WhistleAI

A Marketplace to Anonymously Buy and Sell High Value Information

Version 1.0

Overview

WhistleAI seeks to attack corruption on the world stage by creating a decentralized platform wherein whistleblowers can safely and anonymously "blow the whistle" without fear of retaliation. WhistleAI is a decentralized marketplace that provides whistleblowers anonymity and allows them to monetize information that, until now, would never have seen the light of day. For buyers of information, such as journalists, new media, or watchdogs, WhistleAI uses a unique combination of artificial intelligence (AI) and crowdsourcing to verify the accuracy of the information being sold, while keeping the substance of the information private and unknown to anyone but the seller.

The Problem

When crime or corruption take place, there are often witnesses or people who have knowledge of the act. These witnesses fall into two categories: the bystander and the whistleblower.

Being a bystander is easy. Do nothing or say nothing and life will continue on as usual. In many cases, there may even be a financial or other benefit for remaining silent.

Being a whistleblower is hard. There's very rarely a real incentive to expose the corruption of powerful people or organizations, and the prospects of career suicide and even violent retaliation are real.

First, a solution would be to protect the whistleblower's anonymity. That would protect him from retaliation.

Second, fighting corruption requires that blowing the whistle be more lucrative than remaining a bystander. How can we make it more economically appealing to be a whistleblower rather than a bystander?

These two problems can be solved at once with a blockchain solution that simultaneously protects the identity of the whistleblower while making it economically feasible to expose corruption. Introducing WhistleAI, the solution to corruption in the world.

How do we protect whistleblowers while financially incentivizing them to do the right thing? How do we build a model that will turn more bystanders into whistleblowers? Until fairly recently, an immediate answer to questions like these was not forthcoming.

First is the problem of anonymity. We must make it so that the whistleblower cannot be identified. Let's take a look at a typical information release where the whistleblower is kept anonymous.

An employee of XYZ corporation discovers evidence that XYZ Corp is dumping hazardous materials illegally. The corporation is saving millions of dollars by doing this, but the whistleblower can't stand

what is going on, so she decides to reveal her information to the world. How does she do this while protecting her identity?

In the current paradigm, our brave whistleblower can go to the press. She can take documents that prove the fraud beyond doubt and show them to a reporter. However, the whistleblower is now faced with the fact that she is risking everything: her career, her reputation and possibly her safety. Everything depends on the reporter, a person whom she has never met. Will this person protect her identity at all costs? Would this person turn down a million dollar payment from XYZ Corp in return for revealing his source? Would this person go to jail for the whistleblower if necessary?

Unfortunately, most whistleblowers decide that putting their lives in the hands of another human being is too risky. While we do hear about the occasional reporter who is willing to do jail time to protect a source, this is the exception, not the rule.

There are also many situations today where a newspaper may not have any interest in covering the information the whistleblower is trying to expose. This is where WhistleAI can have far more impact than what's currently available. WhistleAI can act as a connector between whistleblowers and a worldwide community of prospective buyers. In our scenario, perhaps the journalist isn't interested in covering the story, but there are environmental groups that would be extremely interested. Perhaps it's a vlogger or an online magazine who wants the story.

While media organizations and watchdog groups are clear examples of interested parties, we believe the real, true market potential for whistleblowers' information remains largely untapped.

New Media

This untapped market is new media. Today, traditional or legacy media are being disrupted before our eyes. Millions of citizen journalists now exist throughout the world, many of whom specialize in thousands of different subject matters. Many garner more views online than their legacy media counterparts. We believe that new media is ultimately a force for good and another step toward a more decentralized world. Whatever the subject matter is, there an alternative outlet searching for information that can break a story. We believe that WhistleAI can empower new media in ways never before seen. It will give them a marketplace to shop for stories and information that has always existed but has never been readily available until now.

