Deuro: Democratize AI ecosystem on the blockchain

August 2, 2018

Abstract

With the exponential advancement of GPU computing power, enormous information, Internet of things, sensors, and different fields in the course of recent years, artificial intelligence has started to take off. According to AngelList, there were over 5000 AI startups worldwide sta (a). MIT technology review documented that $15.2 billion invested in AI startups globally in the 2017 yea. One might say that after the ascent and fall of a few times ever, the time of artificial intelligence has finally arrived!

In the past few years, the Deuro core team has been deeply exploring the cutting-edge technology of artificial intelligence. The core team of Deuro came from both AISense Inc. and Skylight Investment. Backed by top VCs from the bay area, our team have developed better far-field multi-person speech recognition software than Google, Amazon, Microsoft. And sell our speech product to Zoom(video conferencing Unicorn val) and Bridgewater Associates(the $160B world largest hedge fund (ass))

When we develop our AI software solutions, we realize that approximately 10% to 30% of our budget goes into the artificial intelligence cloud computing infrastructure, and at least couple of million dollars for the labeled training data. In addition, we also experienced the trust and security concerns raised by our enterprise clients. Zoom requires us to do an on-premise distribution on their server to protect their user recordings. Bridgewater requires us to pass the software security compliance requirement. This is where Deuro comes in. Deuro is the world first and only artificial intelligence ecosystem on any blockchain. With decentralization, open-governance, cybersecurity and token economy on the mind, Deuro can help AI companies to dramatically reduce their computation cost, shorten the AI system development cycle and increase the production data privacy as well as security.

Contents

1. Introduction

1.1. Applied Artificial Intelligence

1.2. General Artificial intelligence

1.3. AI SAAS solutions

2. Problems in today's AI company

2.1. Production Data security & privacy

2.2. Lack of the training data

2.3. Algorithms

2.4. Low adoption rate

3. Solutions provided by Deuro

3.1. gStorage: Privacy-Preseving Decentralized Artificial Intelligence client-driven data storage system

3.1.a. Feature extraction (Data preprocessing)

3.1.b. Decentralized pseudo-anonymous multi-party key encryption and decryption

3.1.c. AI Data(Conversation) as an asset

3.1.d. Enterprise facing solution

3.1.e. Consumer facing solution

3.1.f. Modeling set repository

3.1.g. Predicting data storage

3.1.h. Models hub

3.2 gCrawl: Decentralized high performance algo-generated training data solution

3.2.a. Crawling task as a transaction

3.2.b. Crawled data as an asset

3.2.c. Fault-tolerant distributed high performance crawling

3.2.d. RAFT consensus algorithm

3.2.e. Decentralized data storage

3.2.f. DFSM decoding path

3.2.g. Transfer learning for updating the model parameters

3.3. gPredict: Decentralized gradient learning framework

3.3.a. Training task as a transaction

3.3.b. Trained model as an asset

3.3.c. Floating Point mapping

3.3.d. Hyper-Parameters routing

3.3.e. Results reducing

3.4. gCompute: Internet-scale AI dApps solution

3.4.a. AI model as an asset

3.4.b. Scalable

3.4.c. Controllable

3.4.d. Difficult

3.4.e. Maintainable

3.4.f. Responsive

3.4.g. Concurrent

3.5. Consensus algorithms

3.5.a. DPoS

3.5.b. Vectority - Convex hull powered ASIC friendly PoW

4. Applications and Future Work

4.1. Aviation

4.2. Education

4.3. Finance

5. Development RoadMap

6. Deuro's Economics and Token economy

6.1. Functional utility token

6.2. The Deuro marketplace

6.3. The miners

7. Core Team Members

8. Advisors

A1. Product

A2. Portfolios

Keywords

AI dApps: Decentralized Artificial intelligence Application whose backend is usually the blockchain

Weak-AI: Weak Artificial intelligence / Applied Artificial intelligence

AGI: Artificial General intelligence

Brute-force search: In computer science, brute-force search or exhaustive search, also known as generate and test, is a very general problem-solving technique that consists of systematically enumerating all possible candidates for the solution and checking whether each candidate satisfies the problem's statement.

AlphaGo(sea): AlphaGo is a computer program that plays the board game Go.[1] It was developed by Alphabet Inc.'s Google DeepMind in London.

GD: Gradient descent is a first-order iterative optimization algorithm for finding the minimum of a function. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point. If instead one takes steps proportional to the positive of the gradient, one approaches a local maximum of that function; the procedure is then known as gradient ascent.

Supervised-Learning: Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs

Unsupervised-learning: Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses.

NBC: In machine learning, naive Bayes classifiers are a family of simple "probabilistic classifiers "based on applying Bayes' theorem with strong (naive) independence assumptions between the features.

CA: Clustering analysis is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters)

DMR: dimensionality reduction or dimension reduction is the process of reducing the number of random variables under consideration[1] by obtaining a set of principal variables. It can be divided into feature selection and feature extraction

PCA: Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

K-NN: k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The k-NN algorithm is among the simplest of all machine learning algorithms.

SVM: support vector machines (SVMs, also support vector networks[1]) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis

RL: Reinforcement learning (RL) is an area of machine learning inspired by behaviorist psychology[citation needed], concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward.

AI-complete: In the field of artificial intelligence, the most difficult problems are informally known as AI-complete or AI-hard, implying that the difficulty of these computational problems is equivalent to that of solving the central artificial intelligence problem—making computers as intelligent as people, or strong AI

NP-complete: In computational complexity theory, an NP-complete decision problem is one belonging to both the NP and the NP-hard complexity classes. In this context, NP stands for "nondeterministic polynomial time". The set of NP-complete problems is often denoted by NP-C or NPC

DL: Deep learning (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms. Learning can be supervised, semi-supervised or unsupervised

CNN: convolutional neural network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks that has successfully been applied to analyzing visual imagery.

RNN: Recurrent neural network (RNN) is a class of artificial neural network where connections between units form a directed graph along a sequence.

LSTM: Long short-term memory (LSTM) units (or blocks) are a building unit for layers of a recurrent neural network (RNN). A RNN composed of LSTM units is often called an LSTM network. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate

Singularity: The technological singularity (also, simply, the singularity) is the hypothesis that the invention of artificial super intelligence (ASI) will abruptly trigger runaway technological growth, resulting in unfathomable changes to human civilization. According to this hypothesis, an upgradable intelligent agent (such as a computer running software-based artificial general intelligence) would enter a "runaway reaction" of self-improvement cycles, with each new and more intelligent generation appearing more and more rapidly, causing an intelligence explosion and resulting in a powerful super intelligence

GAN: Generative adversarial networks (GANs) are a class of artificial intelligence algorithms used in unsupervised machine learning, implemented by a system of two neural networks contesting with each other in a zero-sum game framework

SAAS: Software as a service is a software licensing and delivery model in which software is licensed on a subscription basis and is centrally hosted.

Colo: A colocation centre (also spelled co-location, or colo) or "carrier hotel", is a type of data centre where equipment, space, and bandwidth are available for rental to retail customers.

