Utonoma: A Model for Incentivizing Productive Collaboration

Adrian Armenta Sequeira
adrian.sequeira@utonoma.com

Gerardo Pedrizco Vela
gerrypvela@utonoma.com

March 8, 2024

Keywords: utonoma, incentive model, collaboration, production, blockchain, smart contracts, cryptocurrencies

Abstract

For communities to progress, improve, and solve their problems more quickly without directly depending on a profitable business model, it is necessary to implement an efficient incentive system that promotes productive collaboration. This is because, to date, such systems depend on the good faith, altruism of participants, and centralized committees that choose which actions are rewarded and which are not. However, there are few success stories, as most run out of resources or allow corruption to creep in.

Therefore, this article proposes the idea of a system based on blockchain and self-managed by the community itself, which is decentralized and resistant to malpractices, such as collusion or corruption. A model where, in addition to good faith and altruism, social recognition and the motivation to grow economically are added as incentives.

This is achievable through *Utonoma*, a social network aimed at users posting content and documentary evidence of their actions with the aim of being rewarded. Users are responsible for voting on the blockchain to determine if they are deserving of the reward and to what extent. The prize consists of cryptocurrencies that are generated and awarded to the author.

Introduction

To begin with, it's necessary to present a scenario where an individual sees an opportunity for improvement that they are capable of addressing with their own resources. For example: repairing a sidewalk, painting a wall, or placing a warning sign. Undertaking this task brings benefits to themselves and those around them. However, even though the action may be urgent for themselves and beneficial for society, it's rare that they will take it on, as the materials needed would have to be provided by themselves. Additionally, their act would benefit not only the good members of society but also the undesirable ones, potentially leading to a feeling that the latter are taking advantage

of their effort. For this reason, the person will likely choose to focus on something more rewarding, like their job or an activity that generates money.

As a result, they will leave such activities and improvements to be carried out by the government or organizations sponsored by it. However, bureaucratic processes are very slow and can take years to implement the necessary actions to benefit the community. Wouldn't societal improvements be resolved more efficiently and effectively if each community member directly contributed according to their capabilities?

That is precisely what *Utonoma* aims to achieve. Its goal is to accomplish this through the creation of an incentive system for productive collaboration that persuades people to do things they consider good, despite the obstacles they may face in carrying them out. In doing so, it would be encouraging them to take the first step towards solving something within their capabilities. Additionally, it addresses the need for independence from a centralized committee, as the perception of value varies from one person to another, depending on their preferences and needs.

1 Determining Fair Reward

> *"I am fascinated by Tim May's crypto-anarchy. Unlike the communities traditionally associated with the word "anarchy", in a crypto-anarchy the government is not temporarily destroyed but permanently forbidden and permanently unnecessary"*
>
> Cypherpunk Mailing List, 1998, Wei Dai

As mentioned earlier, Utonoma consists of a social network where users can upload content to add value to the network. The rest of the participants are responsible for judging whether the actions published deserve to be rewarded. They do this through voting, which aims to reach a consensus on valuable acts and the economic decisions that must be made to reward their authors. Once an agreement is reached that the content indeed constitutes something of value, the creator will be rewarded with an amount of cryptocurrency proportional to the voting results.

It is important to mention that this entire process is managed using smart contracts [1] and the blockchain, in a decentralized manner. As a result, the mentioned cryptocurrencies are native to the system and are generated at the time of awarding the reward. Initially, they have no intrinsic value. Therefore, the system will only work if the rest of the network users appreciate and recognize the value generated within it by buying the network's cryptocurrencies. This appreciates the reward obtained by the content creator, as it becomes scarce and valuable. It also benefits them as they can keep them to foster their own wealth, which would generate a self-perpetuating and reinforcing economic incentive system, as the increase in the valuation of the reward

[1] A computer program whose agreements are processed in the same way as transactions in the Bitcoin blockchain system.

incentivizes authors to create more and higher-quality new content.

