PERSPECTIVE
published: 28 August 2019
doi: 10.3389/fbloc.2019.00009

The Ethics of Contentious Hard Forks in Blockchain Networks With Fixed Features

Tae Wan Kim[*†] *and Ariel Zetlin-Jones*[†]

The Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA, United States

An advantage of blockchain protocols is that a decentralized community of users may each update and maintain a public ledger without the need for a trusted third party. Such modifications introduce important economic and ethical considerations that we believe have not been considered among the community of blockchain developers. We clarify the problem and provide one implementable ethical framework that such developers could use to determine which aspects should be immutable and which should not.

Keywords: blockchain, cryptocurrency, ethereum, fixed features, ethics, governance

OPEN ACCESS

Edited by:
Natalie Geld,
MedNeuro, Inc., United States

Reviewed by:
Raul Zambrano,
Independent Researcher, New York, United States
Tim Cadman,
Griffith University, Australia

****Correspondence:***
Tae Wan Kim
twkim@andrew.cmu.edu

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *15 May 2019*
Accepted: *15 August 2019*
Published: *28 August 2019*

Citation:
Kim TW and Zetlin-Jones A (2019) The Ethics of Contentious Hard Forks in Blockchain Networks With Fixed Features. Front. Blockchain 2:9. doi: 10.3389/fbloc.2019.00009

1. INTRODUCTION

Blockchain protocols allow a decentralized community of individuals to maintain a ledger of transactions without the need for a trusted third party. However, the consensus algorithms that allow a community of users to maintain trust in a blockchain ledger, as developed by existing protocols, also allow the community to alter or change any of the initial design choices in the protocol. In this article, we argue for the importance of the following question: *Should any aspect of a blockchain protocol be immutable—either for economic, ethical, or other reasons?*

We illustrate the consequences of permitting any changes to a blockchain protocol by examining proposed changes to the Ethereum protocol that arose in April 2018. Other cases of Ethereum hard forks have occurred since 2015, such as Constantinople, and Byzantium. But for simplicity's sake, in this article, we focus on the 2018 case. At that time, Ethereum network developers proposed changes to the blockchain consensus algorithm that would disable newly developed, innovative mining hardware. We then propose an implementable, ethical framework for blockchain protocol designers to decide which aspects of their protocol are immutable and which are not.

2. A BLOCKCHAIN NETWORK WITH CHANGING FEATURES: THE CASE OF ETHEREUM AND ASIC MINING

2.1. Overview

In March 2018, rumors emerged that Bitmain, a large, cryptocurrency miner and designer of Application-Specific Integrated Circuit (ASIC) chips had developed a powerful dedicated hardware chip designed to mine Ether, the crypto-currency associated with the Ethereum protocol. Since this dedicated chip was anticipated to be able to mine Ether faster and more powerfully than existing hardware setups [largely dependent on Nvidia graphics processing units (GPUs)], these rumors, which were confirmed in April 2018, raised concerns about the possibility that a single market participant could control a plurality of mining power in the Ethereum network. With a plurality of mining power, a single market participant would threaten the security and validity of the Ethereum blockchain. In response, many Ethereum users (and GPU-dependent miners) proposed adopting a change in the Ethereum network protocol that would render the rumored ASIC miners obsolete.

In what follows, we describe in more detail the issues surrounding the debated change in the Ethereum protocol and the implications this might have for particular design choices underlying newly proposed cryptocurrencies.

2.2. Ethereum Mining

Ethereum's mining process is in many ways similar to Bitcoin's proof-of-work protocol. Miners listen for broadcasts of transactions that should be added to the Ethereum blockchain. Miners aggregate these broadcasts into a block of transactions and then compete to solve a complicated cryptographic puzzle. If the miner solves the cryptographic puzzle first, then the miner broadcasts the new block across the Ethereum network and is rewarded with newly minted Ether.

The primary difference between the Ethereum mining process and the Bitcoin mining process is the nature of the cryptographic puzzle that miners must solve. Specifically, the two protocols differ in their choice of a hash function. Ethereum relies on the EtHash algorithm, whereas Bitcoin relies on the double SHA-256 hash function. Bitcoin's hash function, SHA-256, is a cyptographic function that takes as an input a string of characters of any length and outputs a 256-bit hash. To mine a block of Bitcoin, miners append the block of data to be added along with the existing metadata from the blockchain and a conjectured nonce into a single string. They then input this string into SHA-256 and examine whether the resulting hash meets the existing Bitcoin hash target (i.e., the hash has a certain number of leading zeros). If so, they have successfully mined the block of Bitcoin; if not, they try another nonce and repeat the above process.

