority is sovereign" (Philpott, 2016). When understood in this way, there have always been competing sources of authority for a state (e.g., the Church), but these institutions were largely overshadowed. However, the structure of the digital as complex globalised networks that transcend state borders leaves the idea of state supremacy over digital technologies largely redundant, particularly in smaller states with less developed governance capacities. In line with this, states are conceptualised as the *primus inter pares* because of their continued centrality to the organisation of contemporary life, but importantly, not the only digital sovereign (Roberts et al., 2021).

Thus, when speaking of digital sovereignty here, it is not in reference to a state's claim of supreme authority over digital technologies within a territory; something that is analytically redundant considering the realities of the 21st century. Instead, it is an umbrella term encompassing the competition and cooperation taking place to assert legitimate control over the stack of digital layers that underpin computation on a planetary scale. These "fights" for digital sovereignty are not necessarily territorially bounded and differ in their character based on the specific technologies and actors partaking. Unlike definitions which simply focus

[5] Most arguments based on output being insufficient sources of legitimacy are in reference to the state's absolute authority to use coercive force and as noted above, the focus here is on a lower threshold of coercion.

on control, the approach taken here is inherently normative as while some actors may assert coercive control over specific technologies, if there is not a high degree of legitimacy, then they cannot reasonably be said to possess or share digital sovereignty over that technology. They can only be described as quasi-sovereigns.

Big tech as sovereign actors over AI

To operationalise this normative approach to digital sovereignty, this section will consider the case study of "Big Tech" companies – such as Amazon, Apple, Meta, Google, and, Microsoft –as potential sovereign actors over artificial intelligence (AI).[6] Big Tech Companies are focused on because of their significant influence through providing crucial infrastructure that governments, citizens, and other private actors depend on (Rahman, 2018). No one feature has led to the success of these platforms. A new logic of "colonialist" accumulation based on data collection and profiling has facilitated a highly profitable business model based on the commodification of individuals' data (Couldry & Mejias, 2019; Zuboff, 2019). A networked market structure combined with resource-intensive products has hindered new competitors entering the market (Eisenmann et al., 2006; Luchetta, 2014). Platforms have also been able to reinvest a substantial proportion of their considerable profits, made from monetising users' data and harnessing their attention, into research and development (R&D) efforts, further reinforcing their market lead. As a result, it is increasingly difficult for governments and citizens to opt-out of using key services provided by these platforms (Taylor, 2021).

AI is focused on because it is a layer of the digital stack where Big Tech companies are increasingly focused, with a significant share of their profits invested in developing advanced AI systems or acquiring smaller AI firms, as was the case with Google's acquisition of DeepMind in 2014. Developing cutting-edge AI systems is a priority for Big Tech companies as these technologies allow them to compete in multiple sectors of the economy by offering data- and not expertise-driven insights (Sharon, 2020). Amazon, for instance, has utilised AI for personalisation in its e-commerce business, offers "off the shelf" AI tools as part of its cloud services, and relies on AI for its voice assistant Alexa to understand and respond to consumer requests. This trend towards Big Tech competing in multiple sectors has been exacerbated by the emergence of new "foundation models"; systems which are trained on broad data and are designed to be integrated into various downstream applications (Bommasani et al., 2022). As just one example, a purported 92% of Fortune 500 companies are using OpenAI's[7] products (Murgia & Hammond, 2024). This is not to say that Big Tech are the only companies developing AI, with other organisations also developing task-specific systems; nonetheless, Big Tech maintain an outsized influence over these technologies.

Developing AI products gives Big Tech companies significant control over access to cutting-edge technologies, as well as over many aspects of rulemaking related to these systems, including for highly value-laden decisions, such as what constitutes a "fair" outcome from a classification AI system or "hate speech" from a large language model output (Lazar, 2024). This concentration of power, combined with these companies' influence across sectors and jurisdictions, means any governance decisions made will have an amplified influence (Lazar, 2024; Taylor, 2021). Part of the reason Big Tech companies have been left with this governance responsibility can be explained by the complexity of cutting-edge systems and the opacity of business models which often leaves regulators ill-equipped to develop specific rules for AI (Aitken et al., 2022). Another element is the huge sums Big Tech spends lobbying governments to avoid regulatory oversight, with over €113 million spent lobbying the European Union in 2023 (Jones, 2023).