WhistleAI stands to be a conduit between two massive markets that have never connected. As of right now, there are millions of new media journalists and content creators all over the world. The subjects they create content around are almost infinite. Some deal with politics, some with various social issues, others with food, others with animal rights and the list goes on. This represents a massive market of untapped buying potential for information that is relevant to their passion or cause. We see new media

as a smoldering bundle of sticks, waiting for fuel to fully ignite the fire. WhistleAI, in this scenario, will be gasoline. By connecting these content creators with millions of people in everyday situations where they witness wrongdoing, we can give them potentially endless sources of content to build their following. The age-old market of gossip finally has a buyer. It's easy to see the potential for WhistleAI at a higher level of crime, Wall Street, government, big business, but think about the possibilities on a more micro scale. Think about the waiter at the restaurant who repeatedly sees the cook intentionally dropping peoples food on the ground before serving it. A relatively small issue in the grand scheme of things. But what if that waiter, who's tired of seeing this happen, can now sell pictures he took of this to a local food blogger for $50? Don't confuse micro for small however, when it comes to impact. Imagine this scenario playing out all over the world, in every industry? We feel that the untapped market potential that exists between new media and everyday witnesses to wrongdoing is enormous.

Anonymity

One of the problems with making whistleblowing safe and lucrative is that once the damaging information is revealed to anyone, it loses its value. How do whistleblowers expose information to the marketplace while simultaneously retaining anonymity and also being able to sell this information to the highest bidder, or outlet of their choice - all while keeping the information secret to any one person until the information has been successfully sold?

At WhistleAI, we believe we have the framework for a solution to this problem. A peer-to-peer network serves three powerful functions that will both protect and compensate whistleblowers. First, it can protect the anonymity of a whistleblower more securely than any existing method through a blockchain-based token geared toward privacy. Second, through a combination of artificial intelligence and crowdsourcing, we can verify the information without any one source having access to the totality of the information. And third - and most important - this anonymous verification process can be used to provide powerful financial incentives for exposing corruption.

Having sensitive information exposed to multiple potential buyers presents a big problem though. It's not like selling a car or a laptop. You can show your car online and let people bid on it, but you can't do the same thing with a document proving malfeasance. Once the document's information is disclosed - especially if that information is newsworthy - the document loses most of its value. How does a whistleblower shop this document around without destroying its value? The answer lies in minimum disclosure authentication.

The idea behind minimum disclosure is the concept of breaking up a secret into fragments, and allowing anonymous community members to verify the secret piecemeal, so that no one person knows the secret, but the secret is still verified and its value authenticated.

Crowdsourcing algorithms work well for authenticating general information, but what if that data is sensitive and requires privacy? In these cases, we need to figure out how to prove authenticity without full disclosure of the information.

In computer science, this problem set is referred to as a zero-knowledge proof. The idea behind the zero-knowledge proof is to create a verification mechanism that allows a separate party to verify underlying information without having to view the information in its entirety.

Minimum-disclosure protocols, or minimum-disclosure proofs of knowledge, are a subset of zero-knowledge proofs. This protocol takes information and chops it into smaller discernible yet incomplete pieces of information. The algorithm then discloses each piece, one at a time, to a verifier until confidence in the underlying information is reached.

Using a peer-to-peer network, we take the idea of minimum disclosure and combine it with crowdsourcing to achieve a decentralized, private information authentication algorithm. The algorithm splits information into chunks prior to any encryption of the information and sends those pieces to different nodes within the peer-to-peer network, asking for independent verification of each chunk. No one user will be able to verify the object is absolutely what the asset owner states it is; however, if any chunk is obviously different from what the owner states it should be, then the entire asset's authenticity is invalidated.

Here's an overly simplified example.

> *Suppose a whistleblower has a photo of John Doe stealing a cookie out of a jar. For the purposes of this hypothetical, let's assume that the cookie theft is of great importance to a lot of people and that the fact that John Doe is the culprit is a valuable secret.*
>
> *Once the world knows that John stole a cookie, the information is no longer valuable and will be impossible to sell. But if the network can verify the secret without actually revealing it, then the photo can be sold at a good price, thus keeping the financial incentive for whistleblowing intact.*
>
> *The network would have built-in protocols that would be able to parse the expression "This is a photograph of John Doe stealing a cookie from a jar." It would then break that statement up into two parts: "This is John Doe" and "This is a cookie being stolen."*
>
> *It would then create two versions of the photo. The first would blur out John Doe's face. Members of the community would be compensated with cryptocurrency for merely answering the question: "Is this a photo of someone stealing a cookie from a jar?" If a sufficient number of users answer yes, then the first part of the secret is then considered authentic.*
>
> *The second version would blur out everything but John Doe's face, and ask the members of the community "Is this a photo of John Doe?" If a sufficient percentage answers yes, then the second part of the secret will be considered authentic.*
>
> *If both parts of the secret are successfully authenticated, then all of a sudden we have a verified secret with great monetary value, but without having revealed the secret to any one person. This*

