Protocolizing the Governance of Public Goods: Using Social Media as an Example

1. Introduction

The tragedy of the commons, a concept introduced by Garrett Hardin in 1968, describes the depletion of shared resources due to individuals acting independently and in their own self-interest, ultimately harming the collective good (Hardin, 1968). This phenomenon has been observed across various sectors, including environmental management, economics, and even the digital realm (Ostrom, 1990). As the world becomes increasingly interconnected and digitalized, the tragedy of the commons has manifested itself in new ways, revealing limitations in traditional governance mechanisms and highlighting the need for novel approaches to resource management and regulation.

Traditional mechanisms, such as centralized control and privatization, have struggled to address the unique challenges posed by the tragedy of the commons in the digital age. One notable example of failure in this context is social media, which has become an essential platform for communication, information dissemination, and social interaction (Tufekci, 2017). However, the unchecked proliferation of misinformation, privacy concerns, and the monopolistic practices of major social media platforms have exposed the inadequacy of existing regulatory frameworks in preventing the negative consequences of the tragedy of the commons in the digital sphere (Zuboff, 2019).

In light of these challenges, protocolization has emerged as a potential solution to mitigate the tragedy of the commons in digital environments (De Filippi & Loveluck, 2016). As a process of formalizing governance structures and rules through algorithms, protocols, and shared standards, protocolization offers the prospect of more equitable and transparent management of shared resources. By applying Elinor Ostrom's Institutional Analysis and Development (IAD) framework to common-pool resource problems (Ostrom, 2005), this study aims to explore the possibilities and limitations of protocolization in addressing the governance challenges of social media platforms. In doing so, the paper delves into the paradox of programmability and irreproducibility, examining the potential risks of systematizing and institutionalizing bias (Crawford, 2016), while considering alternative approaches that reimagine the relationship between technology and democracy (Winner, 1986).

1.1. The tragedy of commons

Many of today's pressing global issues can be attributed to the tragedy of the commons, a phenomenon where individuals acting in their own self-interest ultimately deplete a shared resource (Hardin, 1968). This concept, prevalent across various sectors of the economy and society, can exacerbate existing problems and create new challenges in resource management and distribution (Ostrom, 1990).

The tragedy of the commons is evident in radical markets, where monopolies over land hinder the development of essential infrastructure like transportation, limiting access to crucial services (Posner & Weyl, 2018). In the realm of technology, concentration effects are observed, with a few key players dominating markets such as email services and internet access via IPv6 (Zuboff, 2019). These monopolistic practices raise concerns about privacy, control, and equitable distribution of resources. Furthermore, environmental issues like deforestation exemplify the

tragedy of commons, as uncoordinated exploitation of forests by multiple stakeholders degrades vital ecosystems (MEA, 2005).

One example from Ostrom's work is the management of fisheries (Ostrom, 1990). In an open-access fishery, fishers might act in their own self-interest by catching as many fish as they can, without considering the impact on the fish population. This overfishing can lead to a decline in fish stocks and eventually the collapse of the fishery, which is detrimental to the livelihoods of fishers and the sustainability of the ecosystem (Worm et al., 2009).

1.2. Problems in Traditional Mechanism

The dichotomy between markets and government has long been considered the optimal structure for managing resources (Coase, 1960). Markets excel at producing and exchanging private goods, while a single government department is believed to provide public goods most efficiently (Samuelson, 1954). This perspective posits that a hierarchical government is necessary for guiding citizens and officials towards the efficient production of public goods, such as peace and security. However, this binary approach fails to account for the complex dynamics within private enterprises and the diverse range of institutional arrangements that humans use to manage public goods and common-pool resources (Ostrom, 1990).