On-premise distribution: On-premises software (sometimes "on-premise" or abbreviated as "on-prem") is installed and runs on computers on the premises (in the building) of the person or organization using the software, rather than at a remote facility such as a server farm or cloud

DFSM: Deterministic finite state machine is a finite-state machine that accepts and rejects strings of symbols and only produces a unique computation (or run) of the automaton for each input string

MSE: Mean squared error

BFT: Byzantine fault tolerance

Force-align: Forced alignment refers to the process by which orthographic transcriptions are aligned to audio recordings to automatically generate phone level segmentation

MFCC: Mel-frequency cepstral coefficients - computing feature representation of audio

1.Introduction

The statistic shows the size of the artificial intelligence market worldwide, from 2016 to 2025. In 2017, the global AI market is expected to be worth approximately 2.42 billion U.S. dollars. Some current major uses of artificial intelligence include image recognition, object identification, detection, and classification, as well as automated geophysical feature detection. Today, there exists no decentralized way for company to train, evaluate, test and deploy their AI models at scale. Our mission is to decentralize and disrupt the whole ML industry by creating open market inclusive for all key players, which will stimulate synergy and speed up development of artificial intelligence.

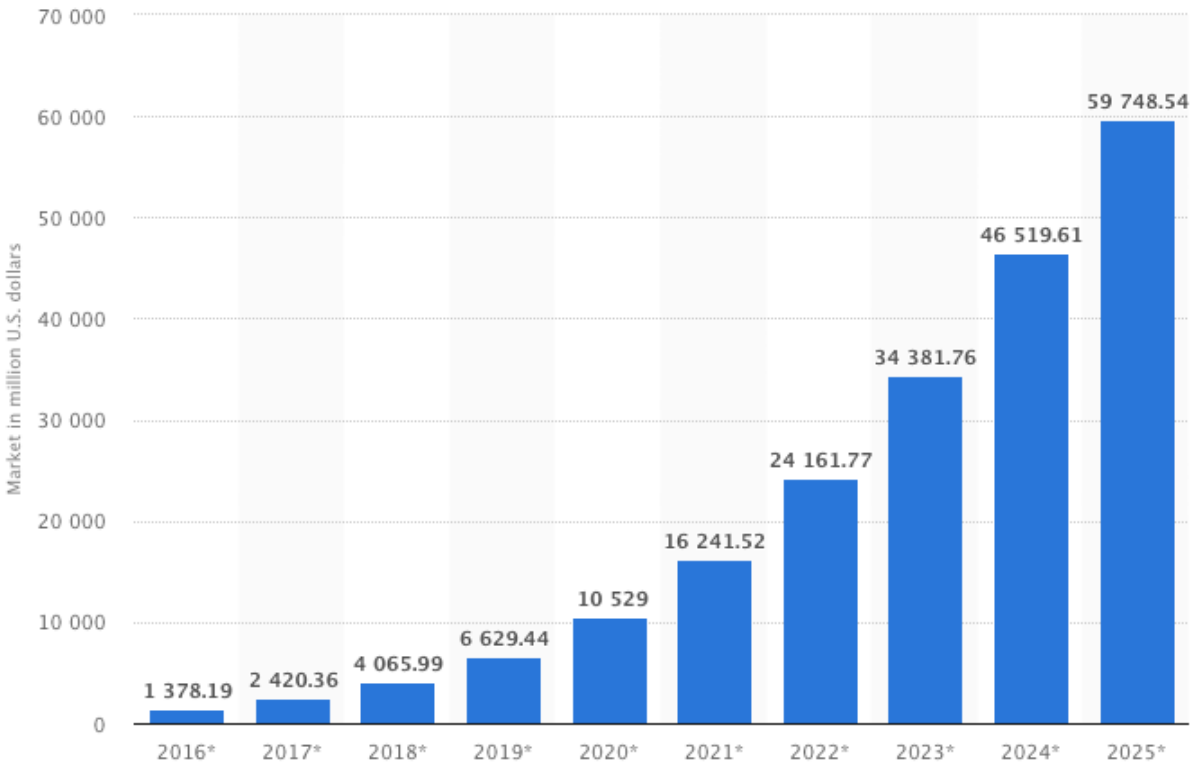


Figure 1: **Global AI market size 2016-2025 sta (b)**

1.1 Applied Artificial Intelligence

Most of the AI today is Weak-AI. One of the key element that distinguish us, human-being, from everything else on the planet is insight. This capacity to comprehend, apply information and enhance aptitudes has assumed critical part of our development and setting up human civilisation. Yet, numerous individuals (counting Elon Musk) trust that the progression in innovation can make

super intelligence that can surpass human presence. The reason that we develop AI is because some of the problem today can't be solved in polynomial time, so we have to use AI to solve the NP-complete problems.w

> **"Within thirty years, we will have the technologicalmeans to create super-human intelligence. Shortly after, the human era will be ended."** Vernor Vinge

The majority of the AI frameworks set up today are Weak Artificial Intelligence, which were intended to tackle a particular issue. Indeed, even AlphaGo, which could beat human champions in table game Go is thought to be a tight AI (Weak AI). Go is a more mind boggling amusement than chess (Brute-constrain hunt won't help this time), its played on a 19x19 board. What makes Go complex is the way that the game is played absolutely by intuition and lessons from past experience are not by strict standards that specialists can clarify. Likewise in Go there is unlimited potential outcomes for each given move.

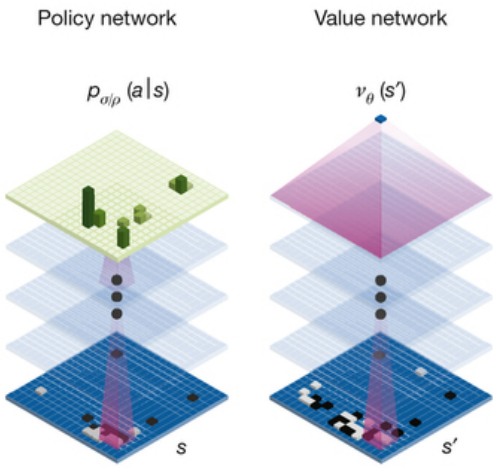


Figure 2: AlphaGo's core PN & VN

1. Value networks:

The purpose of 'value networks' is same as our heuristic function. DeepMind says that 'Ideally, we would like to know the optimal value function under prefect play. In practice, we instead estimate the value function for our strongest policy, using the RL policy network'. It is actually smart that they approach this evaluation problem in an indirect way.

2. Policy networks:

Supervised learning of policy networks(SL) and Reinforcement learning of policy networks(RL) are the first and second stage of their training pipeline. SL is learned from human Go expert. RL is learned by self-playing. Deep neural network is improved through those two steps instead of our hard-coded alpha-beta pruning or mini-max algorithm.

3. Monte Carlo algorithms:

Use random algorithm and statistics to give reasonable estimates about some hard to answer questions. The tricky part is the prior as well as the posterior probabilities. In the case of Go, since the winning probability at each position is so small, DeepMind team may need lots of random samplings and customized smoothing techniques, which are limited by infrastructure like network bandwidth and computer hardware. A classical Monte Carlo algorithm consists a big loop that includes selection, expansion, evaluation and backup.