secret would then be offered to a network of buyers with the simple statement "Whistle has verified a photo of John Doe stealing a cookie from a jar." With the network having authenticated the photo, the monetary incentive for disclosing criminality will have been preserved.

Figure 1: A Minimum Disclosure Protocol Example Breaking Information into Four Pieces

Whistle AI Network

Traditional whistleblowers give value to information by disclosing their personal identity as well as their relationship to the information so that it can be branded as trustworthy. With the anonymous sale of information, we must find another way to replace this idea of trustworthiness to extract the value from the information without attaching it to any individual as the source. The answer is the Whistle AI Network.

Under the hood, it is is a peer-to-peer network that connects machine learning with human crowdsourcing to facilitate the mechanics of the minimum disclosure protocol algorithm.

The network itself is comprised of a group of nodes communicating together to authenticate data that has been submitted to the network. The authentication process is made up of four components: input processing, machine learning, human analysis and accuracy oversight.

Input Processing

A request to authenticate a piece of information must be submitted through a peer-to-peer node. The information must be an accepted file format (png, pdf, mp3, mp4, avi ,docx, etc.) in order for the request to be processed. Based on the digital file type, the network then creates a sequence order of required authentication steps to be taken and keeps track of the progress of the request.

Machine Learning

In the first phase of authentication a digital file will be passed through different machine learners using algorithms to filter for specific things on specific types of data. These filters can range from relevancy filters to determine the piece of information will be valuable on the network, all the way to facial or voice recognition that helps identify an individual's identity.

Human Analysis

As part of the minimum disclosure protocol, the digital data file is altered for privacy and anonymity of the individuals in the pieces of information by machine learning algorithms. It is then passed to our human nodes whose function is to group analyze the same piece of information and extract a yes or no answer about the piece of information.

Not all participating human nodes will vote on every piece of information that comes through the system. Human work assignments are delegated based on an ordering algorithm, called Priority Rank, which takes into account current time spent on the platform as well as an historical accuracy score of work done on the network (see Priority Rank details below under the WhistleAI Gig Economy section).

Accuracy Oversight

As the final step before each piece of information can be deemed fully verified it must pass through the oversight measures we have put in place in order to increase our rate of accuracy. The primary means of oversight is by a group of pre-defined moderators, called Merit Officers. A credentialed Merit Officer has earned their participation level through consistently high-quality authentication work on the network. The main function of a Merit Officer is to monitor for flagged authentication requests as they are processed whether it is by a machine learning algorithm or a human.

Additionally, we will also include an escrow and arbitration service for added assurance that a user gets exactly what she was expecting when she purchases a piece of information. In these cases a buyer spends an additional 1% in cost for the option to use an arbiter if the piece of information they bought is fraudulent. After the purchase, the funds of the sale will be held in escrow for 48 hours before the seller can claim it. During that time, if the information is deemed "not acceptable" by the buyer, the buyer can open a complaint. Once a complaint is opened, a Merit Officer is randomly assigned to the case through a round robin algorithm. The Merit Officer then looks at the information in question to determine

whether or not the seller acted in good faith. The Officer's arbitration ruling on the complaint is final. Both parties must acknowledge this finality ahead of time.

A more detailed explanation of the technical details can be found in the separate WhistleAI Full Technical white paper.

Application Security

Peer-to-peer networks face different security challenges than the centralized client/server networks such as amazon.com, facebook.com, and other modern web applications. Client/server application architectures have a central security model which means all data remains on a protected server. This generally creates a scenario where only a single access point needs to be secured. The common web protocol used to secure these application access points is HTTPS.