Therefore, Utonoma creates a cooperative environment where the well-being of each party depends on the general well-being. Thus, if the cryptocurrency begins to be hoarded, then content authors will be incentivized to create more, and consumers will enjoy better quality content. On the other hand, if users do not correctly moderate content through their votes, those holding the cryptocurrencies will end up getting rid of them because who would want to make rich the content creators of a social network where only garbage is generated? Therefore, the network will only work if each of the parties properly performs its job and it will do so because the greatest benefits for each one are found in correct behavior, rather than incorrect. Without the need for the creation of a strict regulation or a trusted third party for its application, the network will find its point of balance on its own.

2 Content Hosting to Combat Censorship

Currently, social networks are a means to disseminate entertainment, educational content, and opinions. Although freedom of expression is often assumed to be inherent in their use, the reality is that a large number of people feel they have been censored on the internet, as pointed out by Emily A. Vogels (2020). In 2022 alone, an estimated 4.2 billion individuals experienced this situation (Surfshark, 2022). Therefore, the following question arises: *If the content of Utonoma were hosted on a centrally managed server, could pressure be exerted through bribes or threats to have it removed?*

To avoid such censorship, the best way to host its content so that its dissemination cannot be stopped is through an IPFS (Interplanetary File System), also known as an Interplanetary File System, which consists of a user-to-user (Peer-to-peer) file protocol, similar to *BitTorrent*, that is addressable by its content and not by its location.

It is important to remember that a *hash* can be obtained from each content, meaning a mathematical algorithm can be applied to transform data strings of any length into a new fixed-length data string; the result is unique, as there are no two input data strings that can produce the same *hash*. Thus, this system consists of files that in turn reference others successively, creating sets of *hypermedia* that can form complete web pages.

Therefore, users can ensure they are not receiving an apocryphal copy of the file in question, so it is considered that the user-to-user method allows a decentralized distribution of the system's files; to deny access to them, it would be necessary to eliminate all nodes that possess them.

Using an IPFS network, the hosting of *Utonoma's* content would be as follows:

1. Users upload content to the IPFS network and obtain a *hash* with which they can later retrieve the file.
2. They write the *hash*, obtained in the previous step, in the *Utonoma* smart contract.
3. The application that functions as a client accesses the smart contract and retrieves the file *hash*.
4. Subsequently, the *hash* is used to obtain the file, through the IPFS network, and it is displayed to the user.

It is important to mention that the client application can avoid showing explicit content that infringes copyright, in addition to establishing its own moderation rules. Thus, if

a user believes that a client application is exercising censorship, then they can switch to another; any will work using more or less the same structure, as they will be feeding on the information from the *Utonoma* smart contract, where no one can delete the uploaded content, as mentioned earlier, it will be hosted in a system that combines *blockchain* and IPFS, and any user can upload IPFS *hashes* to the smart contract, which will function as a database.

3 Fine Content Moderation

> *"People who do not usually express their emotions are unhappier and feel more isolated. Moreover, apparently suppressing the expression of these emotions does not reduce, and can even increase, the experience of negative emotions, such as disgust, anxiety, sadness, and shame."*
>
> The Economy of Caresses, Alex Rovira

Research conducted at Stanford University by James Gross (2002) concluded that suppressing the expression of emotions carries high psychological, social, and health costs. Therefore, individuals who usually suppress the expression of their feelings generally experience more negative than positive experiences. Furthermore, the lack of expression of feelings generates greater psychological stress, both in the person who suppresses them and in the person they interact with. Additionally, the suppression of emotional expression is associated with a decrease in physiological immunity.

For the above, the competition for audiences on the internet is so arduous that creating quality content is not enough, as it becomes necessary to "ignite the networks" for it to be consumed. That's where recommendation algorithms do their job, as they more frequently show controversial content since it generates more reactions and keeps users on the social network for longer.

This model makes content creators willing to do anything to generate controversy. It also encourages the use of hate speech and makes outrage marketing campaigns effective, deliberately seeking to offend in order to generate controversy with the aim of increasing their reach and reaching their target audience.

Consequently, although client applications can filter content in a broad sense, fine-tuning is much more difficult to achieve. Normally, social network companies use real employees, and their job demands that in a few seconds they decide whether content is just a bad-taste joke, hate speech, fake news, or a debate of a reasonable doubt. However, these moderators will inevitably apply bias in their choices. As a result, the final outcome will not be something natural, but rather crude, being too strict on some occasions and too permissive on others.