The EtHash process is more complicated and relies on a pseudorandom dataset which itself is initialized by the current Ethereum blockchain length. This pseudorandom dataset is called DAG and regenerates every 30,000 blocks. To mine a block of Ethereum, a miner once again creates a string out of the block of data to be added, the existing metadata from the blockchain, and a conjectured nonce. The miner then inputs this string into a SHA-256-like function to generate a 128 byte "mix," denoted as Mix 0. This mix is used to determine which page of the existing DAG must be retrieved. Mix 0 is then combined with the retrieved DAG page using a special mixing function to generate the next Mix. This process is repeated 64 times, yielding Mix 64. The miner then examines the hash of the resulting mix to see if it meets the existing Ethereum hash target. If so, then the miner has successfully mined the block of Ethereum; if not, the miner must repeat the process with a new nonce.

The critical difference (besides the added complexity) between these protocols is that the entire DAG must be maintained in the memory of the processor executing the mixes (since these are randomly retrieved). As of July 2018, the current DAG size was over 2.5 GB and will continue to grow in size. Because Ethereum mining requires many relatively simple computations but a significant amount of memory, miners have largely turned toward hardware builds that rely on many graphics processing units that access a central memory location.

2.3. ASIC Mining and Bitmain

Following the introduction of Bitcoin, mining began on Bitcoin's blockchain on individuals' CPUs. As interest in mining grew, miners moved to GPU processing power to more quickly mine blocks of Bitcoin, and this innovation was followed shortly by ASIC dedicated hardware chips. These ASIC chips quickly outperformed CPU or GPU-based miners and have now dominated Bitcoin mining for the past several years. As of July 2018, Bitmain-controlled mining pools (largely presumed to consist of ASIC miners) mined over 36% of all Bitcoin blocks (according to data compiled by https://coin.dance/blocks).

Ethereum's protocol—the EtHash specifically—was developed with the direct intent of trying to limit the power of ASIC mining hardware and ensure that general-purpose computers would retain a mining advantage[1]. In April 2018, Bitmain announced that it had successfully developed an ASIC mining chip capable of processing Ethereum transactions (and, implicitly, the EtHash)[2]. Bitmain's release specifications state that its Antminer E3 chip is capable of mining 180 MH/s, roughly 4–6 times more hashes than common GPU-based mining setups. While the Antminer chip requires roughly the same energy consumption as typical GPU-based setups (on a per Megahash per second basis), its speed allows it to more efficiently mine blocks of Ethereum than standard, GPU-based setups[3].

2.4. Ethereum Community Responses

Immediately after rumors of Bitmain's Ethereum ASIC mining chip emerged, prominent Ethereum protocol developers began asking whether the Ethereum protocol should be modified to render ASIC mining obsolete[4]. The mechanical way developers would limit the usefulness of ASIC mining chips would be by way of a "hard-fork" in the Ethereum blockchain.

It is useful here to describe the hard fork process in more detail. First, a "fork" in a blockchain is a regular or common occurrence. Forks occur any time two miners find a block at approximately the same time. Once this fork happens, blockchain network users (and miners) now have two copies of the same blockchain (that differ in their terminal block). However, these users treat a block as confirmed only once six additional blocks have been appended to the blockchain. This means that at the instant of the fork, the terminal block of each fork is not confirmed. Only when one fork of the blockchain becomes sufficiently long do (most of) the network users confirm the original block. Hence, these types of blockchain forks typically expire relatively quickly.

However, the blockchain protocol also admits "hard forks," where network developers introduce software upgrades to the protocol. Essentially every aspect of a specific blockchain protocol may be changed through a software upgrade. When a hard fork is introduced, any network users that continue to run the older version of the software will treat any new transactions (i.e., blocks) as invalid and unconfirmed. Hence, for a hard fork software upgrade to be successful, a large enough segment of

[1] Not surprisingly, part of the reason for this intent was to ensure that mining power would not become concentrated in the hands of a few miners.