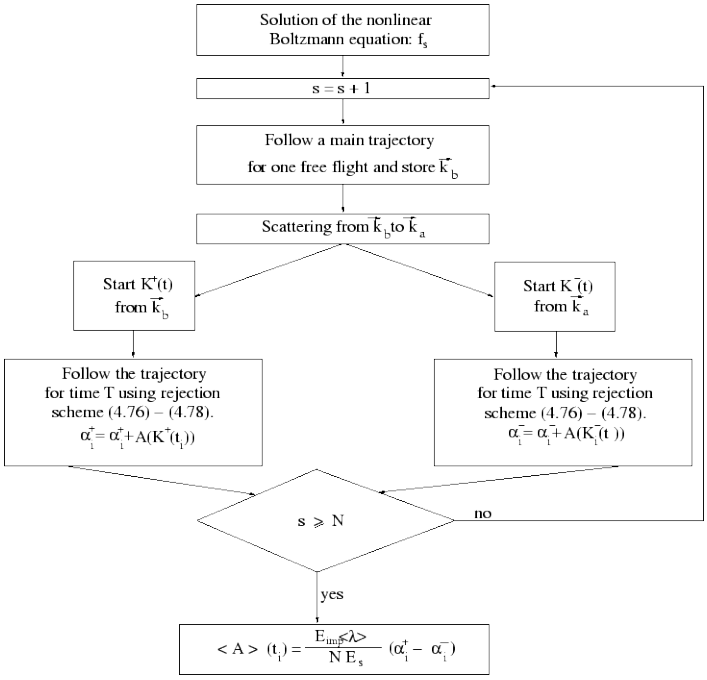


Figure 3: Monte Carlo Algorithms

4. Deep neural networks:

Interconnected web of notes, and edges that join them together. Receive a set of inputs, perform a complex calculation in the middle layer or hidden layer and then use output to solve a problem. For classification problem, we can use each layer to extract different features that interest the algorithm. (The layer inside can be non-linear.) Where GD is the key to unleash the power of DL. The popular DL system includes CNN, RNN, LSTM and GAN.

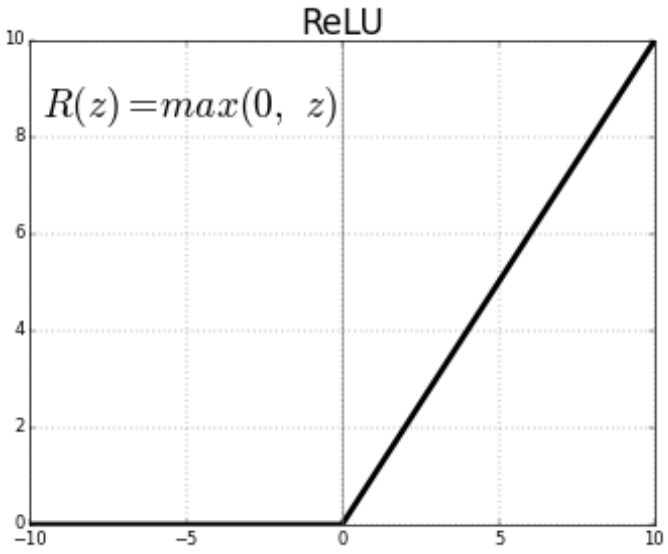


Figure 4: ReLu activation function

5. Supervised-learning:

Human Go expert plays against DeepMind AlphaGo agent. And interact with the system by telling the goodness of each game. This SL policy network can give an acceptable response in a very short period of time.

6. Reinforcement learning

AlphaGo plays against itself for self-learning and self-improvement. This RL policy network improves the SL policy network by optimizing the final outcome of games of self-play.

Google DeepMind has made exciting progress in the Go world. The ways that they structure the solution, take advantage of deep neural network, gather as well as formatting and cleaning training data, optimising agent performance at a hardware level and training AlphaGo agent in such short period of time, show that DeepMind team is really good at generalising this AI technique and is ready to apply it to other fields.

1.2 Artificial General intelligence (AGI)

Artificial General Intelligence (AGI) is the intelligence which can be as intelligent as human beings. And perform intellectual tasks as we can do. The Singularity Summit (2012) predicted this may happen around 2040 based on inputs from experts. There is no predefined definition for an Artificial General Intelligence. But in general, it should be able to learn, represent knowledge, plan, make decisions under uncertainty, communicate in a natural language and use these skills towards a common goal(s) to be AI-complete.

Tests for confirming human-level AGI(wik)

The Turing Test (Turing)

A machine and a human both converse sight unseen with a second human, who must evaluate which of the two is the machine, which passes the test if it can fool the evaluator a significant fraction of the time.

The Coffee Test (Wonazik)

A machine is required to enter an average American home and figure out how to make coffee: find the coffee machine, find the coffee, add water, find a mug, and brew the coffee by pushing the proper buttons.

The Robot College Student Test (Goertzel)

A machine enrolls in a university, taking and passing the same classes that humans would, and obtaining a degree.

The Employment Test (Nilsson)

A machine works an economically important job, performing at least as well as humans in the same job.The flat-pack furniture test (*Tony Severyns*)A machine is required to unpack and assemble an item of flat-packed furniture. It has to read the instructions and assemble the item as described, correctly installing all fixtures.

1.3 AI SAAS Solutions

Automation

AI permits more capacities which may beforehand have had a manual segment to be robotized. This shows itself in different ways – common cases incorporate utilizing machine figuring out how to computerize parts of client benefit (particularly self-serve), yet we can likewise envision utilizing AI to enhance, say, the onboarding procedure of a SaaS item, or refamiliarisation after a noteworthy refresh, through computerization.

Personalisation

AI will likewise dynamically increase the personalisation segment of SaaS administrations, as well. Already, UIs have turned out to be exponentially more unpredictable as every cycle of a bit of programming packs in more capacities and menus onto the screen.

Shipping code even faster

On the engineer side, client and aggressive weights both imply that the iteration cycle for SaaS items has been crunched down from a while to minutes. With great tools like Docker, new code can be conveyed in seconds – with certainty that it will scale superbly to a huge number of clients.

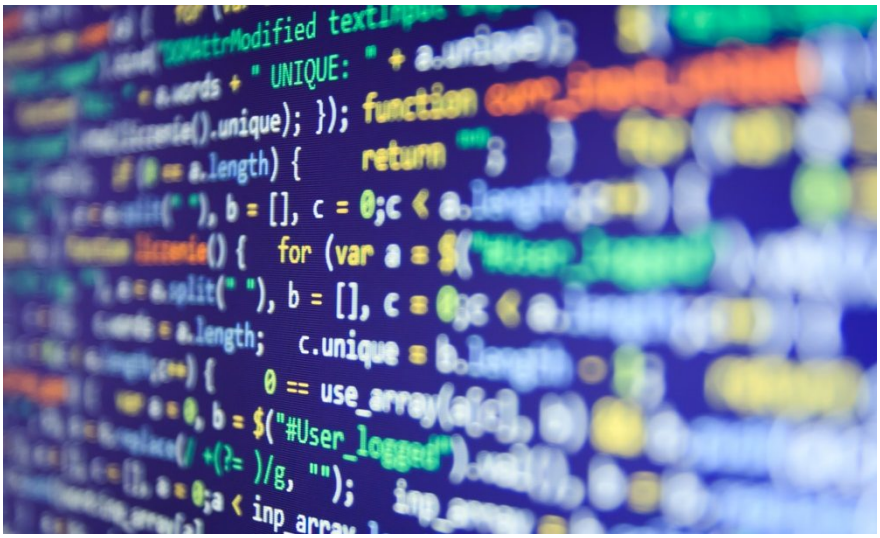


Figure 5: JavaScipt code

2. Problems in today's AI company

The major problems in the AI companies today are lack of labeled data, the data security/privacy issues, algorithms bottlenecks and high computing cost.

