# Building Web 4.0: An Economy for Digital Content and Public Goods

White Paper (v2.0.2)

Michael Natanzon
michael@buildingweb4.io
www.buildingweb4.io

**Abstract.** the new digital economy will be based on (1) the concept that online content can be viewed as assets that have value, (2) that we can leverage existing technologies to produce a digital economy based on programmable rules for the production and distribution of digital currency in relation to user input, (3) that digital assets can be ranked based on their credibility and relative importance, and the product of these rankings can be used to establish a standard unit for digital asset valuation – the basis for the digital currency, (4) that the more accurate the ranking of the assets the greater the value of the standard unit, (5) that the creator(s) of digital content (as well as the sources it relies on) should get the credit associated to the content, (6) that creators can use their credit to invest in online projects in exchange for a percentage of the resulting publication's returns, (7) that creators can exchange their credit for other currencies at a market rate. Based on the framework above we are proposing an open-source, decentralized system with algorithmic transparency and with both a network-based and a cryptographic approach to preventing tampering with user data and asset credit. This system will create an economy for digital public goods and all online content, lead to a more efficient distribution of digital resources, proliferation of high-quality reliable content, and proper compensation for content creators.

# Table of Contents

# 1 Introduction

Online content can be defined as any material available on the web, including articles, posts, images, software, videos, and so on. Online content has the properties of a public good in that it is both non-exhaustible and non-excludable. Due to these properties, it is difficult for the market to price online content the same way it can price other products or services in the economy and the result may be market failure.

In economics, a market failure is commonly defined as a condition in which resources are distributed inefficiently in the economy, or where individuals' economic incentives do not lead to efficient results. As such, it is evident that when it comes to online content - and reliable, quality content in particular - all the current business models of the market economy (advertising, subscription, licensing, sponsoring, donations, and self-funding) result in market failure.

## 1.1 Advertising

The advertising-based business model, which is the dominant model for media today, affects both the quality and reliability of the content. Since this model relies on ratings to generate revenue from advertising, content creators have an incentive to prioritize quantity of content over quality. Intense competition for ratings results in a race to the bottom where content creators increasingly rely on divisive, sensational, hyperbolic, and controversial tactics to boost their ratings at the expense of quality and credibility. Such tactics essentially "crowd out" content creators who wish to maintain a high quality of content - which is usually more expensive to produce - and result in the proliferation of low quality, unreliable content. Additionally, when an advertiser directly sponsors a content creator the content is more likely to promote the interests of the advertiser, thus negatively affecting the reliability of the content.

## 1.2 Subscription and Licensing

Both the subscription-based business model, which is the second most popular model for media today, and the licensing-based business model – popular with software distribution – primarily affect efficient resource distribution. Since online content can be distributed in abundance and at virtually no cost (other than the initial cost of production), the laws of supply and demand prescribe that the cost of content should be low as well. Yet, partially due to competition with other business models, subscription and licensing costs are a few orders of magnitude above expected market rates. This means that only a small fraction of people have access to the content, making distribution extremely inefficient.

Since these models do not rely as heavily on ratings as the advertising-based business model there is less pressure to rely on disreputable tactics to boost ratings. However, a content creator is likely to be pressured by his or her base of subscribers or users to align with their views or ideology or suffer a loss in revenue, thus affecting the content creator's intellectual independence and credibility.

Similarly, software companies expend a substantial amount of resources (and countless work-hours) devising mechanisms to restrict the use of their software or prevent people from pirating it. These efforts are rarely successful, but they still contribute to the inefficiency of the licensing business model.

## 1.3 Sponsorship

In the sponsor-based business model - where a corporation, wealthy tycoon or government sponsor a publication – the primary effect is on the credibility of content. In this model distribution may be at no cost and the content may be of high quality – these are due to ample funding from the patron. However,

the reliability of the content is highly suspect as the content creator is largely viewed as promoting the interests of its patron.

### 1.4 Open-Source and Donation

The donations-based business model is perhaps the most favorable among current business models when it comes to efficient distribution of resources, quality and reliability of content. The main drawback of this model however is its lack of economic viability. Indeed, the few notable exceptions where the model works successfully at scale – Wikipedia perhaps being the most prominent one – are the exceptions that prove the rule.