Traditional solutions have encountered challenges in addressing issues that the government/market framework is ill-equipped to handle. One such challenge is the limited understanding of governments in certain situations, such as farmers' decision-making processes regarding arable land allocation (Feder & Feeny, 1991). Additionally, market-provided solutions can lead to severe negative externalities that hinder social collaboration and further development (Arrow, 1963). As a result, alternative approaches have been explored to better address these shortcomings.

Elinor Ostrom's work offers valuable insights into managing common-pool resources more effectively (Ostrom, 1990). She introduces a "fifth" game design, where a consensus-based penalty mechanism deters individuals from betraying the common interest. Moreover, communication and trust play a crucial role in fostering cooperation among resource users (Ostrom & Walker, 1991). Lastly, the concept of radical markets challenges traditional notions of private property, suggesting that resources should automatically flow to the most efficient user groups (Posner & Weyl, 2018). These alternative perspectives highlight the importance of rethinking traditional government/market frameworks in order to develop more nuanced and effective resource management strategies. Her Nobel Prize-winning research demonstrates that local communities can effectively manage resources through agile and responsive governance structures (Ostrom, 2009). The ability of individuals to "vote with their feet" and exit situations that do not work for them fosters loyalty and adherence to communal rules, mitigating the tragedy of commons (Tiebout, 1956). Consequently, a mixed approach that combines centralized and decentralized governance structures may hold the key to addressing the tragedy of commons and achieving sustainable resource management (Ostrom, 2009). By balancing the strengths of both governance systems, we can work towards overcoming the detrimental consequences of the tragedy of commons and build a more equitable and sustainable world.

1.3. Social Media as an Example of Failure

The relative uprising of the materialized and atomized lifestyle in post-modern society, coinciding with the massive invasion of centralized social media platforms (e.g., Twitter, Facebook,

Instagram), is not a coincidence. Currently, social media monopolies control the ownership of social graphs, meaning all data regarding bidirectional and multidirectional relationships are stored in centralized servers, enabling them to profit from monetizing the social graph (Zuboff, 2019). As Ovadya (2022) states, traditional social media attempts to extend the length of users' engagement on platforms by predicting content that triggers intense emotions, including hysteria, fear, anger, and anxiety, as user volume positively correlates with advertisement spot fees. Although some argue that it is rational and legitimate for businesses to maximize profits by all means, distributing toxic content has negative externalities on society by distracting citizens from significant social discourse and intensifying social polarization.

The true representation behind them is profit-mongering MNEs, the exact opposite of public interest. As a seemingly obvious consequence, we witness a decay of the public sphere: an arena where citizens come together, exchange opinions regarding public affairs, discuss, deliberate, and eventually form public opinion (Habermas, 1991). As more activities and community gatherings shift from in-real-life to virtual spaces, digital spaces become increasingly important venues for civic activities and the public sphere, places where individuals gather as equals to promote democratic participation and accountability (Dahlgren, 2005).

The solution to the above problem does not lie in communism, where the regime forces individuals to surrender individual happiness for the welfare of the whole society using authoritarian power, without realizing a fatal problem - the optimization of resource allocation must maximize the aggregate welfare of every individual. Otherwise, the "servant to the pleasure of the people," as the "communist servant" puts it, can easily be manipulated into the "servant to the authoritarian regime" (Arendt, 1951). "People" without flesh and blood are just numbers. Numbers do not speak for themselves. Thus, the answer is neither prioritizing ethics and national interest over individual welfare nor calling for appeasement of commercialism's penetrating force. Instead, we should admit the existence of corruption of private interest into the public sphere by tokenizing the attention economy (Lanier, 2013).

2. Is protocolization a solution?

2.1. What is protocolization?

Protocolization refers to the process of formalizing governance structures, rules, and decision-making processes through the use of protocols, algorithms, and shared standards. In essence, protocolization transforms complex social, political, or economic interactions into a system of structured rules and procedures that can be enforced and executed through digital means (De Filippi & Hassan, 2016).