2.1 Data security & privacy

While a dominant part of organizations is utilizing SAAS solutions, there are still concerns, dangers, and misguided judgments with respect to their administrations. It's fundamentally in light of the fact that utilizing SAAS regularly suggests not depending on an interior IT division for information stockpiling. Also, that, fundamentally, can be an origin of stress.

The problem includes but not limited to data access risk, instability, lack of transparency, identity theft, uncertainty about data's location, commitement of long-term payment, encryption practice of the user data, access-control of user production data and updates of security standards.

Security

> [(tra)]Over four decades, Mr. Dalio, 68, has built Bridgewater, which has $160 billion in assets, into the largest hedge fund firm in the world — bigger than the next two largest hedge funds combined. He manages money for some of the largest companies, big public pensions, sovereign wealth funds and even some central banks. He has become a financier-statesman, of sorts, consulting with political leaders in China, the Middle East and elsewhere.

Figure 6: BridgeWater: Ray Dalio

> He has also built an unusual and confrontational workplace at Bridgewater, where employees hold each other to account by following a strict set of rules that he created, "Principles." ..... All of the rules celebrate what Mr. Dalio calls "radical transparency" in the workplace, and the search for the ideal employee.

Everything is recorded at BridgeWater to ensure the radical transparency. Every meeting they have is like arguing, and they speak at the speed of 2x. They are okay to let our server process and manage the production recording data. After passing their initial technology assessment and getting them as both our clients and investors, we experienced a huge security compliance concern from BridgeWater. As the largest hedge fund in the world, BridgeWater manages more than $160 assets. Under the current legal framework, if we leaked out any of our their recordings, somebody needs to go to jail(think about insider news)

For this problem and concern, we have to invest millions of dollars into our security infrastructure and consult companies like PaloAlto Networks to pass all the security compliance requirements.

But is there a better way? Can we solve this trust problem using blockchain? We think the answer is yes and that's why we build the Deuro.

Privacy

> Zoom has partnered with AISense to launch a new feature that allows their customers to get automatic transcriptions of their meetings. With *Recording Transcripts*, Zoom is taking smart meetings to the next level. This free, industry-first feature creates a searchable transcript of a Zoom cloud recording. Recording Transcripts convert all the speech from a recording into text and even identify each speaker. It allows users to search the resulting transcript for keywords, then jump to that keyword in the cloud video recording playback. Recording Transcripts save money and time by eliminating the need for participants to take notes during a meeting and by capturing information needed for training, content creation, legal depositions, sales calls, shareholder meetings, and more.

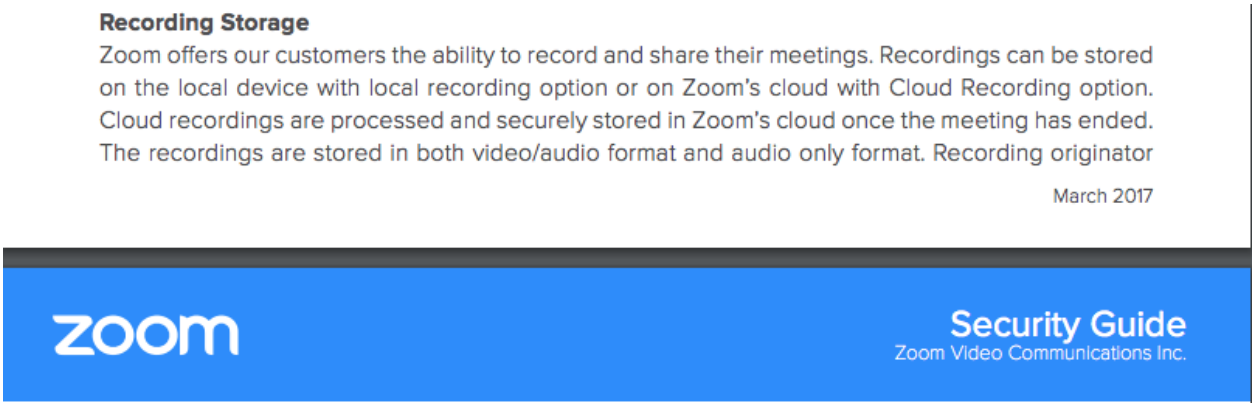


Figure 7: Zoom(Unicorn)

Most of the internet companies need to build their own server room, as the cost of cloud computing grows too quick. As Zoom stated in their user agreement: Cloud recordings are processed and securely stored in Zoom's cloud once the meeting has ended. In other words, they can't let user's recording leave their cloud because of the rules in the agreement. Is there any ways we can let SAAS companies like Zoom realize the huge benefits of opening up their production data for third parties like us to access and gain insights to improve our speech product in a secure manner using blockchain? We think the answer is yes and that's why we build the Deuro.

2.2 Lack of the training data

Data is currently the biggest bottleneck for the majority of AI companies. The ability to quickly gather huge volume, in-domain and accurately labeled data become the major differentiation factors. Today, just tech giants like Google or Facebook can gain and process a huge amount of information to enhance their AI models, while new companies can scarcely to do so because of the high cost of the data cleaning.

2.3 Algorithms

We know that the big tech giants(like Google, Microsoft, Facebook etc) dominate the AI game, but what is their secret sauce? We think it is the combination of huge in-domain labeled training data, highly advanced algorithms, and great in-house proprietary ML tools. Even though a lot of people argue that more data can constantly beat complex algorithm, but that's not always the case. In other words, beyond a certain scale, the size of dataset does not matter anymore. The essential thought is that there are two reasons (and relatively inverse) reasons a model won't perform well. Some of the linear machine learning model includes NBC, PCA, SVM, K-NN, CA, DMR)

Variance or Bias

> **Error due to Bias**: The error due to bias is taken as the difference between the expected (or average) prediction of our model and the correct value which we are trying to predict. Of course you only have one model so talking about expected or average prediction values might seem a little strange. However, imagine you could repeat the whole model building process more than once: each time you gather new data and run a new analysis creating a new model. Due to randomness in the underlying data sets, the resulting models will have a range of predictions. Bias measures how far off in general these models' predictions are from the correct value.
> **Error due to Variance**: The error due to variance is taken as the variability of a model prediction for a given data point. Again, imagine you can repeat the entire model building process multiple times. The variance is how much the predictions for a given point vary between different realizations of the model.