The situation is similar for open-source projects; though these projects score high on efficient distribution of resources, quality, and reliability, there is currently no viable business model for developers to make money directly off of their open-source projects. Certainly, there are ways for developers to make money around the open-source project – paid support, Open Core, selling additional features, and so on – but none of these are based on developers profiting from the open-source project directly. As a result, unless they are privileged to have the financial means to support their open-source work, developers cannot devote significant time and effort to produce high-quality open-source projects, and these projects are not nearly as robust as they would have been had a viable business model was in place.

### 1.5 Summary

Thus, all the current business models of the market economy result in market failure in one form or another; either the business model results in content that is unreliable or of low quality, the model inefficiently distributes resources by restricting the vast majority of people from accessing the content, the model is not economically viable, or some combination of these. Even when a content creator combines multiple business models this does not eliminate the drawbacks but rather compounds them.

The system we are proposing in this paper leverages existing Web 2.0 (Frontend) and Web 3.0 (Backend) technologies to valuate online content, incentivize the creation of high-quality content that is credible and reliable, and efficiently distribute online resources by allowing everyone the broadest access to content. The system is based on four integral components: (1) valuation tool for all digital content, (2) collaboration platform, (3) investment medium, and (4) exchange market.

## 2 Background

Let's consider what the system is trying to achieve: the idea is to have a system that (1) establishes a business model that promotes and facilitates the creation of high-quality reliable content (articles, video, applications, and so on), and (2) helps people determine the credibility of online content.

### 2.1 Quality and Reliability

How does the system determine what content is of high-quality and reliable? We define reliable content as content that is truthful and accurate. We define high-quality content not merely as content that is written or presented well, but rather as content that contributes to people's lives and to the economy and to society as a whole. As such, high-quality content is content that people consider important (to themselves and to society). Each post online (article, blog post, video, open-source project, podcast and so on) receives a Credibility Score and an Importance Score that are based on users' input. These scores are then displayed for each online post (through a dedicated browser, browser extension or mobile app).

## 2.2 Scoring

The Credibility Score is multiplied by the Importance Score to produce a Content Score which represents the value of the post. These content scores are then tokenized and are credited to the creator(s) of the post. If the post is influenced by other sources those sources receive a part of the credit in proportion to their contribution to the post. Thus, creators are incentivized to openly share their content as they are credited for their contribution to others.

## 2.3 Token Value

To make the system a viable business model for content creation the token must be valuable. What makes tokens valuable is the entire digital content ecosystem that they support:

- *Improves the web*. The system creates a positive feedback loop for creating high-quality content, as well as helping people determine the credibility of content and of creators. Some examples include: helping expose scams and fraud online, showing the track record of analysts and pundits, distinguishing those who are experts in their field from trolls, eliminating the need for advertising, indicating the credibility of science and news stories, and much more. The more accurate the rating of credibility and importance of content the more people want to use and participate in this system. The system also provides users with an ad-free search engine, where content can be ranked by its importance and credibility, and filtered by various categories.
- *New content business model*. The system creates a new business model where content creators can directly profit from the content they create or from reviewing the content of others. There are no barriers to entry, no middlemen that receive a cut of the profits, and no restrictions on free-speech. Each content creator simply benefits in proportion to the importance and credibility of their posts or reviews.
- *A new era for software development*. The system creates a viable economic model for open-source projects. Developers will have a strong economic incentive to create applications that greatly benefit society (since those projects will have a higher Content Score). Developers also benefit when others build applications based on their code, and when they collaborate with others to improve and customize the applications. Thus, the system will help usher in a new era of software development centered on creating value for people.
- *Collaboration platform*. The system creates a platform for creators to collaborate with others on larger projects (research articles, applications, books, films, magazines, and so on) where each individual benefits in proportion to their contribution to the project.
- *Investment mechanism*. The system creates a mechanism for individuals to invest in the content of others. Producing high quality content (scientific articles, books, investigative journalism and so on) often requires a significant investment. The system allows individuals to request funding for their project in exchange for a percentage of the tokens generated from the final product. The individual would post a proposal to the system with a description of the project, the expected Importance Score, requested funding amount of tokens, and percent of return offered. Others can then review the proposal and refine the expected return. People can then bid on the proposal (the percent return they're willing to accept). Investors can either use the tokens they generated in the system to invest in projects, or exchange other currencies for tokens to do the same.
- *Medium of exchange*. The system allows the exchange of tokens for other currencies, thus allowing individuals to invest in the system as well as use the tokens to purchase goods and services.