[2] See Bitmain's Twitter announcement: https://twitter.com/BITMAINtech/status/981169456411021312.

[3] See https://cryptoslate.com/bitmain-e3-asic-ethereum-miner/ for some of this data.

[4] See https://github.com/ethereum/EIPs/issues/958 or https://twitter.com/VladZamfir/status/979060233430552576 for examples.

network users must agree to adopt the new blockchain fork. If all users agree to use only the blockchain resulting from the hard fork, then the forked blockchain network replaces (effectively) the old one. If instead many users agree to use both the original blockchain and the hard forked blockchain, then the fork essentially creates a new blockchain network.

One of the most prominent such hard forks occurred on August 1, 2017 when the Bitcoin blockchain experienced a hard fork. This hard fork was introduced as a way for Bitcoin network developers to increase the blocksize of each block in the Bitcoin blockchain. This increase in blocksize would allow for more transactions to be written into a single block of the blockchain. The implementation of the hard fork yielded a new blockchain (associated with Bitcoin Cash) in addition to the original Bitcoin blockchain. Each holder of Bitcoin at the time of the hard fork then owned an equal amount of Bitcoin and Bitcoin Cash. Even though this hard fork was intended as a software upgrade to replace the existing Bitcoin blockchain, because users adopted and continue to update both blockchain networks, the hard fork resulted in an effective doubling of the supply of (original) Bitcoin (from 21 to 42 billion coins). Of course, over time as relative demand for tokens on the two networks changed, the relative price of Bitcoin and Bitcoin Cash changed as well.

Returning to Ethereum, developers in April 2018 were proposing an upgrade to Ethereum's network software. Specifically, there were calls to modify the EtHash function so that the newly developed ASIC miners, with chips dedicated to implementing the existing EtHash function, would be rendered useless. To adopt this change, developers proposed implementing the software upgrade by way of a hard fork to replace the existing blockchain. This hard fork has not been implemented to date, largely due to the belief that future planned changes to the Ethereum network protocol would have the same effect. Such a hard fork would have protected the efficiency of existing GPU-based mining setups while devaluing the usefulness of newly developed ASIC miners, which, as described above, are capable of mining a block of Ethereum in a faster amount of time. Such changes by their very nature involve a redistribution of costs and benefits across miners and, potentially, network users. Many proposed this hard fork not purely to protect their (sunk) investments in GPU-mining setups but rather to prevent a concentration of mining power in the hands of Bitmain—a significant threat to the trust needed to maintain a decentralized ledger.

What we hope is clear from the above discussion is that hard forks, in principle, allow network users and developers to modify almost any aspect of a blockchain protocol. Narrowly, these changes have been limited to software upgrades intended to improve the performance of the network, such as changes in the hashing algorithms used to verify blocks in the blockchain. However, as illustrated above, some proposed changes may involve a redistribution of value across network users. More generally, such "upgrades" could be used to change the pre-determined supply of cryptocurrency. Or, in principle, it would appear that users could agree to remove all cryptocurrency held in a specific wallet and re-allocate it to other users in the network. Again, these are feasible changes that could be implemented by way of a hard fork. The only limitation to these hard forks is whether there would be a consensus among network users to adopt such changes.

3. A NORMATIVE FRAMEWORK FOR BLOCKCHAIN DESIGN WITH FIXED FEATURES

Which features of a blockchain protocol should or should not be alterable? To answer this question, we need a normative framework[5]. Our framework is twofold: the substantive and the procedural. The substantive consists of two ethical principles: The generalization principle and the utility-enhancement principle. The procedural has three principles: publicity, revision and appeals, and regulation. All the principles are necessary conditions. The procedural principles help to collectively examine whether any application of the two substantive principles are reasonable. The set of the five principles as a whole is in line with the broadly Kantian deontological approach to justice and democracy (Kant, 1785). In particular, we are partly indebted to Daniels and Sabin (2002) procedural approach to fair allocations of limited resources. Yet, our framework is different from theirs in several ways: the particular context we deal with is different, we replace the controversial "relevance" condition with our own representation of the Kantian generalization principle, and we add the utility-maximization principle. Although we do not offer a fully fledged normative analysis of the given issue, we propose a *possible* normative framework for cryptocurrency communities.