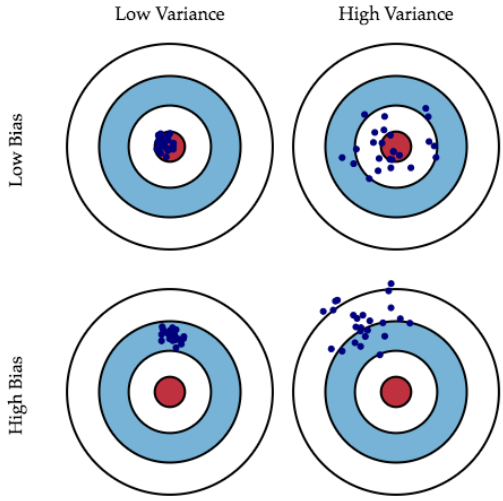


Figure 8: Graphical representation of bias and variance

Mathematical Definition

If we denote the variable we are trying to predict as Y and our covariates as X, we may assume that there is a relationship relating one to the other such as $y = f(x) + \gamma$ where the error term is normally distributed with a mean of zero like so $\gamma \sim \mathcal{N}(0, \sigma)$. We may estimate a model $\hat{f}(x)$ of $f(x)$ using linear regressions or another modeling technique. In this case, the expected squared prediction error at a point x is: $Err(x) = E[(Y - \hat{f})^2]$.

This error may then be decomposed into bias and variance components:

$Err(x) = (E[\hat{f}(x) - f(x)^2]) + E[(\hat{f}(x) - E[\hat{f}(x)])^2] + \sigma_e^2$

$Err(x) = Bias^2 + Variance + IrreducibleError$

That third term, last bumble, is the noise term in the bona fide relationship that can't on an exceptionally essential level be reduced by any model. Given the certified model and unbounded data to adjust it, we should have the ability to diminish both the inclination and vacillation terms to 0. Regardless, in a world with blemished models and restricted data, there is a tradeoff between constraining the inclination and constraining the change.

More data does not always help

In the main case, we may have a model that is excessively complicated for the amount of data we have. This circumstance, known as high variance, causes the model to overfit. We realize that we are confronting a high variance problem when the training error is much lower than the test error. High variance problems can be solved to by reducing the amount of features. These are models with many features as compared to the training examples. However, in the contrary case, we may have a model that is excessively straightforward, making it impossible to clarify the data we have.

All things considered, known as high bias, including more data won't help. See beneath a plot of a real production system at Netflix and its accuracy as we add more training datas.

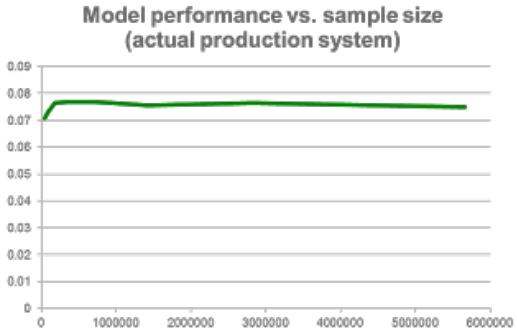


Figure 9: Netflix model

2.4 Low adoption rate

AI adoption in 2017 stays low with dominant part of significant examples of overcoming adversity coming just from the biggest tech players in the business (Google, Baidu, Apple, and so forth). McKinsey Global Institute gauges that in 2016, tech monsters put $20-30 billion into AI, while littler organizations inside and out contributed an expected $6-9 billion.

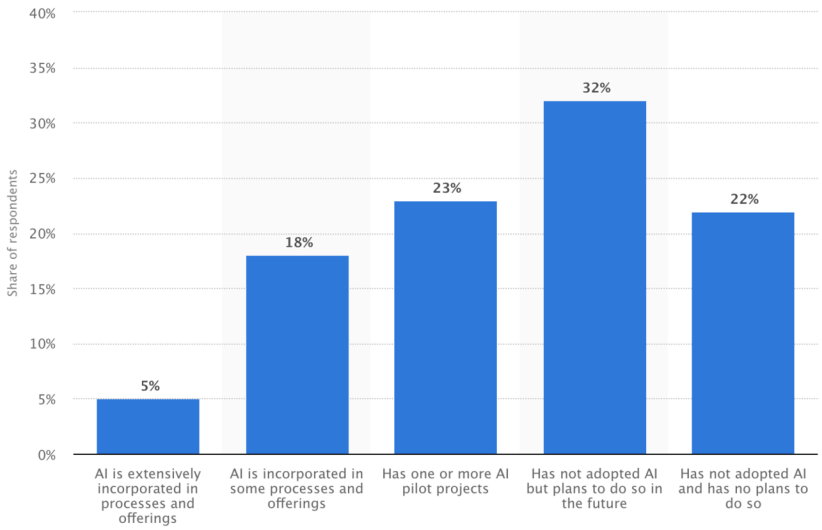


Figure 10: Adoption level of artificial intelligence (AI) in business organizations worldwide, as of 2017

This measurement introduces the level of AI adoption among organizations around the world,

starting at 2017. Five percent of the respondents said they had embraced AI broadly in their associations, in contrast with 22 percent of respondents that had neither received AI nor had plans to. Source: Statista.

3. Solutions provided by Deuro

Here we propose our innovative AI solutions based on the blockchain to solve the problems mentioned above.

gStorage: For the data security and privacy problem, we propose the gStorage solution. With AI special use case and blockchain smart contract support in mind, we introduced the distributed end-to-end encrypted solution to store/retrieve the training data, the production prediction data, and the model files.

gCrawl: For the lack of the labeled data problem, we propose the innovative gCrawl architecture. The gCrawl system is composed of fault-tolerant distributed high-performance data crawling system, stateful NoSQL database(deduplication as well as metadata storage), decentralized storage, simplified model decoding path and transfer learning for model parameters to finish the feedback loop.

gPredict: For the algorithm accuracy problem, we proposed an innovative way to train/evaluate/test the model performance on the blockchain. gPredict includes floating point mapping, hyper-parameters routing, and decoding results reducing.

gCompute: For the high hardware as well software costs when developing, testing and deploying AI systems, we introduced the gCompute module. The gCompute module is scalable, controllable, difficulty, maintainable, responsive and concurrent.

3.1. gStorage: Privacy-Preseving Decentralized Artificial Intelligence client-driven data storage system

In order to securely store the user recording in a trustless way, we have to ensure privacy and security for our customer. Here we propose an innovate way to store user recording data on the blockchain. Below is the workflow chart for how gStorage works.