- *Future development*. The system is flexible to allow further development and integration of additional tools to facilitate content creation.

## 2.4 Incentive Structure

Since the Credibility and Importance Scores (and the resulting tokens) rely on the accuracy of users' input, what prevents users from trying to game the system? Can't users just conspire to boost each other's scores to get more tokens, or give high scores to posts they "like" regardless of the importance and credibility of these posts? The answer is that what prevents (or rather, discourages) users from trying to game the system is the incentive structure built into the system. Since what really matters is the token's exchange rate vs other currencies, for the token to have any value vs other currencies the Credibility Scores (CrS) and Importance Scores (IS) of posts and creators have to be meaningful. For these to be meaningful they have to be accurate. The more accurate these scores are, the more value the token would have vs other currencies. It is obviously better to have 1000 tokens with a value of $500 each than a million tokens worth $0.01 each.

The system also requires users to provide sources in their reviews, which means that the scores in the reviews must be justifiable. Even if multiple users try to coordinate a disinformation campaign, if users provide false information all this means is that they're creating an opportunity for many other users to jump into the discussion and make money from refuting their claims (and at the same time such users will keep losing credibility until their misinformation will end up having no effect on the discussion, since their Credibility Score will be below 0). It simply doesn't pay to spread disinformation in the system. And it doesn't matter if there are a hundred, a thousand, or ten thousand users on the side who do that, since their influence is directly proportional to their credibility.

Therefore, each user has the incentive to make their reviews as accurate as possible, both when reviewing posts and when reviewing the comments/reviews of others. Everyone also has an incentive to catch others who try to game the system (by doing so they would increase their own score, reduce the score of the cheaters, and increase the accuracy of the system). This incentive structure promotes the accuracy of scores throughout the system, which makes the Credibility and Importance Scores meaningful, and increases the value of the token as a whole.

## 2.5 Security

What prevents individuals from hacking the system and setting fraudulent Credibility and Importance Scores, thus giving themselves a lot of tokens as well as great influence in the system? While the simple answer may be that the system is built on blockchain technology and therefore difficult to tamper with, what really makes the system nearly impossible to hack is its mechanism for generating scores.

Since each user's scores are the result of other users' inputs, in order for a hacker to fraudulently generate scores for any user in the system a hacker would have to alter all the scores that the user reviewed (since these reviews incorporated the user's original content score). The hacker would then have to alter the scores of all the reviewed users, the scores of all their reviews, and so on. Therefore, unless the hacker adjusts the affected scores of nearly every user throughout the system – a practically impossible task – any alteration would be easily detected through an audit of a user's scores (since all scores in the system have to be internally consistent).

# 3 System Design

The system's core objective is to establish a new business model for content creation based on the value associated with the credibility and importance of online content. To realize this objective the system was designed in such a way that Credibility Scores and Importance Scores throughout the system are meaningful in the real world.

## 3.1 Overview

**Tokens** in the system are generated based on the creation and review of content. Creating high-quality (high **Importance Score**) and reliable (high **Credibility Score**) content results in more tokens (high **Content Score**) than otherwise. Users can invest or trade their tokens.

Users can establish their expertise in different topics (**Categories**) based on posts/reviews they make in those topics or in closely related topics. Users' reviews of other posts are then weighted based on the users' expertise in the post' topics (a post may be associated with multiple topics). Users determine the importance of each topic to their lives and to the overall economy, thus encouraging everyone to contribute to topics that have more significance in people's lives (since posts in those topics are likely to receive a higher Importance Score overall).

Posts are reviewed based on the text (or code) itself as well as their context. The text is broken down into sentences. Users can then rate the credibility of each sentence, comment and provide sources to back up their claims. Users then comment on the context in which the post was made, to provide a full picture of the post's credibility. Each user's review is weighed based on the user's expertise in the topics related to the post. Users can also state their confidence level in their comment, so that their review accurately reflects their knowledge of the topic. Users also rate the importance of the post compared to other posts in relevant categories.

A user can also review other users' comments and reviews for their credibility, thus affecting the weight of each review as well as the Credibility Score of other users. This incentivizes all users to be as accurate as possible in their posts and reviews.