Returning to *Utonoma*, a more dynamic way of moderating content would be one in which users can also vote against content, with a predominantly negative vote being an indication that the community disapproves of it, considering it to be valueless and,

therefore, should be removed from the smart contract so that it can no longer be listed by client applications. When this happens, the author will be penalized with the aim of making content creators think twice about the medium they use to convey their message, as using offense can lead to unfavorable voting and the removal of content without any reward for the author.

The purpose of this is to promote the use of more appropriate language that encourages dialogue and avoids the creation of ideological bubbles, maintaining a pluralistic spirit and respecting the right to freedom of dissent.

This idea of content moderation based on the morality of each individual stems from the belief that, as long as the majority of members remain honest, individuals in societies can mutually moderate their behavior.

4 Measuring Coin Issuance

"Nothing is more useful than water: but it will purchase scarcely anything; scarcely anything can be had in exchange for it. A diamond, on the contrary, has scarcely any use-value; but a very great quantity of other goods may frequently be had in exchange for it."

The Wealth of Nations,
Adam Smith

The mostly favorable voting of content should generate an amount of cryptocurrency proportional to the result of the difference between positive and negative votes, to be given to the creator as a reward.

This is represented by the following formula:

$$V_u = V_p - V_n \qquad \text{Useful Votes} \tag{1}$$

Donde:
$V_p \geq V_n$
V_p = *Positive votes*
V_n= *Negative votes*

Consequently, this awarding implies the correct circulation of new coins. However, if this increase is excessive, it can generate inflationary pressures. To delve into this point, imagine a situation where the network has 1,000 users, and an author manages to create content for which 1% of the network, i.e., 10 users, voted favorably. If the reward were one coin for each useful vote and assuming no unfavorable votes, 10 coins would be generated and given to the author.

Now let's consider a scenario with the same conditions, but where the network had 100,000 users. The 1% would be 1,000 coins as a reward, which would cause a significant devaluation of the pre-existing coins and an unfair situation for the first user who only obtained a few coins for their work.

In other words, **if the reward for each useful vote obtained is constant and if the network exponentially increases its use, the monetary mass will also do so. Consequently, it will be more common to obtain a favorable vote when there are many users in the network than when there are not. This issuance of new coins in an uncontrolled manner will end up undermining the wealth of the system, as the coins would no longer be scarce. Similar to what would happen with diamonds if they were very common.**

Therefore, to control the monetary supply and preserve the existing wealth, while continuing to issue new coins to reward content creators, it is proposed that the use of the network be associated with the number of coins obtained for each useful vote. So that, the more the network is used, the fewer are issued.

According to Robert Metcalfe's law, the value of a telecommunications network increases proportionally to the square of the number of system users (n^2). Therefore, to obtain the theoretical value of a single coin in the Utonoma network, the previous equation (which is equivalent to the total value of the network) should be divided by the total cryptocurrency supply:

$$P = \frac{n^2}{S_T} \qquad \text{Price} \tag{2}$$

Where:
n = *Number of users*
S_T = *Total supply*

Then, the reward for obtaining a positive vote should be multiplied by a factor that causes the number of new coins issued to decrease as more value is in the network. Thus, from equation 1 and Robert Metcalfe's law, the reward is:

$$R = B(\frac{V_p - V_n}{n^2}) \qquad \text{Reward} \tag{3}$$

Where:
$V_p \geq V_n$
B = *Base reward (an arbitrary constant that marks the amount of reward obtained when there is only one user in the network, if it were possible to claim a reward under these conditions).*

From equation 3 it can be seen that when there are few users in the network, the reward for each useful vote is greater. This is necessary because it seeks to motivate new users with more coins to join and contribute to the network early. Later, when there are more users, the reward for each useful vote will be smaller, preventing inflation from being so strong.

5 Preventing Fraud and Bots

Earning a reward for favorable voting is an excellent way to associate the value of the network with the value perceived by the Utonoma user community. However, it can also be an incentive for vote rigging through collusion, as many users might mutually vote positively for each other's content without making a real value contribution to the network. Therefore, it is necessary to add a fee to the voting cost, so that if a user votes infinitely for their own content, they cannot earn more than they spend.