Peer-to-peer networks have no central server, so security is a much harder problem to solve. Every communication between users needs to be independently secured so that it can be trusted. The current best practice of secured peer-to-peer networks is asymmetric encryption, which is also known more commonly as public key encryption. In this model, two unique keys are created to encrypt and decrypt the messages being sent through the network. If one of these keys remains private to an individual peer within the network, and the other shared with all other users, communications can be privately encrypted by other users. Also, messages can be guaranteed to have been generated by the source. These properties create two useful security protections.

First, a user can send a message to another user encrypted with the public key so that only she will be able to decrypt it with the unknown private key. Second, messages encrypted with the private key are verifiable to anyone with the public key. This reversal is useful if one wants to verify that the message originated from the expected source. This latter capability is called a digital signature, so that all messages can be verifiably pointed back to their original destinations. When requiring user-to-user authentication instead of user-to-server authentication, digital signatures are key to secure communications.

Security of peer communications and anonymity of payments are both crucial to the success of WhistleAI. We are fully aware that vital information only maintains its value while it remains private. To protect the integrity of the information's value for the seller, it must be stored securely on our network and can only be passed along to trusted nodes. Within the network, the only nodes that should receive a full piece of information are the node of the owner of the information herself, as well as the machine learning computational nodes that obfuscate the information prior to human analysis.

To ensure that the information remains only within the intended nodes, we use digital signatures to safely ensure accuracy of the source and public key encryption to all but guarantee that the information is only decryptable by the intended receiver. For this functionality, we will use a Blake256 hash function[1]

[1] https://en.wikipedia.org/wiki/BLAKE_(hash_function)

, which is currently in use by the Decred cryptocurrency project to secure their blockchain. Blake256 is designed as a safer future alternative to the SHA256 hash function that is used by Bitcoin.

The token used by WhistleAI shall be deployed using blockchain technology that has been production battle tested for security as well as proven anonymizer algorithms to protect the identity of our users.

WISL Token

In order to preserve a sellers anonymity and allow them to be paid for their contributions, WhistleAI requires an anonymous compensation method. A precedent already exists for private payments in the form of privacy based cryptocurrencies such as Monero, Dash, Zcash, Verge and most recently Komodo and Bitcoin Private. The Whistle Platform uses its own type of privacy tokens, named Whistles (WISL). WISLs serve a dual purposes as they will be the single payment method for buying and selling assets on the network, but they will also be used by human crowdsourcing participants to stake while they help authenticate information on the network.

Market Size

Corruption never seems to be in short supply, in fact it seems to be one of the worlds most abundant, renewable resources. And while there has always been an interest in exposing corruption, a true market for it has never been able to form. Legacy media, the traditional buyer for such information, always represented a centralized “bottleneck” for information, which limited the size of the potential buying market for information. In the mid 20th century, Journalist A.J. Liebling stated that "freedom of the press is guaranteed only to those who own one." Meaning that only those voices powerful enough to produce a book or a print a newspaper could be heard. Today, the power of the press is being disrupted and redistributed among the people. This represents a massive shift in the economy of information. We are witnessing the creation of a worldwide market for information, this likes of which this world has never seem. The potential size of this market is unknown yet, but we estimate it to be in the billions of dollars, with exponential growth following.

However there are still challenges facing this new media. “Fake News” as it’s been labeled, has plagued New Media and now Legacy Media is even coming under fire. The problem is that “trust” in the information being provided by media has traditionally been granted through the authority of the alogarkeys that control it. As New Media is coming online, there is no authority to vouch for the information being provided, as the world has had in the 20th century. And as New Media and Legacy Media create conflicting content, both are feeling the pinch of an erosion in trust by the viewer.

We have seen a similar phenomenon with the evolution of Ecommerce. A shift in trust from the big box retailer to the individual seller of goods had to take place. In the Ecommerce world, it happened through a combination of mechanisms such as online reviews and frameworks, like Ebay, which lent legitimacy to the transaction, without actually being the buyer or the seller in the transaction. Buying

something before physically inspecting it is a very risky proposition. Allowing potential buyers to sample trusted customer reviews before purchasing created enough of a comfort level to allow consumers to buy things online. Buying something online will always be more a convenient proposition than having to go to a brick and mortar store, as long as you can trust the seller. Customer reviews of products supplied that trust, and as the public trust in e-commerce has grown, it has become over a trillion dollar per year market.