In the context of common-pool resources and shared goods, protocolization offers a potential solution to address the challenges associated with their management and distribution. By automating certain aspects of governance and providing clear, transparent rules, protocolization aims to promote fair, efficient, and decentralized management of shared resources (Ostrom, 1990). This approach can help mitigate issues related to the tragedy of the commons and foster cooperation among diverse stakeholders.

2.2. Solution to the Common pool resources problems: applying IAD framework

First, the polycentric structure of the protocol can be set to have different ranges of impact, providing a framework for multi-layered governance systems, enabling various levels of impact to be established (Ostrom, 2010). For example, small "protocol," such as the expectation of behavior in greeting between two individuals, and large "protocol," like the nuclear launch protocols, can have different effects, often coordinating and mediating the evolving relationships between different stakeholders. Boundaries between related spaces are delineated, allowing for effective collaboration and decision-making. Within this framework, local communities are often the most knowledgeable actors, possessing unique insights into the specificities of their respective domains (Ostrom, 1990).

Secondly, the protocol exhibits a natural tendency towards decentralization, which offers a more reliable alternative to human-automated systems (Swan, 2015). In individual sovereignty, Dale (1999) describes the essence of the modern state as a monopoly of violence and creating coordination between individuals. This use of force, both domestically and internationally, has often been justified by the need to protect national interests or uphold a particular ideology. Moreover, fiat money, a form of currency that is not backed by a physical commodity, has allowed governments to manipulate economies and exert influence over their citizens. Inflation, for example, can erode the value of money and diminish individual purchasing power, further consolidating the state's authority. On the contrary, automated protocols encourage voluntary participation, with agents choosing to enter the system based on a shared commitment to consensus, avoiding reliance on violence or centralized control (Buterin, 2014). This fosters a more inclusive and democratic environment, in which individuals can work together towards common goals without the imposition of hierarchy or external control.

Thirdly, protocol offers a unique approach to managing complexity by creating action situations tailored to diverse contexts (Agrawal, 2001). By specifying and devising a vast array of rules suited to the specificities of local resource systems, the protocol is capable of transforming initially chaotic environments into orderly frameworks. This process entails identifying optimal rules to absorb variety and manage complexity. Ashby's Law of Requisite Variety underscores this principle, stating that a system must possess the necessary variety to handle complexity; otherwise, chaos will ensue (Ashby, 1956).

Fourthly, the protocol's inherent mutability enables it to adapt to changing circumstances and foster coherence between different protocols (De Filippi & Hassan, 2016). This coordination between different protocols is achieved through the implementation of triggering mechanisms between protocols, which standardize interactions in terms of input and output. Meta-governance, designed to coordinate various layers of the protocol, serves as a guiding principle in this process, ensuring that the system remains agile and responsive to evolving contexts (Sørensen & Torfing, 2009).

Finally, in cases where objections arise, the protocol offers a practical solution in the form of voluntary exit. In the book of exit, voice, and loyalty, Hirschman (1970) states that exit enables individuals to exercise their autonomy and freedom of choice. By offering the option to leave, people can preserve their agency and ensure that their participation remains voluntary and consensual. This option empowers individuals and groups to opt-out of the system if they find it incompatible with their values or objectives. This mechanism not only ensures the preservation of individual autonomy but also contributes to the protocol's overall resilience and adaptability, as it continuously evolves to accommodate diverse perspectives and needs.

3. The paradox of programmatically and irreproducibility

3.1. Schelling point, power of imagination

In the article 'The Unreasonable Sufficiency of Protocols', Rao et al. (2023) state that protocols have the potential to inspire sufficient voluntary commitment and participation to counterbalance the forces of withdrawal and exit. The coordination between protocols may move the status quo to a Schelling point, which leads to a more intricate pattern of coordinated contribution and engagement.