If sovereignty is understood in descriptive terms as control over these technologies, then it is reasonable to conclude that Big Tech companies are playing a sovereign role. However, when the normative criterion of legitimacy is considered, the status of Big Tech as sovereigns becomes more ambiguous. States typically possess some degree of input and/or output legitimacy from their social contracts with citizens, which encompasses control over digital technologies (Schmidt, 2020). Big Tech companies do not have an equivalent social contract with citizens from which to derive their legitimacy, giving rise to what could be called a "legitimacy gap" unless public consent is provided through other means. Without this consent, AI could be controlled in a way that is misaligned with citizens' interests, leading to the potential erosion of rights, increasing inequalities, and so on. To determine whether Big Tech companies can be considered sovereign actors over AI, it is necessary to assess the degree to which they possess input and output legitimacy regarding these technologies.

In terms of *input*, efforts to ensure decision-making over how these technologies are controlled is inclusive have been limited. Those deciding how AI should be designed,

[6] AI is understood here in line with the widely adopted definition put forward by the OECD: "An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment" (OECD, 2024).

[7] While companies like OpenAI and Anthropic are newer players, they rely on partnerships with, and funding from, traditional Big Tech.

deployed, and governed are a narrow, demographically undiverse subset of the population (Harrison, 2019). It could be argued that input legitimacy is grounded in the terms and conditions that users sign up to, with individuals simply able to use a different product if they are dissatisfied with these terms. But as has been shown by numerous studies, these terms and conditions are a form of "convention consent" where people cooperate despite not necessarily agreeing rather than an "endorsement consent" where people agree to the authority due to believing it is justified (Taylor, 2021). This is because in practice, a meaningful alternative is not always present or the "switching costs" of moving to an alternative service are high for consumers. Furthermore, in many cases, those subjected to unethical practices are not the consumer of the product, meaning they are unable to consent (Stone & Mittelstadt, 2024). Take the law enforcement space, where AI systems developed by private companies are increasingly procured by police forces to profile and allegedly "predict" offences and offenders, frequently resulting in discrimination and due process violations (Angwin et al., 2016). In this case, it is police forces and not those who are being subjected to the systems that make procurement decisions.

To strengthen input legitimacy, some efforts have been made by Big Tech companies to more directly solicit consumer opinions on how AI should be governed. However, engagement exercises have generally been small-scale and typically unrepresentative.[8] Even if these exercises were expanded to be more inclusive, there is reason for scepticism, as companies have typically focused public engagement on vague issues like value-alignment rather than questioning core business practices. By foregrounding these cosmetic details while ignoring *a priori* questions over whether the technology or business practice itself is legitimate (Barocas et al., 2023), this type of "democratisation" can be considered little more than "ethics washing" (Bietti, 2020). A further criticisms of this type of democratisation is that it is simply not a suitable method for legitimising private sector use of AI, as it is resource intensive and not effective for addressing morally thick issues, like inequality (Himmelreich, 2023).

In the absence of direct democratic procedures, improving transparency and accountability could be looked to as mechanisms that foster greater input legitimacy. AI companies have taken steps to make inputs to their governance processes more transparent, for instance, by releasing ethical principles or mission statements that the companies theoretically should be held to. Nevertheless, scepticism here is also warranted. Studies have highlighted how the governance of AI by Big Tech companies – from inputs into making AI systems, to the properties of a system, and potential downstream impacts – are highly opaque (*Foundation Model Transparency Index*, n.d.; Stone & Mittelstadt, 2024). When external researchers have attempted to scrutinise systems or practices, they have often been met with bans from platforms or threats of legal action (Brandom, 2021). Accountability mechanisms are also present, like oversight boards designed to hold companies accountable to their principles or statements. But the effectiveness of internal accountability mechanisms can also be questioned, particularly when commercial incentives run counter to interventions, as was shown with the failed firing of OpenAI's CEO Sam Altman by the organisation's Board.