First, we submit that whether or not it is permissible to change some aspect of the protocol should not be a matter of power. Instead, this normative judgment should be a matter of reasonableness/legitimacy. In our framework below, we define reasonableness/legitimacy as the maximization of both fairness and economic utility.

In our framework, a change (or class of changes) to the protocol should be permissible, if all of the following five requirements are met at the same time:

I. Substantive principles

1. **Utility-Enhancement Principle:** The change in the protocol maximizes (or least enhances) the overall value of the cryptocurrency in question[6].

[5] Normative statements are stipulative and action-guiding, whereas descriptive statements are reportive of states of affairs. Examples of the latter are "The grass is green" and "Many people find deception to be unethical." Examples of the former are "You ought not murder" and "Lying is unethical." Normative statements usually figure in the semantics of deontic (obligation-based) or evaluative expressions such as "ought," "impermissible," "wrong," "good," "bad," or "unethical" (Kim and Donaldson, 2018). Next, in this article, we use the word "normative" as a synonym for the expression "morally normative" and not in the weaker sense adopted by many social scientists, that is, meaning merely practical guidance. In particular, "normative" refers to concepts that are prescriptive in a categorical rather than hypothetical way. The proposition "If you want good reputation, then never lie" is prescriptive, but only in a hypothetical sense. In contrast, the proposition "You ought not to lie" is intended as an all-things-considered, non-hypothetical guide to action. Such an imperative qualifies as "normative" (Kim et al., 2018).

[6] Unlike in utilitarianism, the utility principle in our proposal is a deontic principle (see Hooker and Kim, 2018).

2. **Generalization/fairness Principle:** The change is something that everyone is allowed to do in similar circumstances, if they can (Nagel, 1986; Korsgaard, 1996; O'Neill, 2014).

II. Procedural principles

3. **Publicity Principle:** Decisions and rationales behind the decisions must be publicly accessible to all coin users.
4. **Revision and Appeals Principle:** There must be mechanisms for challenge and dispute resolution regarding decisions, and, more broadly, opportunities for revision and improvement of policies in light of new evidence or arguments.
5. **Regulation Principle:** There is either voluntary or public regulation of the process to ensure that conditions 1–4 are met.

Utility-Maximization is straightforward, although applying the condition needs to be further elaborated once a specific real context is given. In typical circumstances, utility would refer to the financial value of a coin. Any change must enhance the expected monetary value of the coin. But it is not always an easy task to forecast the expected value. The best evidence and theoretical tools must be used.

One might ask: "Is the utility principle really necessary? Network participants have an economic incentive to maximize the value of the coin. Moreover, making wealth-maximization for token-holders the ethical goal could cut against, say, creating a public commons, or financial inclusion, or some other values." This is an apt question. But the question comes from a misunderstanding of the term "utility." Utility *per se* is selfless. Utility is an intrinsically valuable state of affairs (e.g., welfare) of all involved bodies, so it is a good proxy for the common good. Second, the utility principle is a necessary condition, so any proposal that meets the utility principle but not the fairness/generalization condition cannot pass the substantive test (Hooker and Williams, 2012; McElfresh and Dickerson, 2017).

The Generalization Condition is a well-established ethical test to secure fairness. The test is useful to avoid cheaters/free-riders and to promote sustainability. If necessary, the generalization condition can be formalized. Hooker and Kim (2018) formalized the principle with quantified modal logic, to teach the principle to artificial intelligence.

We clarify the application of the Generalization Condition using a hypothetical situation. Suppose that a single coin user amasses a large share of the coin. In response, all other users of the coin could hard fork the protocol to re-distribute the single user's coins evenly to all other network users. In principle, such a hard fork is feasible (on existing cryptocurrency networks). Should this be permitted on the coin network?

The Generalization Condition requires that the user's action is generalizable. A premise underlying the condition—a so-called, "universality of reason"—is that if a reason justifies an action for me, it justifies the same action for anyone to whom the reasons apply. This corollary is a functional expression of a normative value, fairness, in ethics and, simultaneously, an epistemic value, consistency in logic and mathematics.