When individuals or groups converge on a Schelling point, they can create positive externalities by collectively generating benefits for others. This can lead to a more efficient allocation of resources, improved social welfare, and the reinforcement of social norms and values. Conversely, the presence of positive externalities can help to establish and maintain Schelling points, as the benefits generated by such externalities can incentivize individuals to cooperate and coordinate their actions around a common point of understanding or agreement.

However, in order to shift the current situation to the Schelling point, it partially depends on the creation of a positive expectation first among a small, like-minded group of people, then spread into the majority of people, forming the basis of a transformative society and evolving society. This contagious power of imagination is associated with specific time, space, and environment, which cannot be fully captured by language. The fundamental idea is that any algorithm, essentially a form of written language, cannot encompass all aspects that humans value, especially with respect to the relation of connection between humans, the environment, and common memories.

3.2. Complexity of value

The concept of fragility of value posits that the loss of even a small portion of the rules that constitute our values may lead to outcomes that the majority of people would consider unacceptable. This notion can be likened to the situation where dialing nine out of ten phone digits correctly does not connect an individual to a person 90% similar to their friend. An example illustrating the fragility of value is a future where, except for novelty, all our values might lead to a reality filled with individuals perpetually reliving a single optimal experience. This demonstrates the inherent complexity of value and the challenges in capturing the full range of human values.

No algorithm can comprehensively capture all aspects that are important to humans, which reflects the complexity of value. Consequently, it is impossible for any algorithm to determine what is "right" or "wrong" in a meaningful, human sense of these terms. This limitation highlights the inherent difficulties in developing algorithms or systems that can effectively address the diverse and nuanced values held by individuals and societies.

The creation of value in a public context often involves individuals making decisions in terms of the behavior of other agents co-existing in the environment. This means that all of our values, except novelty, might yield a future full of individuals replaying only one optimal experience through all eternity. In other words, the process of value co-creation is associated with a massive amount of randomness and aggregation of personal choice, which cannot be programmed into a protocol.

3.3. Risk of systematizing and institutionalizing bias.

Protocols, while inherently indiscriminate and arguably fairer than human judgment, are only as effective as the data they are trained on. Therefore, when designing algorithms, it is essential to ensure that the underlying data and designs contain minimal bias. When protocols are designed to become the foundation for an increasing number of social systems and institutions, it is vital to recognize and address the inherent risk of bias.

Systems designed by a small, homogeneous group may not function effectively for a diverse population. It is important to remember that protocols represent consensus of social interaction, power structure and distribution mechanism. As we continue to grant more agency to these systems, we must take the risk of bias seriously; otherwise, we risk creating a new world order that disproportionately benefits a select group of individuals while disregarding the needs and interests of others.

4. Protocolizing the governance of social media

4.1. Social media as public sphere in nature

Platforms, including search engines and social media, construct an essential part of citizens' online civic participation in the digital space, which can be classified as public goods due to their non-excludable and non-rivalrous nature. The tragedy of common goods seems unavoidable if we assume all the agents in the space are completely rational and self-interest-oriented - similar to the physical world, the only difference is that the resource here is attention rather than a tangible existence.

The Leviathan method practiced by the current sovereignty government aims to move the game result from Pareto-inferior to Pareto-optimal, assuming an omniscient state and complete control. Section 230 is an extension of such practice to stream power from authority to centralized platforms out of necessity (core algorithm and database ownership).

The problem with Section 230 is that the exertion of protection from responsibility does not necessarily turn a profit-mongering MNE into a provider of public goods. By default, the moderation power and the role of guardian of public interest should be owned by an entity resulting from a voting process and consensus. In authoritarian countries, it is often controlled by IT security, eliminating all speech against the authority.

Content moderation is closely related to the First Amendment, but it would be too naive for online citizens to trust any centralized entity serving as a safeguard. The fifth game design in Elinor Ostrom's article describes an outcome of automatic penalty terms embedded out of consensus before the game starts. Combined with the cybernetic governance surface and the decentralization-first principle, it offers a new way in decentralized web design (Ostrom, 1990).