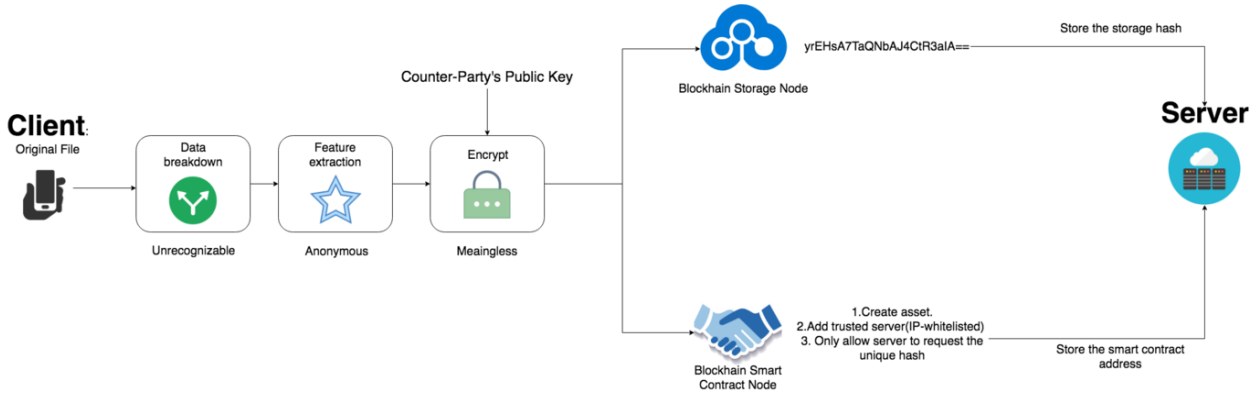


Figure 11: gStorage workflow

3.1.a. Feature extraction (Data preprocessing)

Here we showed the common features used in both speech recognition and image recognition. Audio Features: MFCC, i-vector. Image features: edges, corners/interest points, blobs/region of interest points, ridges.

MFCC

Mel-Frequency Cepstral Coefficients. MFCC are mostly used features in state-of-art speech recognition system.

```
class MFCC(object):
    def __init__(self, nfilt=40, ncep=13,
                 lowerf=133.3333, upperf=6855.4976, alpha=0.97,
                 samprate=16000, frate=100, wlen=0.0256,
                 nfft=512):
        # Store parameters
        self.lowerf = lowerf
        self.upperf = upperf
        self.nfft = nfft
        self.ncep = ncep
        self.nfilt = nfilt
```

```
self.frate = frate
self.fshift = float(samprate) / frate

# Build Hamming window
self.wlen = int(wlen * samprate)
self.win = numpy.hamming(self.wlen)

# Prior sample for pre-emphasis
self.prior = 0
self.alpha = alpha

# Build mel filter matrix
self.filters = numpy.zeros((nfft/2+1,nfilt), 'd')
dfreq = float(samprate) / nfft
if upperf > samprate/2:
   raise(Exception,
         "Upper frequency %f exceeds Nyquist %f" % (upperf, samprate/2))
melmax = mel(upperf)
melmin = mel(lowerf)
dmelbw = (melmax - melmin) / (nfilt + 1)
# Filter edges, in Hz
filt_edge = melinv(melmin + dmelbw * numpy.arange(nfilt + 2, dtype='d'))

for whichfilt in range(0, nfilt):
   # Filter triangles, in DFT points
   leftfr = round(filt_edge[whichfilt] / dfreq)
   centerfr = round(filt_edge[whichfilt + 1] / dfreq)
   rightfr = round(filt_edge[whichfilt + 2] / dfreq)
   # For some reason this is calculated in Hz, though I think
   # it doesn't really matter
```

```
        fwidth = (rightfr - leftfr) * dfreq
        height = 2. / fwidth

        if centerfr != leftfr:
            leftslope = height / (centerfr - leftfr)
        else:
            leftslope = 0
        freq = leftfr + 1
        while freq < centerfr:
            self.filters[freq,whichfilt] = (freq - leftfr) * leftslope
            freq = freq + 1
        if freq == centerfr: # This is always true
            self.filters[freq,whichfilt] = height
            freq = freq + 1
        if centerfr != rightfr:
            rightslope = height / (centerfr - rightfr)
        while freq < rightfr:
            self.filters[freq,whichfilt] = (freq - rightfr) * rightslope
            freq = freq + 1

    # Build DCT matrix
    self.s2dct = s2dctmat(nfilt, ncep, 1./nfilt)
    self.dct = dctmat(nfilt, ncep, numpy.pi/nfilt)

def sig2s2mfc(self, sig):
    nfr = int(len(sig) / self.fshift + 1)
    mfcc = numpy.zeros((nfr, self.ncep), 'd')
    fr = 0
    while fr < nfr:
```

```
        start = round(fr * self.fshift)
        end = min(len(sig), start + self.wlen)
        frame = sig[start:end]
        if len(frame) < self.wlen:
            frame = numpy.resize(frame,self.wlen)
            frame[self.wlen:] = 0
        mfcc[fr] = self.frame2s2mfc(frame)
        fr = fr + 1
    return mfcc

def sig2logspec(self, sig):
    nfr = int(len(sig) / self.fshift + 1)
    mfcc = numpy.zeros((nfr, self.nfilt), 'd')
    fr = 0
    while fr < nfr:
        start = round(fr * self.fshift)
        end = min(len(sig), start + self.wlen)
        frame = sig[start:end]
        if len(frame) < self.wlen:
            frame = numpy.resize(frame,self.wlen)
            frame[self.wlen:] = 0
        mfcc[fr] = self.frame2logspec(frame)
        fr = fr + 1
    return mfcc

def pre_emphasis(self, frame):
    # FIXME: Do this with matrix multiplication
    outfr = numpy.empty(len(frame), 'd')
    outfr[0] = frame[0] - self.alpha * self.prior
```

```
        for i in range(1,len(frame)):
            outfr[i] = frame[i] - self.alpha * frame[i-1]
        self.prior = frame[-1]
        return outfr

    def frame2logspec(self, frame):
        frame = self.pre_emphasis(frame) * self.win
        fft = numpy.fft.rfft(frame, self.nfft)
        # Square of absolute value
        power = fft.real * fft.real + fft.imag * fft.imag
        return numpy.log(numpy.dot(power, self.filters).clip(1e-5,numpy.inf))

    def frame2s2mfc(self, frame):
        logspec = self.frame2logspec(frame)
        return numpy.dot(logspec, self.s2dct.T) / self.nfilt

def s2dctmat(nfilt,ncep,freqstep):
    """Return the 'legacy' not-quite-DCT matrix used by Sphinx"""
    melcos = numpy.empty((ncep, nfilt), 'double')
    for i in range(0,ncep):
        freq = numpy.pi * float(i) / nfilt
        melcos[i] = numpy.cos(freq * numpy.arange(0.5, float(nfilt)+0.5, 1.0, 'double'))
    melcos[:,0] = melcos[:,0] * 0.5
    return melcos

def logspec2s2mfc(logspec, ncep=13):
    """Convert log-power-spectrum bins to MFCC using the 'legacy'
    Sphinx transform"""
    nframes, nfilt = logspec.shape
```

```
    melcos = s2dctmat(nfilt, ncep, 1./nfilt)
    return numpy.dot(logspec, melcos.T) / nfilt

def dctmat(N,K,freqstep,orthogonalize=True):
    """Return the orthogonal DCT-II/DCT-III matrix of size NxK.
    For computing or inverting MFCCs, N is the number of
    log-power-spectrum bins while K is the number of cepstra."""
    cosmat = numpy.zeros((N, K), 'double')
    for n in range(0,N):
        for k in range(0, K):
            cosmat[n,k] = numpy.cos(freqstep * (n + 0.5) * k)
    if orthogonalize:
        cosmat[:,0] = cosmat[:,0] * 1./numpy.sqrt(2)
    return cosmat

def dct(input, K=13):
    """Convert log-power-spectrum to MFCC using the orthogonal DCT-II"""
    nframes, N = input.shape
    freqstep = numpy.pi / N
    cosmat = dctmat(N,K,freqstep)
    return numpy.dot(input, cosmat) * numpy.sqrt(2.0 / N)

def dct2(input, K=13):
    """Convert log-power-spectrum to MFCC using the normalized DCT-II"""
    nframes, N = input.shape
    freqstep = numpy.pi / N
    cosmat = dctmat(N,K,freqstep,False)
    return numpy.dot(input, cosmat) * (2.0 / N)

def idct(input, K=40):
```

```
    """Convert MFCC to log-power-spectrum using the orthogonal DCT-III"""
    nframes, N = input.shape
    freqstep = numpy.pi / K
    cosmat = dctmat(K,N,freqstep).T
    return numpy.dot(input, cosmat) * numpy.sqrt(2.0 / K)

def dct3(input, K=40):
    """Convert MFCC to log-power-spectrum using the unnormalized DCT-III"""
    nframes, N = input.shape
    freqstep = numpy.pi / K
    cosmat = dctmat(K,N,freqstep,False)
    cosmat[:,0] = cosmat[:,0] * 0.5
    return numpy.dot(input, cosmat.T)
```