## 3.2 Technology

The system leverages existing Web 2.0 technologies (for Frontend user interactions and review of any content on any webpage) and Web 3.0 technologies (for Backend data security and token programming) to produce a decentralized, trustless economy for digital content.

Users post and review content through a user interface in a dedicated browser, a browser extension, or on a mobile app. Each post or review is stored on a distributed database and its cryptographic hash is added to a multi-layer blockchain. Tokens are minted according to the Content Scores of posts, associated to the creator(s) of the content, and can be transacted on the blockchain.

## 3.3 System

The total number of tokens in the system ($\mathbf{Tokens_{Total}}$) is the sum of tokens that each user has ($\mathbf{\sum User_{Tokens}}$), as shown here:

$$\mathbf{Tokens_{Total} = \sum User_{Tokens}}$$

## 3.4 Users

For each user, tokens (**$User_{Tokens}$**) are made up of the user's total Content Score (**$User_{Total}$ CoS**) plus the tokens the user exchanges from other currencies and the user's investments. This is expressed in the following equation:

$$\mathbf{User_{Tokens} = User_{Total}\ CoS +/-\ Tokens}$$

The only two ways to produce tokens in the system are by creating content and reviewing the content of others. The Credibility Score and Importance Score of posts is what produces a Content Score which generates the tokens. There must be a certain delay period between the time a post is created and initially reviewed and the time tokens for the post are minted – this delay period allows other users to review the post and reduces the possibility of users trying to game the system. Once the tokens are minted, they can then be invested or traded with other users in the system.

### 3.4.1 Credibility and Importance Scores

Why do we need both a Credibility Score and an Importance Score? The simple answer is that to determine the Credibility Score of a user we have to make sure that the numbers reflect something meaningful in reality.

Let's take for example a Financial Analyst – if we don't have an Importance Score (measuring the importance of the claims the person makes), the Financial Analyst can make 9 claims about the values of stocks in the past (e.g. "the stock price of stock XYZ was $157 last December") which are 100% accurate (Credibility Score = 10, on a scale from -10 to 10), and one claim about the value of a stock in the future (e.g. "the stock price of Netflix will be $7000 next year"). If the Financial Analyst is wrong his Credibility Score will decline to 8.0

$$\frac{(9 \times 10) + (1 \times -10)}{(9 + 1)} = 8.0$$

$$\frac{\text{[accurate claims]} \times \text{[\% accurate]} + \text{[inaccurate claims]} \times \text{[\% inaccurate]}}{\text{[\# of claims]}} = \text{[Credibility Score]}$$

This equation can be simplified to the following:

$$\frac{\sum (\ \text{[claim]} \times \text{[\% accuracy]}\ )}{\text{[\# of claims]}} = \text{[Credibility Score]}$$

If our credibility score was determined as described above, that would completely defeat the purpose of the Credibility Score and allow users to manipulate their Credibility Score by making irrelevant (but accurate) claims to offset relevant (but inaccurate) claims.

The purpose of the Importance Score, therefore, is to make the Credibility Score **meaningful** (to give the proper "weight" to each claim). Accurate claims about the values of stocks in the past are valued a lot less by the general public than accurate predictions of future stock prices. Let's say the public values accurate claims about the values of stocks in the past at 0.04 IS (Importance Score), while accurate predictions of future stock prices are valued by the public at 20 IS. Then the person's Credibility Score would be:

$$\frac{[(9 \times 0.04) \times 10] + [(1 \times 20) \times -10]}{[(9 \times 0.04) + (1 \times 20)]} = -9.65$$

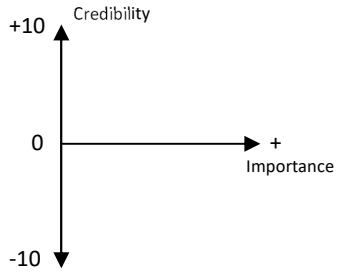
[accurate claims]X[weight of claims]X[% accurate] + [inaccurate claims]X[weight of claims]X[% inaccurate] = [Credibility Score]

$$\sum ([\text{claim}] \times [\text{weight of claim}])$$

This equation is more accurately given by the algorithm:

$$\frac{\sum ( [\text{claim}] \times [\text{weight of claim}] \times [\% \text{ accuracy}] )}{\sum ( [\text{claim}] \times [\text{weight of claim}] )} = [\text{Credibility Score}]$$

With the help of the Importance Score, an 8.0 "credibility score" (which doesn't reflect much in reality) is reduced to a -9.65 Credibility Score (which gives a more accurate representation of the person's track record).