4.2. Twitter v.s. Mastodon

Social media as the modern "agora" for citizens to gather should not be owned by any centralized entity but should be built as a public good infrastructure. This can be realized through

decentralized social graph protocols such as Nostr, Facaster, ActivityPub, etc., which guarantee that all content posted by users is recorded on-chain as unmodified data. Based on these protocols, decentralized storage systems such as IPFS and Filecoin enable data privacy and move forward towards individual sovereignty. Most importantly, the ownership of social graphs, indicated by the individual digital footprint (likes, following, comments, views), can be returned to the users themselves. For instance, Lens protocol enables users to mint NFTs (non-transferable tokens under ERC-721) as followers, Disco creates multi-dimensional social graphs by enabling bidirectional verification between nodes, and Mask creates profile graphs by organizing users' personal Twitter activity (Slepak, 2021; Wilkinson, 2021).

Twitter and Mastodon are both social media platforms that enable users to share short messages and engage in conversations with others. However, they differ significantly in terms of their underlying structures and philosophies. The most notable difference between the two platforms is their organizational structures, with Twitter being a centralized platform and Mastodon being part of the decentralized Fediverse.

The Fediverse is a federated network of independent servers that host various social media platforms, each with its own rules and moderation policies, and based on ActivityPub. This promotes autonomy, as each instance within the network is independently operated, allowing communities to develop their own rules and standards based on their unique needs and preferences. This is similar to the community-designed rule for better understanding and more comprehensiveness of commons than the centralized system in Ostrom's work (Ostrom, 1990). Users can join any instance (server) within the Mastodon network and interact with users from other instances, creating a more diverse and decentralized experience. This federated structure encourages local communities to develop their own rules and standards and allows users to enter or exit the server based on their personal preferences and experience. When entering the server, the user essentially agrees upon the rules set by the server, which is the consensus of the community. The reward or penalty policy is set to be automatically executed once triggered by certain behavior. Also, policies aimed at broader users may coordinate between different communities, build bridges of communication, and enable a larger impact of positive externalities.

4.3. Plural public: reimagining technology and democracy

The traditional centralized social media and engagement-based algorithms, aiming to extract messages that stimulate intense emotional responses among users, have had detrimental effects on human communication, as they can lead to misinterpretation and misrepresentation of information. Large foundation models, such as GPT, have the potential to exacerbate this issue by decontextualizing information, ultimately threatening information integrity. To combat these challenges and maintain trust in information, privacy-preserving tools are essential for upholding contextual integrity.

Plural publics represent digital spaces where individuals can collaborate and generate common knowledge while preserving context. Protocol-governed digital commons, with technological primitives such as identity certificates and distributed ledgers, can contribute to achieving higher information integrity. Additionally, deniable messaging techniques can aid in verifying the authenticity of information and maintaining its context. As a result, plural publics can foster inclusive conversations and interactions, ensuring diverse perspectives are represented in the digital realm. The rebuilding of social media as a protocol-governed public sphere has a significant impact on fostering healthy communication between different communities and repairing democracy in Western countries. Finding bridging topics between different communities

and identifying efforts that maximize combined welfare between various social groups, as well as facilitating the common understanding of social discourse, are the fundamental values when designing the protocol governing social media as digital commons.

5. Conclusion

In conclusion, the development of protocol-governed digital commons and decentralized social media platforms offers a new way to address the challenges posed by centralized platforms and their effects on human communication and democracy. By leveraging decentralized technologies and designing systems that uphold contextual integrity, it is possible to create more inclusive, diverse, and democratic digital spaces that better align with the values and needs of individuals and communities.

The implementation of such systems requires the collaboration of various stakeholders, including technologists, policymakers, and users, in order to develop protocols and infrastructures that are resilient, adaptable, and capable of fostering trust and collaboration. Furthermore, ongoing research and innovation in decentralized technologies, such as blockchain, distributed ledgers, and privacy-preserving tools, will be critical to realizing the full potential of digital commons and plural publics in the digital age.