On the surface, *output* appears a more promising source of legitimacy for Big Tech companies' control over AI. Their products have addressed public needs, such as through supporting positive rights, like improving education and healthcare, while also arguably enhancing negative rights like freedom of information and expression. These are benefits for both individuals and society more broadly. Indeed, when governments have attempted to increase oversight over Big Tech, there have been several noteworthy examples of consumers rallying against regulatory proposals perceived to threaten their access to beneficial products. For example, Uber users have voiced their opposition to regulatory interventions in some localities due to fears over less availability and affordability (Culpepper & Thelen, 2020). Yet, it is important not to confuse consumer support with output legitimacy, with the former focused on the opinion of individuals rather the moral permissibility of the practice itself (Simmons, 1999). When considering the track record of platform companies, a trend can often be observed of these companies initially running unprofitable business models subsidised by venture capital, which makes them popular among consumers, while also making it unviable for other companies to compete. When sufficient market share has been achieved, policies are changed to ensure business profitability, which can be resented by consumers who no longer have meaningful market choice (Lehdonvirta, 2022).

Furthermore, although AI systems developed by Big Tech have brought about numerous benefits, their failures to effectively govern these systems have also led to myriad harms. This includes facilitating bias and discrimination in hiring decisions, disinformation that has contributed to ethnic violence, the erosion of privacy from data collection and surveillance, to name just a few. Certainly, governments have been imperfect in their efforts to strengthen output legitimacy through increasing oversight over Big Tech companies. Chomanski (2021), for example, argues that greater government control of digital technologies does not necessarily lead to better outcomes on account of parochial interests of competing government departments and

[8] For example, see OpenAI's Democratic Inputs to AI Grant Programme (OpenAI, 2024).

limited technical competencies of regulators. In turn, this could lead to overregulation and harmful unintended consequences, like solidifying the position of establishment firms, reducing investment for local startups, and reducing consumer welfare. Nonetheless, when considered in the context of the potential harms associated with AI, it is reasonable to consider most regulatory initiatives that have emerged as serving the public benefit to at least some degree, given the alternative of flawed industry self-regulation.

Not all the Big Tech companies are the same, yet when their control over AI is considered as a whole and compared with the authority over AI generally held by states, a legitimacy gap can be observed on account of limited citizen input and ambiguous public benefits from the outputs produced. Because of this, Big Tech companies can only currently be considered as quasi-sovereigns over AI as they miss the crucial ingredient of legitimacy that is held by genuinely sovereign actors.

Conclusion

Understanding digital sovereignty in terms of control facilitates a rich descriptive account of how different actors are competing to control digital technologies, but it provides no mechanism for evaluating the legitimacy of these actions. This is problematic as no distinction is made between coercive and legitimate assertions of control over digital technologies. Labelling actors who use coercive power to control digital technologies as "sovereigns" does a disservice to the term, as it ignores the duties that a sovereign is expected to perform. Failure to distinguish between these types of control is also concerning given the trends identified in this article, including the growing remit of the private sector, the apparent decreasing capacity of governments to keep Big Tech in check, the wide-ranging domains which are underpinned by digital technologies, and the harms brought about Big Tech companies' products.

In contrast, conceptualising digital sovereignty as authority (i.e., legitimate control) foregrounds assessments of the desirability of actions rather than just describing their nature or efficacy. This normative approach considers digital sovereignty as something shared between different actors rather than framing it in binary terms. It centres analysis on identifying which areas of control over digital technologies are lacking in legitimacy and in turn, what can be done to rectify "legitimacy gaps". Such a normative framing supports practical policy making designed to ensure more societally beneficial outcomes from digital technologies.