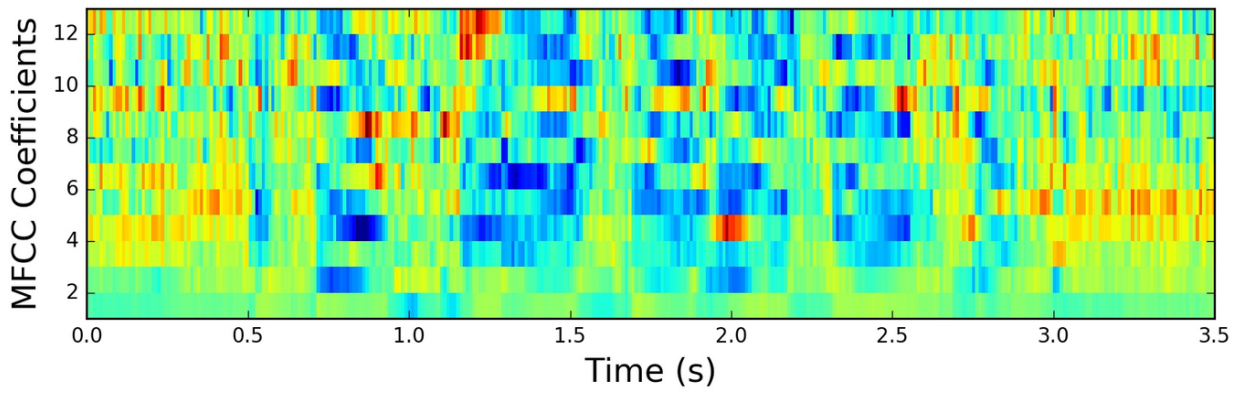


Figure 12: MFCC visuliazatio

i-Vector

An i-vector system uses a set of low-dimensional total variability factors (w) to represent each conversation side. Each factor controls an eigen-dimension of the total variability matrix (T), and are known as the i-vectors.

s(Conversationside supervector) = m + Tw(Total-variability matrix and i-vector)

1. To train T, run exact training procedure used to train V, but treat all conversation sides of all training speakers as belonging to different speakers

2. Given T, obtain i-vectors (w) for each conversation side
3. For channel compensation of i-vectors, perform LDA then WCCN (techniques empirically determined to perform well) on i-vectors. Denote channel-compensated i-vectors as ω.
4. Perform cosine distance scoring (CDS) on channel-compensated i-vectorsω for a pair of conversation sides

$$score(\omega_1, \omega_2) = \frac{\omega_1^* * \omega_2}{\| \omega_1 \| * \| \omega_2 \|} \quad = \cos(\theta_{\omega_1, \omega_2})$$

Figure 13: cosine distance scoring (CDS)

Edges

Edges are focuses where there is a limit (or an edge) between two picture locales. As a rule, an edge can be of relatively discretionary shape, and may incorporate intersections. By and by, edges are generally characterized as sets of focuses in the picture which have a solid angle extent. Moreover, some regular calculations will then chain high inclination directs together toward frame a more total depiction of an edge. These calculations more often than not put a few limitations on the properties of an edge, for example, shape, smoothness, and inclination esteem. Locally, edges have a one-dimensional structure.

Corners / interest points

The terms corners and intrigue focuses are utilized to some degree conversely and allude to point-like highlights in a picture, which have a neighborhood two dimensional structure. The name "Corner" emerged since early calculations initially performed edge discovery, and afterward broke down the edges to discover quick alters in course (corners). These calculations were then grown with the goal that express edge identification was never again required, for example by searching for abnormal amounts of shape in the picture inclination. It was then seen that the purported corners were likewise being recognized on parts of the picture which were not corners in the conventional sense (for example a little brilliant spot on a dim foundation might be identified). These focuses are every now and again known as intrigue focuses, however the expression "corner" is utilized by custom.

Blobs / regions of interest points

Blobs give a correlative portrayal of picture structures as far as locales, instead of corners that are more point-like. By and by, blob descriptors may regularly contain a favored point (a neighborhood most extreme of an administrator reaction or a focal point of gravity) which implies that numerous blob identifiers may likewise be viewed as intrigue point administrators. Blob finders can distinguish territories in a picture which are too smooth to be in any way recognized by a corner indicator.

Think about contracting a picture and afterward performing corner identification. The finder will react to focuses which are sharp in the contracted picture, yet might be smooth in the first picture. It is now that the distinction between a corner locator and a blob identifier turns out to be to some degree dubious. To a vast degree, this qualification can be cured by including a proper thought of scale. In any case, because of their reaction properties to various sorts of picture structures at various scales, the LoG and DoH blob finders are additionally said in the article on corner discovery.

Ridges

For stretched items, the thought of edges is a characteristic apparatus. An edge descriptor processed from a dim level picture can be viewed as a speculation of an average pivot. From a useful perspective, an edge can be thought of as a one-dimensional bend that speaks to a hub of symmetry, and what's more has a quality of neighborhood edge width related with each edge point. Tragically, be that as it may, it is algorithmically harder to separate edge highlights from general classes of dark level pictures than edge-, corner-or blob highlights. By and by, edge descriptors are as often as possible utilized for street extraction in flying pictures and for removing veins in restorative pictures—see edge identification.

3.1.b. Decentralized pseudo-anonymous multi-party key encryption and decryption

Here we will discuss the core of the gStorage. The decentralized pseudo-anonymous multi-party key encryption and decryption algorithms.

1. It's decentralized in a sense that we will use blockchain as the backend and broadcast the signed transaction to the whole network without relaying on a trusted centralized party, which may cause single point of failure.
2. Then, It's pseudo-anonymous in a sense that we only protect the extracted feature and expose the encrypted hash to the whole network. Just like how Bitcoin works(user signed the transaction locally using their private key, and then broadcast the transaction hash to the whole network to verify).
3. It involves multi-party in a sense that we will need to know the counter-party public key ahead of time so that the counter-party(server, IOT device, website, IP address, smart contract) can decrypt the encrypted hash later using their own private key.