Consider the scenario: Multiple users of the coin intend to hard fork to gain financially. Their proposal is "Hard forking is okay when the reason for it is financial gain and all other users support the proposal." To be generalizable, the change must be something that every similar size group is allowed to do. However, a simple game theoretic thought experiment implies that it is irrational to believe that the change is something that every similar size group is allowed to do. For the sake of argument, suppose that every such group were allowed to make such a change to the protocol. In that case, users would be concerned that if a coin user were to accrue any sizable account of coins, then remaining users would propose a change to redistribute the single user's coins. In the end, no user would want to accrue coins, and the coins would not have value—all of which yields the conclusion that users cannot gain from the ability to make such changes.

If it is permissible for anyone to engage in a practice (hard forking to redistribute), no one can take financial advantage of that practice. As a result, the practice in question is said to be *ungeneralizable*. That is, the practice in question is not a principled action. Ungeneralizable acts are forms of free-riding. They are also unsustainable practices.

Now consider a real case: In the case of the DAO exploit in 2016, one user stole $50 million from a decentralized application. The miners eventually agreed to fork Ethereum to return the funds. The proposal was "Hard forking is okay when the reason for it is to remedy financial harm caused by a crime and all (or most of) the miners support the proposal." Now imagine that the proposal were endorsed. What would happen? The hard forking would work without any contradiction or self-defeating results. Thus, the miners' hard forking passed the generalization condition in 2016.

How about the Ethereum anti-ASIC proposal? The proposal is "It is okay to hard fork if the reason for it is to prevent a single miner from threatening the security and validity of the Ethereum blockchain and the majority of the participants support this proposal." Now imagine that this proposal were really endorsed. What would happen? Whenever it is reasonably expected that a single miner's power is likely to threaten the security and validity and most of the participants are afraid of that, the participants would hard fork to control the single user. Thus, such a proposal would accomplish its own goal without any self-defeating results. So, such a proposal passes the generalization principle.

Now consider the case of GPU miners: Developers have incentives to develop and implement new technologies that mine crypto-currencies faster and with less energy dependence. The adoption of these new technologies, however, poses a threat to the profitability of mining operations that implement older technological vintages. More broadly, how does our normative framework deal with technological innovations? We show how the principles in the framework can provide an approach to deal with innovations.

Suppose that a miner (you) wants to use an innovative tool X (a brand new GPU) to increase your financial gain. The action is "Using X." The reasons are (i) "I can use X" and (ii) "to increase my financial gain." Now, apply the universality of reason (i.e., the first step of the generalization principle). Then, "Anyone who can use X and wants to financially gain is okay to use X." This step is to transform "a rule that an individual uses" to a principle that governs everyone. Now do the second step. Imagine what would happen if the principle were adopted? Third, ask, "Would it be still rational to believe that you could financially gain by using X in the imagined world?" If the answer is negative, your using X in the real world is unethical. But we cannot answer the question, because without more details it is difficult to predict how many miners could afford and use X once the principle were adopted. If you are the only one who can use X, then you will gain. If many can use X, there is no advantage to use X, so the action plan is self-defeating and un-generalizable. This issue is in part an empirical matter. The base line is this: fairness requires that if you are okay to use X, everyone must be okay to use X.

One might say that when only some people can afford X, while others not, allowing X is unfair. But this is not a directly relevant issue in fairness/generalization. Using X in our context is an investment choice.

Now, let us use the utility principle. If you are the only person who can use X, will your using X increase the total amount of utility of all involved parties including people outside the community? If yes, it passes the test. The utility principle requires us to see the social value aspects. Consider, for example, a technological innovation in mining. If this innovation, if used by a single miner, allows the mining network to validate the same number transactions as under the previous technology but at a lower energy cost, then total utility of all parties should rise and the innovation would pass the utility principle. If however this innovation is a way to simply redistribute the financial gains from current mining without reducing the energy costs—as in many arms races—then the innovation would fail the utility test.

The Publicity Condition is useful to ensure the transparency of decision-making procedures, which allows involved parties to examine whether these procedures are coherent, sound, and grounded by evidence. Having the rationales publicly accessible can also make proposers clarify and double-check their rationales and relate them to involved stakeholders. Such transparency would ensure that decisions are open to scrutiny and debate by the people, which, in turn, can contribute to the quality of public deliberation and facilitate social learning. Decision-making processes that meet the Publicity Condition demonstrate that the decision-making is principled and responsive to the people, in particular to those who are affected by the decisions, thereby providing legitimacy to the decision-makers.