Consider the graph above – what we want is for the Credibility Score to give a high score to someone who posts content of high credibility and high importance, and punish someone who posts fraudulent or misleading information about high-importance subjects. Similarly, we want the reward/punishment to be **proportional** both to the **importance** of the subject matter and to the level of **accuracy** of the post.

### 3.4.2 Measuring the Importance Score

The total number of Importance Score points in the system is always equal to the number of posts. The total number of credits represents 100% of the value (importance) of all posts in the system. Logically, the average rating of a post in the system is 1 IS (ie. dividing the total number of Importance Score points by the total number of posts should give us exactly 1). By normalizing the average Importance Score of a post to 1 IS we get a better sense of how much a post should be worth. Similarly, if a certain user has 150,000 posts, but only has 50,000 IS, we can tell that the user has below average influence.

Normalizing the Importance Score to the average post also allows people to more easily determine the appropriate exchange rate between 1 token and other currencies, since people will be able to see the "average" post (a post with 1 IS) and determine how much the post should be worth in other currencies.

The user's total Content Score (**$\text{User}_{\text{Total}}$ CoS**) is equal to the user's total Credibility Score (**$\text{User}_{\text{Total}}$ CrS**) multiplied by the user's total Importance Score (**$\text{User}_{\text{Total}}$ IS**). It is also equal to the sum of all the user's category Content Scores (**∑Cat CoS**), as shown here:

$$\textbf{User}_{\text{Total}}\ \textbf{CoS} = \textbf{User}_{\text{Total}}\ \textbf{CrS} \times \textbf{User}_{\text{Total}}\ \textbf{IS} = \textbf{∑Cat CoS}$$

The user's total Importance Score (**$\text{User}_{\text{Total}}$ IS**) is made us of the sum of all the user's category Importance Scores (**∑Cat IS**)

$$\textbf{User}_{\text{Total}}\ \textbf{IS} = \textbf{∑Cat IS}$$

The user's total Credibility Score (**$User_{Total}$ CrS**) is equivalent to the average Credibility Score (**Avg. Cat CrS**) across all categories, which is the sum of category Importance Scores multiplied by category Credibility Scores (**∑ (Cat IS x Cat CrS)**), divided by the sum of all category Importance Scores (**∑Cat IS**)

$$User_{Total}\ CrS = Avg.\ Cat\ CrS = \sum(Cat\ IS \times Cat\ CrS) / \sum Cat\ IS$$

## 3.5 Categories

A user's category Content Score (**Cat CoS**) is equal to the user's category Credibility Score (**Cat CrS**) multiplied by the user's category Importance Score (**Cat IS**), or the sum of all Content Scores for posts (and reviews) in a category (**$\sum Post_{Cat}$ CoS**).

$$Cat\ CoS = Cat\ CrS \times Cat\ IS = \sum Post_{Cat}\ CoS$$

A user's category Importance Score (**Cat IS**) is equivalent to the sum of all Importance Scores for posts (and reviews) in a category (**$\sum Post_{Cat}$ IS**).

$$Cat\ IS = \sum Post_{Cat}\ IS$$

A user's category Credibility Score (**Cat CrS**) is equal to the average Credibility Score for posts (and reviews) in the category (**Avg. $Post_{Cat}$ CrS**), which is equal to the sum of all Credibility Scores multiplied by Importance Scores for user posts (and reviews) in the category (**∑ ($Post_{Cat}$ CrS x $Post_{Cat}$ IS)**) divided by the sum of all Importance Scores for user posts (and reviews) in the category (**$\sum Post_{Cat}$ IS**), as shown in the equation below:

$$Cat\ CrS = Avg.\ Post_{Cat}\ CrS = \sum (Post_{Cat}\ CrS \times Post_{Cat}\ IS) / \sum Post_{Cat}\ IS$$