In the case of digital sovereignty over AI analysed in this article, Big Tech companies are asserting significant control but have severe deficiencies in input legitimacy and possess a questionable amount of output legitimacy. To rectify this legitimacy gap, it is necessary to strengthen the legitimacy of Big Tech companies' practices and/or for governments to increase oversight of these technologies based on the legitimacy they possess. Strengthening input into decision-making is central, as it will likely also have a downstream impact on the legitimacy of outputs (Piattoni, 2010). In particular, mechanisms that meaningfully improve transparency and accountability, like third party technical audits and disclosure requirements, would support scrutiny ofbusiness practices and systems in a manner that is most likely to facilitate changes that address public needs.

For digital sovereignty to be an analytically useful term for addressing digital challenges in the 21st century, it is integral that the twin traps of understanding the term as solely applying to the state on the one hand, and simply referring to control over the digital on the other, are avoided. A normative approach to digital sovereignty achieves this.

Acknowledgements I would like to thank Josh Cowls, Emmie Hine, Rosaria Taddeo, and Luciano Floridi for discussions and collaborative projects on digital sovereignty over many years, which were invaluable in shaping this paper.

Data availability The author confirms that all data generated or analysed during this study are included in this published article.

References

Acharya, U. D. (2013). Globalization and Hegemony Shift: Are States Merely agents of corporate capitalism. *Boston College Law Review*, *54*, 937.

Agnew, J. (2005). Sovereignty regimes: Territoriality and State Authority in Contemporary World politics. *Annals of the Association of American Geographers*, *95*(2), 437–461. https://doi.org/10.1111/j.1467-8306.2005.00468.x

Aitken, M., Leslie, D., Ostmann, F., Pratt, J., Margetts, H., & Dorobantu, C. (2022). *Common Regulatory Capacity for AI*. The Alan Turing Institute. https://doi.org/10.5281/zenodo.6838946

Angle, S. (2002). *Human rights in Chinese Thought: A Cross-cultural Inquiry*. Cambridge University Press.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica, May*, *23*.

Arneson, R. J. (2003). Defending the purely instrumental account of democratic legitimacy. *Journal of Political Philosophy*, *11*(1), 122–132. https://doi.org/10.1111/1467-9760.00170

Baezner, M., & Robin, P. (2018). *Trend Analysis: Cyber Sovereignty and Data Sovereignty*. Center for Security Studies (CSS), ETH Zürich. /en/services/digital-library/publications/publication.html.

Barlow, J. P. (1996). *A declaration of the independence of Cyberspace*. Electronic Frontier Foundation. https://www.eff.org/cyberspace-independence

Barocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and Machine Learning: Limitations and opportunities*. MIT Press.

Basu, A., Hickok, E., & Singh Chawla, A. (2019). The Localisation Gambit: Unpacking policy moves for the sovereign control of data in India—The Centre for Internet and Society. *The Centre for Internet and Society*. https://cis-india.org/internet-governance/blog/the-localisation-gambit-unpacking-policy-moves-for-the-sovereign-control-of-data-in-india

Beblawi, H. (1987). The Rentier State in the Arab World. *Arab Studies Quarterly*, *9*(4), 383–398.

Beetham, D. (1991). Dimensions of State Legitimacy. In D. Beetham (Ed.), *The Legitimation of Power* (pp. 117–160). Macmillan Education UK. https://doi.org/10.1007/978-1-349-21599-7_5

Bietti, E. (2020). From ethics washing to ethics bashing: A view on tech ethics from within moral philosophy. *Proceedings of the 2020 Conference on Fairness Accountability and Transparency*, 210–219. https://doi.org/10.1145/3351095.3372860

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., & Liang, P. (2022). *On the Opportunities and Risks of Foundation Models* (arXiv:2108.07258). arXiv. http://arxiv.org/abs/2108.07258.

Bradford, A. (2023). *Digital empires: The Global Battle to Regulate Technology*. Oxford University Press.

Brandom, R. (2021, August 13). *Facebook shut down German research on Instagram algorithm, researchers say*. The Verge. https://www.theverge.com/2021/8/13/22623354/facebook-instagram-algorithm-watch-research-legal-threat

C Demchak, C. (2016). Uncivil and Post-western Cyber Westphalia: Changing interstate power relations of the cybered age. *The Cyber Defense Review*, *1*(1), 49–74. JSTOR.