We feel that today there are two massive markets, information sellers and New Media, that need a conduit in which to transact. A framework that allows buyers and sellers to transact without being directly involved. Just as Ebay first provided the validation and trust needed for online transactions between individuals to take place on a massive scale, WhistleAI will provide the framework and verification for information to be transacted between New Media / Legacy Media buyers and the world of everyday people who witness wrongdoing and corruption.

While it's difficult to put an exact number on this emerging market for information, we feel that many similarities hold true between it and the birth of Ecommerce. We are injecting trust for online transactions when it only existed previously face-to-face. We are also creating buyer convenience and saving content creators and other buyers days or weeks worth of research.

Using Whistle

At its core, WhistleAI is a marketplace for content that connects buyers with sellers.

Figure 2: Flow of Information in the Whistle Marketplace

The Content

When content is uploaded to the system, it includes a digital file, description, initial list price for the sale and any other supporting documents required for authentication.

The description is used to make sure the content follows the intentions of the network per the founding principles to help govern the network. It can also be used as the basis of understanding for the human's within the authentication algorithm to make a judgement call to flag the piece of information as different from what a buyer would expect it to be. If more than 50% of crowdsourced participants flag a piece of information, then it will not pass authentication and not be sellable through the marketplace.

The Buyer

Buyers within WhistleAI have the option to disclose themselves and have their identity verified or remain anonymous. Verified buyers have a distinct advantage because sellers will be able to seek them out with offers for relevant information. Anonymous buyers must search through the network for interesting information themselves. Making it even harder is sellers may requiring the buyer of their information to be verified, making the possible pool of information to buy smaller.

The Seller

The only information stored about the seller through the network is a unique ID, that is just a unique number. This ID must be different from the wallet address to help preserve a sellers anonymity. All other information about the seller should never leave their own peer node and will permanently delete itself after a transaction has been completed.

The seller is ultimately in control of what happens to the information. For instance they can restrict to only sell to buyers that are known and verified so that they can maintain an understanding of what will happen with the formation after the sale.

Transaction Breakdown

Before a piece of information is sellable on the network, it must first be authenticated. The cost to get a piece of information verified by the network is a flat upfront fee plus a portion of the sale price going to the network profit pool.

The flat fee changes on a daily basis and will be tethered to the price of gold set in USD in order to keep the fee dynamic, but also less volatile than the price of the token itself. The flat fee proceeds are split by granting 60% to the human validators and 40% going toward the resources providing the computational power that drives the network.

Once a buyer is found and a piece of information is sold, then the proceeds of the sale are split into multiple initial buckets. The first is the seller, who receives 92% of the buy price. The remaining 8% gets split evenly 4 ways between the development fund, the merit officer payment fund, the pool of machine learning resource holders and the pool of human authenticators.

The human authenticators do not receive a direct kickback from the deals they approve. Instead all approved deal sale proceeds go into a general pool that pays out weekly proportional to each participants Priority Rank for that week. This mechanism is to help remove the temptation for participants to falsely approve bad information in hope of receiving an allocated share of the profits. If all profits are shared in a way that is directly correlated toward positive accuracy and time spent on the network then the end result should be a more truthful authentication and oversight process.

The Human Element of Authentication

Humans will interact with a Tinder-like interface that allows “swipe right” for a positive verification and “swipe left” for negative vote against the piece of information provided. For each human attempt at authentication, there will be a set number of voters assigned a given task. The voters must collectively approve the asset at a percentage above a set threshold for authentication. Each human’s voting result will have either a positive or a negative effect on their accuracy score based on whether they were

among the majority. The higher a user's accuracy, the faster they are likely to receive another task delegated to them from the routing nodes and the higher rank they will be able to receive.

WhistleAI Gig Economy

Humans are an essential aspect of machine learning and Whistle AI Network. A three part incentive system will be built in to reward the human nodes for their time and participation. This work utility will add another dimension to the gig economy made prevalent by companies such as Uber and Airbnb. Our goal is to have a network of over 500,000 people collectively interacting with and teaching the AI's on our platform by 2020. We feel that a large human network of intelligence to learn from will be a game changer for the progress of Artificial Intelligence.