A user's category-related Content Score (**$User_{Cat}$ CoS**) is equal to the user's category-related Credibility Score (**$User_{Cat}$ CrS**) multiplied by the user's category-related Importance Score (**$User_{Cat}$ IS**), which is equivalent to the user's category-specific Content Score (**Cat CoS**) plus the sum of the user's Content Scores from adjacent categories (**$\sum Cat_{Related}$ CoS**)

$$User_{Cat}\ CoS = User_{Cat}\ CrS \times User_{Cat}\ IS = Cat\ CoS + \sum Cat_{Related}\ CoS$$

where the user's Content Scores from adjacent categories (**$Cat_{Related}$ CoS**) is expressed as each category's Content Score (**Cat CoS**) multiplied by a category-relatedness ratio (**$Cat_{Related}$ Ratio**)

$$Cat_{Related}\ CoS = Cat\ CoS \times Cat_{Related}\ Ratio$$

### 3.5.1 Category-relatedness Ratio

What is a category-relatedness ratio (**$Cat_{Related}$ Ratio**)? The purpose for this ratio is to meaningfully capture users' expertise in a field. If a user has many posts in a certain field, but few posts in another closely-related field, without a category-relatedness ratio the user's expertise in the closely-related field will seem to be minimal. The ratio is therefore supposed to adjust the user's expertise so that it meaningfully reflects the user's expertise in the closely-related field.

A category-relatedness ratio (**$Cat_{Related}$ Ratio**) for any two categories is calculated by taking the Importance Scores for all posts throughout the system where both categories are used (**∑ ($Post_{Cat1\&2}$ IS)**) and divide it by the Importance Score for all posts under either category – ie. the sum of all Importance Scores for each post in the first category (**∑ ($Post_{Cat1}$ IS)**), all posts in both categories (**∑ ($Post_{Cat1\&2}$ IS)**), and

all posts in the second category (**∑ ($Post_{Cat2}$ IS)**). The posts' Importance Scores are used to give each post its proper weight.

$$Cat_{Related}\ Ratio = \sum (Post_{Cat1\&2}\ IS) / ( \sum (Post_{Cat1}\ IS) + \sum (Post_{Cat1\&2}\ IS) + \sum (Post_{Cat2}\ IS) )$$

A user's category-related Importance Score (**$User_{Cat}$ IS**) is equal to the user's category-specific Importance Score (**Cat IS**) plus the sum of the user's Importance Scores from adjacent categories (**$\sum Cat_{Related}$ IS**)

$$User_{Cat}\ IS = Cat\ IS + \sum Cat_{Related}\ IS$$

where the user's Importance Scores from adjacent categories (**$Cat_{Related}$ IS**) is expressed as each category's Importance Score (**Cat IS**) multiplied by a category-relatedness ratio (**$Cat_{Related}$ Ratio**)

$$Cat_{Related}\ IS = Cat\ IS \times Cat_{Related}\ Ratio$$

A user's category-related Credibility Score (**$User_{Cat}$ CrS**) is equal to the average of all category-related Credibility Scores (including both the specific category and the related categories). This is equivalent to the sum of the user's Credibility Scores of all related categories multiplied by the user's Importance Scores of all related categories (**∑ ( $Cat_{Related}$ CrS x $Cat_{Related}$ IS )**) multiplied by a category-relatedness ratio (**$Cat_{Related}$ Ratio**), all divided by the sum of the user's Importance Scores of all related categories (**∑ $Cat_{Related}$ IS**) multiplied by a category-relatedness ratio (**$Cat_{Related}$ Ratio**).

$$User_{Cat}\ CrS = Avg.\ Cat_{Related}\ CrS = \sum ( Cat_{Related}\ CrS \times Cat_{Related}\ IS \times Cat_{Related}\ Ratio ) / \sum ( Cat_{Related}\ IS \times Cat_{Related}\ Ratio )$$

## 3.6 Posts

A post's Content Score (**Post CoS**) is equal to the product of a post's Credibility Score (**Post CrS**) and the post's Importance Score (**Post IS**), subtracting the share of the Content Score that is attributed to posts and sources that influenced that post. This is represented by the following equation:

$$Post\ CoS = ( Post\ CrS \times Post\ IS ) \times ( 1 - \sum Influence\ Credit )$$

To make a post's Credibility Score (**Post CrS**) meaningful, the post is reviewed by users based on the credibility and relative importance (within the post) of each sentence in the post, as well as the credibility and relative importance of the context of the post to the post as a whole. Users need to comment on each of the sentences and context that they review and provide sources to back up their claims (it is not necessary however for a user to review all the sentences in a post). Subsequently, other users can evaluate these comments and the sources provided and review those as well. Each user has the incentive to be as accurate as possible in their review process, since providing false or misleading information will result in the user losing credibility.