Chomanski, B. (2021). The missing ingredient in the case for regulating big tech. *Minds and Machines*, *31*(2), 257–275. https://doi.org/10.1007/s11023-021-09562-x

Christiano, T. (1996). *The Rule Of The Many: Fundamental Issues In Democratic Theory*. Routledge & CRC Press. https://www.routledge.com/The-Rule-Of-The-Many-Fundamental-Issues-In-Democratic-Theory/Christiano-Christiano/p/book/9780813314556

Connolly, W. E. (2004). The complexity of Sovereignty. *Sovereign lives*. Routledge.

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking Big Data's relation to the contemporary subject. *Television & New Media*, *20*(4), 336–349. https://doi.org/10.1177/1527476418796632

Couture, S., & Toupin, S. (2019). What does the notion of sovereignty mean when referring to the digital? *New Media & Society*, *21*(10), 2305–2322. https://doi.org/10.1177/1461444819865984

Culpepper, P. D., & Thelen, K. (2020). Are we all Amazon Primed? Consumers and the politics of platform power. *Comparative Political Studies*, *53*(2), 288–318. https://doi.org/10.1177/0010414019852687

Deibert, R., Palfrey, J., Rohozinski, R., & Zittrain, J. (2008). *Access denied: The practice and policy of global internet filtering*. MIT Press.

Demchak, C. C., & Dombrowski, P. J. (2014). Rise of a Cybered Westphalian Age: The Coming Decades. In M. Mayer, M. Carpes, & R. Knoblich (Eds.), *The Global Politics of Science and Technology—Vol. 1: Concepts from International Relations and Other Disciplines* (pp. 91–113). Springer. https://doi.org/10.1007/978-3-642-55007-2_5

Eisenmann, T., Parker, G., & Alstyne, M. W. V. (2006). Strategies for two- Sided markets. *Harvard Business Review*, 12.

Falkner, G., Heidebrecht, S., Obendiek, A., & Seidl, T. (2022). Digital Sovereignty—Rhetoric and Reality. *Framework Paper for the Online Conference 28–29 April 2022*.

Floridi, L. (2020). The fight for Digital Sovereignty: What it is, and why it matters, especially for the EU. *Philosophy & Technology*, *33*(3), 369–378. https://doi.org/10.1007/s13347-020-00423-6

Foundation Model Transparency Index. (n.d.). CRFM Stanford. Retrieved 13 September 2024, from https://crfm.stanford.edu/fmti/May-2024/index.html

Franzese, P. W. (2009). Sovereignty in Cyberspace: Can it exist Cyberlaw Edition. *Air Force Law Review*, *64*(1), 1–42.

Fraser, E. (2016). Data localisation and the balkanisation of the internet. *SCRIPTed*, *13*(3), 359–373. https://doi.org/10.2966/scrip.130316.359

Ganz, A., Camellini, M., Hine, E., Novelli, C., Roberts, H., & Floridi, L. (2024). Submarine Cables and the risks to Digital Sovereignty. *SSRN Scholarly Paper 4693206*. https://doi.org/10.2139/ssrn.4693206

Haché, A. (2014). Technological Sovereignty. *Mouvements*, *79*(3), 38–48.

Harrison, S. (2019, October). Five Years of Tech Diversity Reports—And Little Progress. *Wired*. https://www.wired.com/story/five-years-tech-diversity-reports-little-progress/

Henriksen, A. (2019). The end of the road for the UN GGE process: The future regulation of cyberspace. *Journal of Cybersecurity*, *5*(1). https://doi.org/10.1093/cybsec/tyy009

Himmelreich, J. (2023). Against democratizing AI. *AI & SOCIETY*, *38*(4), 1333–1346. https://doi.org/10.1007/s00146-021-01357-z

Hummel, P., Braun, M., Augsberg, S., Dabrock, P., Erlangen-Nürnberg, F. A. U., & Gießen, J. L. U. (2018). *Sovereignty and data sharing*. *2*, 10.