Three-Part Human Incentive System

Money alone can't motivate every type of person. Recognizing this, we aim to deliver a three-part reward system based on merit, to both incentivize performance and to add the elements of accomplishment and progress to human verifier experience. A person participating in the Whistle gig economy, has three incentives for performing: WISL**,** Ranking and Promotion. When a person joins the Whistle Platform, they begin earning Whistle Coins for accurate verifications. As they verify more and more tasks accurately, they begin to earn badges and ranks within the system. The more badges and the higher one's rank, the more important their role becomes in the Whistle ecosystem. And subsequently, the more they earn per verification. In this model long term loyal workers can get priority access to premium work, but new participants can hang out on the network and built up there priority rank in order to compete with longstanding participants by building up current active participation. A participant's access to quicker and higher paying work is directly correlated to that participants rank relative to other users on the network at that time.

Priority Rank = (Active Participation Multiplier)(Work Accuracy Score)(Total Network Earnings + Amount Staked).

Meritocracy Governance of Participants

Moderation positions to help govern both the crowdsourcing participants and the quality and relevancy of the information being submitted to the network. These moderation positions are called Merit Officers and are paid positions in the Whistle Network. Merit Officers, are the decision makers for the Whistle AI Network and are responsible for keeping the network "White Hat". They are compensated through verification fees.

To become a Merit Officer, a set number of badges and ranks must be achieved. Of those who achieve the proper rank, the top 10% of both total network earnings and work accuracy score will be invited to become Merit Officers.

A White Hat Network

Whistle is fanatical about being a force for good, and will be designed to stay that way. We are acutely aware that sensitive information can be used for good or for evil. Blackmail is a crime for a reason: society doesn't benefit when one citizen threatens to expose another's embarrassing activities - especially activities that harm no one. With that in mind, the WhistleAI contains ground rules that prevents it from becoming a breeding ground for "Black Hat" uses.

The whistle constitution is small set of founding principles and that set present and future moderating guidelines for Merit Officers to enforce in the present and modify in the future. Like any good constitution, the guidelines can be amended by a two-thirds majority vote of the Merit Officers, who are instructed to make sure the guidelines always follow and maintain the ethos of the founding principles.

Founding Principles of The WhistleAI Constitution

1. WhistleAI exists to expose harm being inflicted on others
2. WhistleAI does not exist to inflict harm on others

Founding Set of Amendable Guidelines for Merit Officers Enforcement

1. Abuse of Power is always fair game
2. WhistleAI is not a tool for individual blackmail, so disclosures of affairs, as long as it is not an abuse of power is forbidden.
3. Disclosure of a person's sexual orientation, gender identity are private and not harmful to others so they are strictly forbidden
4. Pedophilia is fair game, but it is forbidden to sell this type of information to the intended target as a form of blackmail.

When content is uploaded to WhistleAI, it must be accompanied by a description. The network will identify and flag descriptions that contain words and phrases that could potentially represent "Black Hat" material. At which point that content will go to a Merit officers for review.
Additionally, as part of the verification process, WhistleAI will compare the content to the description, to ensure that there is a match. If the AI cannot confirm, then the file is flagged and sent for review by Merit Officers.

Conclusion

WhistleAI truly has the potential to change the world for the better. At the sametime, it will create a massive new market for information that has never had an outlet. From multimillion dollar pieces of information on the wrongdoings of the powerful to the $400 piece of evidence that the construction company is cutting corners. WhistleAI has the potential to impact the world for the better, on a global level. Until recent technology advancements, there was never much hope for this kind of change, but today with the decentralized of blockchain, the advent of the cryptocurrency and it's ability to enable secure, private payments as well as the progression of Artificial Intelligence, we are finally reaching a point where technology can provide the anonymity, incentives and verification to allow the truth to come to light. Louis D. Brandeis once said "Sunlight is said to be the best of disinfectants." We believe that WhistleAI can bring sunlight to darkest corners of the world, by incentivizing whistleblowers and simultaneously protecting them from the criminals they expose.