The post's Credibility Score (**Post CrS**) therefore is the average credibility score for the post's context and for each sentence in the post. The post's context Credibility Score (**Context CrS**), and each of the sentences' Credibility Scores (**Sentence CrS**) is multiplied by the relative importance within the post for the context (**Context IS**) and for each sentence's importance score (**Sentence IS**) respectively. This is represented by the following equation:

$$Post\ CrS = Context\ CrS \times Context\ IS + \sum (Sentence\ CrS \times Sentence\ IS) / \sum (Sentence\ IS) \times ( 1 - Context\ IS )$$

A post's context Credibility Score (**Context CrS**) is determined by user reviews. Each user's review is weighted by the user's Content Score in categories related to the post (**$User_{Rev}$ Cat CoS**) and the user's context Confidence Level (**$User_{Rev}$ Context CL**). This is represented by the following equation:

$$\textbf{Context CrS} = \textbf{Avg. Rev Context CrS} = \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL} \times \textbf{Rev Context CrS}\ ) / \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL}\ )$$

Users can assign a confidence level to comments to reflect their level of certainty in their claims (by default, each comment has the highest confidence level). By assigning a lower confidence level the user reduces the weight of their comment to the overall post's score. However, if the user made a mistake and other users review the user's comment, the user will lose less credibility (in proportion to the user's confidence level). Users therefore have an incentive to accurately represent their confidence level in their comments.

Similar to the post's context Credibility Score, the post's context relative Importance Score (**Context IS**) is weighted by the user's Content Score in categories related to the post (**$User_{Rev}$ Cat CoS**), and by the user's context Confidence Level (**$User_{Rev}$ Context CL**). It is measured on a scale from 0 to 1, and represented by the following equation:

$$\textbf{Context IS} = \textbf{Avg. Rev Context IS} = \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL} \times \textbf{Rev Context IS}\ ) / \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL}\ )$$

The calculation for a post's sentence Credibility Score (**Sentence CrS**) and a post's sentence Importance Score (**Sentence IS**) work much the same as the ones for the context scores. These are represented here:

$$\textbf{Sentence CrS} = \textbf{Avg. Rev Context IS} = \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL} \times \textbf{Rev Sentence CrS}) / \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL}\ )$$

$$\textbf{Sentence IS} = \textbf{Avg. Rev Context IS} = \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL} \times \textbf{Rev Sentence IS}) / \sum (\ User_{Rev}\ \textbf{Cat CoS} \times User_{Rev}\ \textbf{Context CL}\ )$$

The post's Importance Score (**Post IS**) is the sum of all the importance scores by category (**$Post_{Cat}$ IS**) multiplied by the rank of that category (**Cat Rank**), as shown here:

$$\textbf{Post IS} = \sum (\ Post_{Cat}\ \textbf{IS} \times \textbf{Cat Rank}\ )$$

A post's category Importance Score (**$Post_{Cat}$ IS**) is the average of users' reviews of post category importance scores (**Rev $Post_{Cat}$ IS**) weighted by users' overall Content Score (**$User_{Total}$ CoS** or **$User_{Rev}$ $CoS_{Total}$**). This is represented by the following equation:

$$Post_{Cat}\ \textbf{IS} = \textbf{Avg. Rev}\ Post_{Cat}\ \textbf{IS} = \sum (\ User_{Rev}\ CoS_{Total} \times \textbf{Rev}\ Post_{Cat}\ \textbf{IS}\ ) / \sum (\ User_{Rev}\ CoS_{Total}\ )$$

A category rank (**Cat Rank**) is the average of users' category ranking (**User Cat Rank**) weighted by the users' overall Content Score (**User $CoS_{Total}$**), as shown here:

$$\textbf{Cat Rank} = \sum (\ \textbf{User}\ CoS_{Total} \times \textbf{User Cat Rank}\ ) / \sum (\ \textbf{User}\ CoS_{Total}\ )$$

## 3.7 Investment

The investment mechanism in the system is based on users providing details on the project they're working on, the expected Importance Score of the project (**$Project_{Cat}$ IS**) based on the project's related

categories. The system then calculates the project's expected Credibility Score (**$User_{Cat}$ CrS**) based on the user's category-related credibility score. This results in the project's expected Content Score (**$Project_{Exp}$ CoS**)

$$\mathbf{Project_{Exp}\ CoS = User_{Cat}\ CrS \times Project_{Cat}\ IS}$$

The user then sets the level of funding he or she needs for the project (in tokens), and the percentage of the expected return the user is willing to provide in return for the investment. The project is then put out for public review. Users in the system can then review the project and provide their estimated Importance Score for the project (each user's review is weighted based on their overall Content Score). The project's expected Content Score then changes based on user review.