To determine the price of the coins awarded as a reward, the reward can be multiplied by the price of the cryptocurrency:

$$P_R = R \bullet P \tag{4}$$

Substituting R and P and developing:

$$P_R = B(\frac{V_p - V_n}{n^2}) \bullet \frac{n^2}{S_T}$$

$$P_R = B(\frac{V_p - V_n}{S_T})$$

To simplify calculations, let's consider the case of a single positive vote and none negative, then the reward price would be:

$$P_R = \frac{B}{S_T} \qquad \text{Reward price} \tag{5}$$

The above equation represents the price of the reward earned per positive vote received. To make defrauding the system unprofitable, the price for voting should be made higher than the reward price, by multiplying the former by a commission constant:

$$P_V = C\frac{B}{S_T} \qquad \text{Price for voting} \tag{6}$$

Thus, to vote, assuming that the cost of the blockchain commission for executing the voting function in the smart contract is negligible, cryptocurrencies equivalent to the voting price should be added to the transaction, so that defrauding the system is not profitable.

Conversely, defrauding through negative voting is also possible, as a user could be censored if a computer bot repeatedly votes against their content, which would result in its removal. Therefore, **this voting price should also apply in the case of negative voting.**

6 Additional Services

Adding a fee for casting a vote makes fraud and manipulation economically unviable. However, a very high voting price might discourage users from participating in moderating and rewarding content. Therefore, a voting encouragement program is proposed. This is because the alternative would be to have no fees, and to achieve this, computer bots could be detected with a centralized validation system (like captcha), resulting in reliance on a trusted third party and reverting to opaque and manipulable models similar to those currently in existence.

Hence, this necessity can be turned into an opportunity for client applications to offer exclusive services that enhance the user experience of frequent voters. Each can devise a different incentive model and compete to offer more and better services. For example, a client application could implement an advertising system and, from the earnings obtained from this, offer a service for music streaming to users who have cast a certain number of votes in the last month.

On the other hand, in equation 6, it can be seen that more cryptocurrencies are entering than leaving when a vote is cast, due to the commission constant. These resources can be used to reward the nodes of the IPFS network that host the content, thus making the network more resilient, ensuring that no file in the network is lost because there are no nodes available to host it, either because they have disconnected from the network or because they have deleted the file to host a new one.

7 Content Removal Algorithm

Although the system of content moderation through morality is fairer and more transparent than allowing a single entity to decide what content should be removed, it is not perfect and can still be subject to abuse. What if, for example, content that over time has a 90% favorable vote is uploaded to Utonoma and immediately receives a negative vote? In this case, 100% of the vote would be unfavorable, but this should not imply its removal, as simply not enough votes have been cast.

The law of large numbers states that as the number of experiments increases, the arithmetic mean of those samples will approach the expected or population mean. In other words, the larger the number of votes, the more certainty there is about what to do with the content: whether to reward it or remove it.

However, if content is harmful to the Utonoma network, it should be removed as quickly and cheaply as possible. Therefore, a minimum number of votes is sought that gives enough certainty about what should be done with it. While it is not feasible to wait for all users to vote, because theoretically the vote has no end, a sample could be obtained at any time from the current vote and define a confidence interval in which the proportion of the population is contained with a margin of error.

If removing content requires at least half of the votes to be unfavorable, then the formula for the confidence interval for the proportion could be used to evaluate whether the lower limit is greater than this amount, assuming that we are working with a standard normal distribution and the current vote is a simple random sample of the total vote:

$$(p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}, p + Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}})$$

Where:
p = *proportion of the sample of negative votes.*
n = *sample size*
$Z_{\alpha/2}$ = *critical value corresponding to the confidence degree 1-*α *of the standardized normal distribution.*
α = *significance level for calculating the interval*