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021). Data sovereignty: A review. *Big Data & Society*, *8*(1), 2053951720982012. https://doi.org/10.1177/2053951720982012

James, A. (1999). The Practice of Sovereign Statehood in Contemporary International Society. *Political Studies*, *47*(3), 457–473. https://doi.org/10.1111/1467-9248.00212

Jones, M. G. (2023, September 11). *Tech companies spend more than €100 million a year on EU lobbying*. Euronews. https://www.euronews.com/my-europe/2023/09/11/tech-companies-spend-more-than-100-million-a-year-on-eu-digital-lobbying

Krasner, S. D. (2001). Abiding Sovereignty. *International Political Science Review*, *22*(3), 229–251. https://doi.org/10.1177/0192512101223002

Krasner, S. D. (2007). Sovereignty. In *the Blackwell Encyclopedia of Sociology*. *American Cancer Society*. https://doi.org/10.1002/9781405165518.wbeoss213

Kukutai, T., & Taylor, J. (2016). *Indigenous Data Sovereignty: Toward an agenda*. ANU. https://doi.org/10.22459/CAEPR38.11.2016

Lazar, S. (2024). Legitimacy, Authority, and Democratic Duties of Explanation. In D. Sobel & S. Wall (Eds.), *Oxford Studies in Political Philosophy Volume 10* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198909460.003.0002

Lehdonvirta, V. (2022). *Cloud empires: How digital platforms are overtaking the state and how we can regain control*. MIT Press.

Luchetta, G. (2014). Is Google a two-sided market? *Journal of Competition Law & Economics*, *10*(1), 185–207. https://doi.org/10.1093/joclec/nht026

Magalhães, J. V., & Couldry, N. (2021). Giving by taking away: Big tech, data colonialism and the reconfiguration of social good. *International Journal of Communication*, *15*, 343–362.

Meyer, P. (2020). Norms of Responsible State Behaviour in Cyberspace. In M. Christen, B. Gordijn, & M. Loi (Eds.), *The Ethics of Cybersecurity* (pp. 347–360). Springer International Publishing. https://doi.org/10.1007/978-3-030-29053-5_18

Morgenthau, H. J. (1948). The Problem of Sovereignty reconsidered. *Columbia Law Review*, *48*(3), 341–365. https://doi.org/10.2307/1118308. JSTOR.

Mueller, M. L. (2019). Against Sovereignty in Cyberspace. *International Studies Review*. https://doi.org/10.1093/isr/viz044

Murgia, M., & Hammond, G. (2024, February 15). *Can OpenAI create superintelligence before it runs out of cash?*https://www.ft.com/content/6314d78d-81f3-43f5-9daf-b10f3ff9e24f

OECD (2024). *Explanatory memorandum on the updated OECD definition of an AI system* (OECD Artificial Intelligence Papers 8; OECD Artificial Intelligence Papers, Vol. 8). https://doi.org/10.1787/623da898-en

OpenAI. (2024, January 16). *Democratic inputs to AI grant program: Lessons learned and implementation plans*. https://openai.com/index/democratic-inputs-to-ai-grant-program-update/

Perritt, H. (1998). The internet as a threat to Sovereignty? Thoughts on the internet's role in strengthening National and Global Governance. *5 Indiana Journal of Global Legal Studies*, *423*(1998)), 52. https://www.repository.law.indiana.edu/ijgls/vol5/iss2/4

Philpott, D. (2016). Sovereignty. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2016). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2016/entries/sovereignty/

Piattoni, S. (2010). Output Legitimacy. In S. Piattoni (Ed.), *The Theory of Multi-level Governance: Conceptual, Empirical, and Normative Challenges* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199562923.003.0012

Pizzul, D., & Veneziano, M. (2023). Digital sovereignty or sovereignism? Investigating the political discourse on digital contact tracing apps in France. *Information Communication & Society*, *0*(0), 1–17. https://doi.org/10.1080/1369118X.2023.2232840

Pohle, J., & Thiel, T. (2020). Digital sovereignty. *Internet Policy Review*, *9*(4). https://policyreview.info/concepts/digital-sovereignty

Rahman, K. S. (2018). *Regulating Informational Infrastructure: Internet Platforms as the New Public Utilities* (SSRN Scholarly Paper 3220737). https://papers.ssrn.com/abstract=3220737

Raz, J. (1995). Rights and politics. *Indiana Law Journal*, *71*(1), 27–44.