Subsequently, the project is put out for bid and users in the system can set the percentage of return they're willing to accept in return for the funding (as well as the full or partial amount of funding they're willing to provide). The system then assigns the lowest percentage of return bid (or bids, for partial funding) provided, and tokens are transferred to the project's owner (either immediately or in stages, per the conditions of the project).

### 3.8 Remarks

The mechanisms and algorithms presented above were designed to provide meaningful importance and credibility scores throughout the system, which then result in the generation of tokens. The value of the tokens in the system reflects how accurately the scores map to reality. Though we have confidence in our approach to valuating content, we do not believe that it is the only approach. In fact, we welcome competition in this field – both to improve the algorithms presented here, and come up with alternative approaches to content valuation. Ultimately, more research and competition in the field will result in better algorithms and a more effective business model for online content creation that will be at the foundation of Web 4.0.

## 4 Implications

The system we are proposing has far-reaching implications in the digital space; it opens the door to incredible progress and innovation in content creation, and may serve as a powerful engine of growth for the broader economy. Let's consider some of the areas where the system may have a major impact:

### 4.1 Social media

In the current ratings-based social media environment the most extreme, controversial, polarizing and outrageous posts tend to grab the most attention, and therefore dominate the conversation. The system we are proposing offers an alternative paradigm that promotes civil discourse and truthfulness. Instead of focusing on getting the most 'likes' people will consider what they can contribute to the conversation.

The system would also reduce the spread of misinformation on social media – without the need for social media platforms to censor or suspend anyone; misleading posts would simply get a low credibility score, while online trolls and bots would have little to no influence in the system due to their low (or negative) content scores.

## 4.2 Analyst track record

One of the built-in features of the system we are proposing is the ability to see the track record of posts a creator makes (either related to a particular post, topic, or overall). This feature allows people to more easily determine the reliability of analysts and commentators who make predictions in fields such as finance, politics, sports and so on. Seeing the track record of financial analysts, for example, will allow individuals to make better decisions with their investments, while at the same time bring more transparency and accountability to the field.

## 4.3 Scientific research

The system offers an alternative to government grants and paywalled scientific journals in scientific research. It allows scientists to directly profit from their research (based on its accuracy and importance to the public), and receive funding from other users through the system's investment mechanism. The system would allow scientists to maintain intellectual independence in their research, allow researchers (and the wider public) to quickly evaluate the quality and reliability of articles in any field, and give the public full access to scientific articles – thus making significant improvements in how scientific research is done and disseminated.

## 4.4 Transform online experience

The digital tools that the system offers have the potential to transform many facets of people's online experience; they allow users to easily identify online scammers and charlatans, as well as help detect if any app has malware (based on user reviews). They provide students with the tools to distinguish between experts and trolls, and between reliable content and crackpot theories – on any topic and on any website. They allow consumers to make informed decisions on the products and services they buy – without having to deal with the conflicts of interest or hidden motives of reviewers in the current online environment. They allow users to search for the information they need, and filter it by category and credibility level – instead of relying on algorithms designed to maximize corporate ad revenue. They promote an online experience where the user enjoys full access to online content, and where content creators can profit handsomely from producing quality content without the need to rely on advertising or subscription.

# 5 Conclusion

The system we have proposed in this paper solves a market failure for online public goods and creates an economic model that results in high-quality and reliable digital content, and efficient distribution of digital resources. It leverages existing Web 2.0 and Web 3.0 technologies to create a decentralized, open-source system with algorithmic transparency. Content creators get credit associated with their digital assets based on the product of the importance and credibility scores of the content. Content creators can exchange their credit for other currencies, or invest it in other content creators. The interconnectedness of user-generated data makes it easy to uncover any tampering with user data. The system is designed to economically incentivize all users to be as accurate as possible in their reviews, and to produce reliable content that benefits the common good – thus, fundamentally transforming the web as we know it.