As we are only concerned with the lower limit of the interval being greater than $\frac{1}{2}$, meaning that the proportion of the sample of negative votes is contained in an interval whose lower limit is at least above $\frac{1}{2}$, and as the upper limit will be greater than the lower, then the latter is irrelevant and can be omitted, leaving only the following part:

$$p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \tag{7}$$

If the lower limit of the confidence interval is above $\frac{1}{2}$, it allows us to affirm that with a certain level of confidence, the number of negative votes is at least $\frac{1}{2}$ of the total vote. Conversely, if the lower limit is below $\frac{1}{2}$, the cases in which content should be removed and in which it should not overlap, and the proportion could be above or below $\frac{1}{2}$, meaning there is a probability that the number of negative votes for the content is less than $\frac{1}{2}$, therefore, the content cannot be removed, and more votes are needed to deliberate.

To remove content, the following condition must be met:

$$p - Z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} > \frac{1}{2} \qquad \text{Condition for content removal} \tag{8}$$

If the above inequality is met, the content should be removed; if not, the author of the content in question can claim their reward, as long as the positive votes outnumber the negative ones.

8 Conclusions

To maintain user attention and generate higher profits, social media companies tend to favor controversial content. With the use of deep fakes and artificial intelligence for content generation, it is now impossible to discern what is true from what is not.

The only way users could get closer to the truth is by listening to both thesis and antithesis so that each one can formulate their own opinion. However, they will not do so if they remain sheltered in their ideological bubbles that protect them from insults, which are a product of the business model of generating controversy. This immerses humanity in a kind of new dark age often referred to as the era of post-truth.

It is paradoxical, as now there is access to a vast amount of information, but users will not access the antithesis of their ideas if it means exposing themselves to dehumanizing language and insults. Therefore, the only way to break the cycle is by finding a way in which promotion through offense economically harms those who use and promote it, instead of favoring them.

Utonoma is a social network that seeks to achieve this by allowing users to vote for or against content and incentivizing authors to contribute positively to society, as the rewards they receive are based on community appreciation of their contributions.

If an author decides to use hate speech, the aggrieved parties will vote negatively, having an economic repercussion, as the awards are calculated using favorable votes and deducting unfavorable ones. In addition, to keep the network free of bots, users must attach a fee to each vote, but this is compensated for by the "premium" benefits the social network can offer to users who vote consistently. It also addresses content moderation and censorship concerns in a democratic and decentralized manner, managing voting, management, and content hosting using blockchain, smart contracts, and the IPFS network.

As these are open-access, other third parties can create their own social networks giving their users access to all the functionalities the smart contract offers (reward generation, voting, penalties, etc.) In addition, Utonoma has an issuance method that adjusts the number of new cryptocurrencies produced based on the level of network activity, thus ensuring sufficient supply while caring for inflation and value preservation. With this, it seeks to promote a more inclusive, fair digital ecosystem that enhances the spirit of the collaborative economy.

References

[1] Gross, J. (2022) *Emotion regulation: Affective, cognitive, and social consequences.* Psychophysiology 39. Cambridge University Press. USA. Retrieved on January 28, 2024, from: https://onlinelibrary.wiley.com/doi/pdf/10.1017/S0048577201393198

[2] Surfshark (2022). *4.2 billion people experienced internet censorship in 2022.* Surfshark. Retrieved on January 28, 2024, from: https://surfshark.com/blog/internet-censorship-2022

[3] Vogels, E. (2020). *Most Americans Think Social Media Sites Censor Political Viewpoints.* Pew Research Center. Retrieved on January 24, 2024, from: https://www.pewresearch.org/internet/2020/08/19/most-americans-think-social-media-sites-censor-political-viewpoints/

[4] Wikipedia (2023). *Ley de los grandes números.* Retrieved on January 28, 2024, from: https://es.wikipedia.org/wiki/Ley_de_los_grandes_n%C3%BAmeros (in Spanish)

[5] Wikipedia (2023). *Ley de Metcalfe.* Retrieved on January 28, 2024, from: https://es.wikipedia.org/wiki/Ley_de_Metcalfe (in Spanish)

UTONOMA

Translated into English from the Spanish original, respecting the aesthetics and trying to preserve the linguistic style.
Proofreading of the original Spanish document by Anayeli Castañeda Covarrubias