Roberts, H., Cowls, J., Casolari, F., Morley, J., Taddeo, M., & Floridi, L. (2021). Safeguarding European values with digital sovereignty: An analysis of statements and policies. *Internet Policy Review*, *10*(3). https://policyreview.info/articles/analysis/safeguarding-european-values-digital-sovereignty-analysis-statements-and-policies

Sargsyan, T. (2016). Data localization and the Role of Infrastructure for Surveillance, privacy, and security. *International Journal of Communication*, *10*(0), Article0.

Savelyev, A. (2016). Russia's new personal data localization regulations: A step forward or a self-imposed sanction? *Computer Law & Security Review*, *32*(1), 128–145. https://doi.org/10.1016/j.clsr.2015.12.003

Schmidt, V. A. (2020). Conceptualizing legitimacy: Input, output, and Throughput. In V. A. Schmidt (Ed.), *Europe's Crisis of Legitimacy: Governing by rules and ruling by numbers in the Eurozone* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198797050.003.0002

Schmitt, M. N. (Ed.). (2017). *Tallinn manual 2.0 on the international law applicable to cyber operations* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781316822524

Sharon, T. (2020). Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-020-09547-x

Sheikh, H. (2022). European Digital Sovereignty: A Layered Approach. *Digital Society*, *1*(3), 25. https://doi.org/10.1007/s44206-022-00025-z

Simmons, A. J. (1999). Justification and legitimacy. *Ethics*, *109*(4), 739–771. https://doi.org/10.1086/233944

Stoker, G. (1998). Governance as theory: Five propositions. *International Social Science Journal*, *50*(155), 17–28.

Stone, J., & Mittelstadt, B. (2024). *Legitimate Power, Illegitimate Automation: The problem of ignoring legitimacy in automated decision systems* (arXiv:2404.15680). arXiv. http://arxiv.org/abs/2404.15680

Storey, D. (2017). States, territory and sovereignty. *Geography*, *102*(3), 116–121.

Strange, S. (1995). The defective state. *Daedalus*, *124*(2), 55–74. JSTOR.

Strange, S. (1996). *The retreat of the state: The Diffusion of Power in the World Economy*. Cambridge University Press.

Taddeo, M. (2017). Cyber conflicts and Political Power in Information Societies. *Minds and Machines*, *27*(2), 265–268. https://doi.org/10.1007/s11023-017-9436-3

Taylor, C. (Ed.). (1985). Legitimation crisis? *Philosophical papers: Volume 2: Philosophy and the Human sciences* (Vol. 2, pp. 248–288). Cambridge University Press. https://doi.org/10.1017/CBO9781139173490.011

Taylor, L. (2021). Public actors without public values: Legitimacy, domination and the regulation of the Technology Sector. *Philosophy & Technology*, *34*(4), 897–922. https://doi.org/10.1007/s13347-020-00441-4

Timmers, P. (2019). Ethics of AI and Cybersecurity when Sovereignty is at Stake. *Minds and Machines*, *29*(4), 635–645. https://doi.org/10.1007/s11023-019-09508-4

Tretter, M. (2022). Sovereignty in the Digital and contact tracing apps. *Digital Society*, *2*(1). https://doi.org/10.1007/s44206-022-00030-2

Walzer, M. (2008). *Spheres of justice: A defense of pluralism and Equality*. Basic Books.

Wellman, C. H. (1996). Liberalism, Samaritanism, and political legitimacy. *Philosophy & Public Affairs*, *25*(3), 211–237. https://doi.org/10.1111/j.1088-4963.1996.tb00040.x

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for the future at the new frontier of power*. Profile Books.