Suppose that a group of coin users announces that its proposed change meets the Utility-Maximization and Generalization Conditions, and the group discloses its reasoning processes for the two conditions, which meets the Publicity Condition. Imagine that a different group believes that the proposed change does not meet the Utility-Maximization condition or the Generalization Condition. Then, through the Revision and Appeals Condition, the different group must be allowed to offer an explanation why it believes that the original group's applications are not correct. If the original group finds that the challenge is wrong, it can explain why in light of new evidence or arguments. If the original group finds that part of the challenge makes sense, it could revise the original proposal and resubmit. Throughout revisions and appeals, proposals for changes can gain more legitimacy.

To ensure the four conditions above, there must be enforcement. Enforcement can be voluntary if the Regulation Condition is designed to focus on financial incentives. The Utility-Maximization Condition and the Generalization Condition both aim to sustainably enhance the financial value of a coin. But some specifically designed incentive or penalty must be added to regulate the behavior of coin users. An alternative is to outsource the Regulation Condition to a governmental body that has policing and legal resources. Much more discussion is needed.

A third possible approach is to program the ethical framework into the blockchain protocols, so that if someone proposes a change that is not permitted, ethically, the protocol is designed to erase itself (Deuber et al., 2019). Part of the ethical framework introduced in this paper has been already represented as a form of quantified modal logic in AI ethics to teach ethics to AI (Hooker and Kim, 2018) and Natural Language Processing (Prabhumoye et al., 2019). More research is needed to find out which programming technique would best serve the context of blockchain.

4. CONCLUSION

An advantage of blockchain protocols is that members of a decentralized community of users may each update and maintain a public ledger without the need for a trusted third party. However, the ability of each individual to update and maintain the ledger also creates the ability to update and modify the protocol itself. Such modifications introduce important economic and ethical considerations, which we believe have not been considered among the community of blockchain developers. Since it is technologically feasible to hard-code certain aspects of a protocol, we believe it is important for protocol developers to consider whether any aspects of their protocol should be immutable by network users. We have provided one implementable framework that such developers could use to determine which aspects should be immutable and which should not.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We are indebted to Kevin Werbach, two reviewers, and the associate editor Natalie Geld.

REFERENCES

Daniels, N., and Sabin, J. (2002). *Setting Limits Fairly: Can We Learn to Share Medical Resources?* New York, NY: Oxford University Press.

Deuber, D., Magri, B., and Thyagarajan, S. A. K. (2019). "Redactable blockchain in the permissionless setting," in *2019 IEEE Symposium on Security and Privacy (SP)* (San Fransisco, CA), 645–659. doi: 10.1109/SP.2019.00039

Hooker, J. N., and Kim, T. W. N. (2018). "Toward non-intuition-based machine and artificial intelligence ethics: a deontological approach based on modal logic," in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18)* (New York, NY: ACM), 130–136. doi: 10.1145/3278721.3278753

Hooker, J. N., and Williams, H. P. (2012). Combining equity and utilitarianism in a mathematical programming model. *Manag. Sci.* 58, 1682–1693. doi: 10.1287/mnsc.1120.1515

Kant, I. (1785). *Groundwork of the Metaphysics of Morals*, *Vol. 4*. Akademie edition.

Kim, T. W., and Donaldson, T. (2018). Rethinking right: moral epistemology in management research. *J. Bus. Ethics* 148, 5–20. doi: 10.1007/s10551-015-3009-2

Kim, T. W., Donaldson, T., and Hooker, J. (2018). Mimetic vs anchored value alignment in artificial intelligence. *arXiv [preprint]*. arXiv:1810.11116.

Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge, UK: Cambridge University Press.

McElfresh, D. C., and Dickerson, J. P. (2017). Balancing lexicographic fairness and a utilitarian objective with application to kidney exchange. *arXiv [preprint]*. arXiv:1702.08286.

Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.

O'Neill, O. (2014). *Acting on Principle: An Essay on Kantian Ethics, 2nd Edn.* Cambridge, UK: Cambridge University Press.

Prabhumoye, S., Mayfield, E., and Black, A. W. (2019). Principled frameworks for evaluating ethics in nlp systems. *arXiv [preprint]*. arXiv:1906.06425.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.