By reimagining the role of technology in fostering democratic values and designing systems that prioritize individual sovereignty and collective action, we can move towards a more equitable and inclusive digital future. This will ultimately contribute to the creation of a more just and sustainable society, where diverse perspectives and voices are not only heard but also valued and respected.

6. Reference

Agrawal, A. (2001). Common property institutions and sustainable governance of resources. World Development, 29(10), 1649-1672.

Arendt, H. (1951). The Origins of Totalitarianism. Harcourt Brace Jovanovich.

Arrow, K. J. (1963). Uncertainty and the Welfare Economics of Medical Care. American Economic Review, 53(5), 941-973.

Ashby, W. R. (1956). An introduction to cybernetics. Chapman & Hall.

Buterin, V. (2014). A next-generation smart contract and decentralized application platform. White Paper, 3(37).

Coase, R. H. (1960). The Problem of Social Cost. Journal of Law and Economics, 3, 1-44.

Crawford, K. (2016). Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics. Science, Technology, & Human Values, 41(1), 77-92.

Dahlgren, P. (2005). The Internet, Public Spheres, and Political Communication: Dispersion and Deliberation. Political Communication, 22(2), 147-162.

Dale, R. (1999). The Taming of the Shrew: The Liberalization of Global Finance. Competition & Change, 4(2), 239-267.

De Filippi, P., & Hassan, S. (2016). Blockchain technology as a regulatory technology: From code is law to law is code. First Monday, 21(12).

De Filippi, P., & Loveluck, B. (2016). The invisible politics of Bitcoin: Governance crisis of a decentralized infrastructure. Internet Policy Review, 5(3).

Feder, G., & Feeny, D. (1991). Land Tenure and Property Rights: Theory and Implications for Development Policy. World Bank Economic Review, 5(1), 135-153.

Habermas, J. (1991). The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society. MIT Press.

Hardin, G. (1968). The tragedy of the commons. Science, 162(3859), 1243-1248.

Hirschman, A. O. (1970). Exit, voice, and loyalty: Responses to decline in firms, organizations, and states. Harvard University Press.

Lanier, J. (2013). Who Owns the Future? Simon & Schuster.

MEA (Millennium Ecosystem Assessment). (2005). Ecosystems and Human Well-being: Synthesis. Island Press.

Ostrom, E. (1990). Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press.

Ostrom, E. (2005). Understanding Institutional Diversity. Princeton University Press.

Ostrom, E. (2009). Beyond Markets and States: Polycentric Governance of Complex Economic Systems. American Economic Review, 100(3), 641-672.

Ostrom, E., & Walker, J. (1991). Communication in a Commons: Cooperation Without External Enforcement. In Laboratory Research in Political Economy (pp. 287-322). University of Michigan Press.

Ovadya, A. (2022) Bridging based ranking.

Posner, E. A., & Weyl, E. G. (2018). Radical Markets: Uprooting Capitalism and Democracy for a Just Society. Princeton University Press.

Rao (2013) The Unreasonable Sufficiency of Protocols

Sørensen, E., & Torfing, J. (2009). Making governance networks effective and democratic through metagovernance. Public Administration, 87(2), 234-258.

Swan, M. (2015). Blockchain: Blueprint for a new economy. O'Reilly Media, Inc.

Tufekci, Z. (2017). Twitter and Tear Gas: The Power and Fragility of Networked Protest. Yale University Press.

Wilkinson, R. (2021). Disco: Trust-minimized discovery for decentralized social graphs.

Winner, L. (1986). The Whale and the Reactor: A Search for Limits in an Age of High Technology. University of Chicago Press.

Worm, B. et al ((2009). Rebuilding global fisheries. Science, 325(5940), 578-585.

Zuboff, S. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. PublicAffairs.