Protect the extracted feature

Once we have the extracted feature representation for AI model, we will need to protect it. We protect it by encrypt the extracted feature into cipher text using counter-party's public key. Here we will use the industry standard public key encryption algorithm.

1. each party has a PAIR (K, K-1) of keys: K is the public key and K-1 is the secret key, such that DK-1[EK[M]] = M

2. Knowing the public-key and the cipher, it is computationally infeasible to compute the private key Public-key crypto system is thus known to be asymmetric crypto systems
3. The public-key K may be made publicly available, e.g., in a publicly available directory
4. Many can encrypt, only one can decrypt

Key generation: Select 2 large prime numbers of about the same size, p and q. Compute n = pq, and $\Phi(n) = (q-1)(p-1)$ Select a random integer e, $1 < e < \Phi(n)$, s.t. $gcd(e, \Phi(n)) = 1$. Compute d, $1< d< \Phi(n)$ s.t. $ed =1 \bmod \Phi(n)$. Public key: (e, n) Secret key: d

Encryption: Given a message M, $0 < M < n$ use public key (e, n) compute$C=M^e \bmod n$

Decryption: $M \epsilon Z_n - \{0\}$ $C \epsilon Z_n - \{0\}$.Given a ciphertext C, use private key (d).Compute$C^d \bmod n=(M^e \bmod n)^d \bmod n=M^{ed} \bmod n = M$

Broadcast the encrypted hash

Once we have the extracted feature encrypted, we will put into into the p2p decentralized database. Then register the unique hash of the encrypted features into the blockchain, and broadcast this encrypted hash into the whole network. The miner will know the encrypted hash of the features, can find all features belong to a single user, but they can not decrypt it. Just like all bitcoin miners can find transaction belong to a single user, but they can not transfer the money into miner's own account, because the miners do not have the private key of the interested account.

Transfer the ownership of asset to counter-party

Once we register the feature asset on the blockchain, server can request the unique hashes of features belonged to a particular user. Then server can retrieve the encrypted cipher text from the p2p decentralized database, and decrypt the cipher text for the features using its own private key.

———-

Solidity ETH access_control_template()

———-

```
pragma solidity ^0.4.23;


contract gStorageAccessControl {
    address admin;
    address server_address;


    modifier onlyByServer { if (msg.sender != server_address) throw; _ }
    modifier onlyByAdmin { if (msg.sender != admin) throw; _ }
     address[] public features;
```

```
    event FindTheUniqueHash(address[] unique_hash);


    function gStorageAccessControl(address server_address_ ) {
        admin = msg.sender;
        server_address = server_address_;
    }
    function registerFeatures(address unique_hash) onlyByAdmin{
        features.push(unique_hash);
    }
    function transfer() onlyByServer {
        //Retrive data from the storage
        FindTheUniqueHash(features);
    }
}
```

Privacy-preserving pseudo-anonymous data sharing

In this way we protect the privacy of the customer by applying the feature extraction. We protect the security of the data using industry standard public key encryption. We also enforce the access control by passing the countery-party's public key into the smart contract.

3.1.c. AI Data(Conversation, Images, Sentences) as an asset

Audio data comes in a sometimes bewildering variety of forms. The number of fundamental ways in which sound can be represented is actually fairly small. The variety of audio file types is due to the fact that there are quite a few approaches to compressing audio data and a number of different ways of packaging the data. We can abstract an conversation recording as an asset on the blockchain. An asset can also be a state machine where the state transition is represented in the metadata. Each time the machine changes its state, a transaction is triggered to update the metadata to the new state (possibility to listen to it with the WebSocket or HTTP rest way). In our case, we can implement the timestamped_private_key in this asset state machine transformation way.

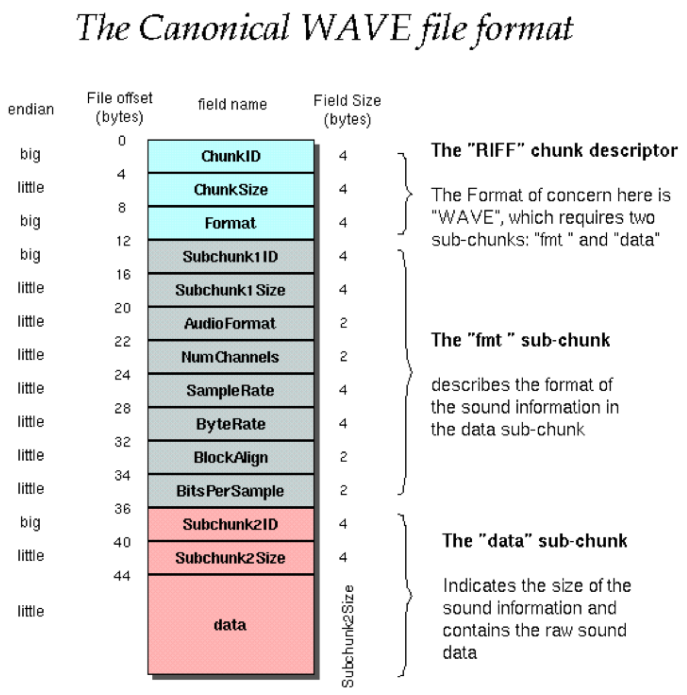


Figure 14: The raw audio file Wav representation

3.1.d. Enterprise facing solution

For the enterprise customer, we will deploy a out-of-the box solution for them. And the customer will need to purchase our token in order to receive our service. Once we finish our services, we will burn their purchased token.

3.1.e. Consumer facing solution

For the consumer facing application, developers can quickly have gStorage working on their app by simply installing our SDK. We will implement a freemium model, once they hit certain usage, we will ask them to purchase certain amount of our token as a reserve. And we will burn their token per their API call.

3.1.f. Modeling set repository

With the above architecture in mind, we can store all the training data easily into the blockchian peer-to-peer database. For audio training data, we can store raw uncompressed or lossless compressed wav file and corresponding proud truth file.

3.1.g. Predicting data storage

For all the production data generated by the AI dApps, it can also go through our existing gStorage pipeline. And allow user to securely retrieve as well as send the files.

3.1.h. Models hub

For the models generated by popular deep learning frameworks like Tensor-flow, Keras, Pytorch, Kaldi, Caffee, CNTK etc, we can also use the above method to securely put and get them from the blockchain database.

3.2. gCrawl: Distributed high-performance algo-generated training data solution

We exhibit a brand new machine learning architecture called "gCrawl" for utilizing unlabeled information in supervised classification tasks. We don't assume that the unlabeled data takes after a similar class names or generative description as the marked data. Hence, we might want to utilize an extensive number of unlabeled pictures (or sound examples, or content records) arbitrarily downloaded from the Internet to enhance execution on a given picture (or sound, or content) order task. We describe an approach to gCrawl that uses simplified decoding path and transfer learning to iterate models from unlabeled data.

Below is an architecture of a gCrawl computing cluster and it is horizontal scalable. In practice, we have 1 masterNode and 11 slaveNode per each cluster , which is capable of processing **1 million crawling tasks a day!** By preprocessing all the tasks in the NoSQL database, we were able to deploy 12 such clusters, resulting **12 million crawling tasks a day!**

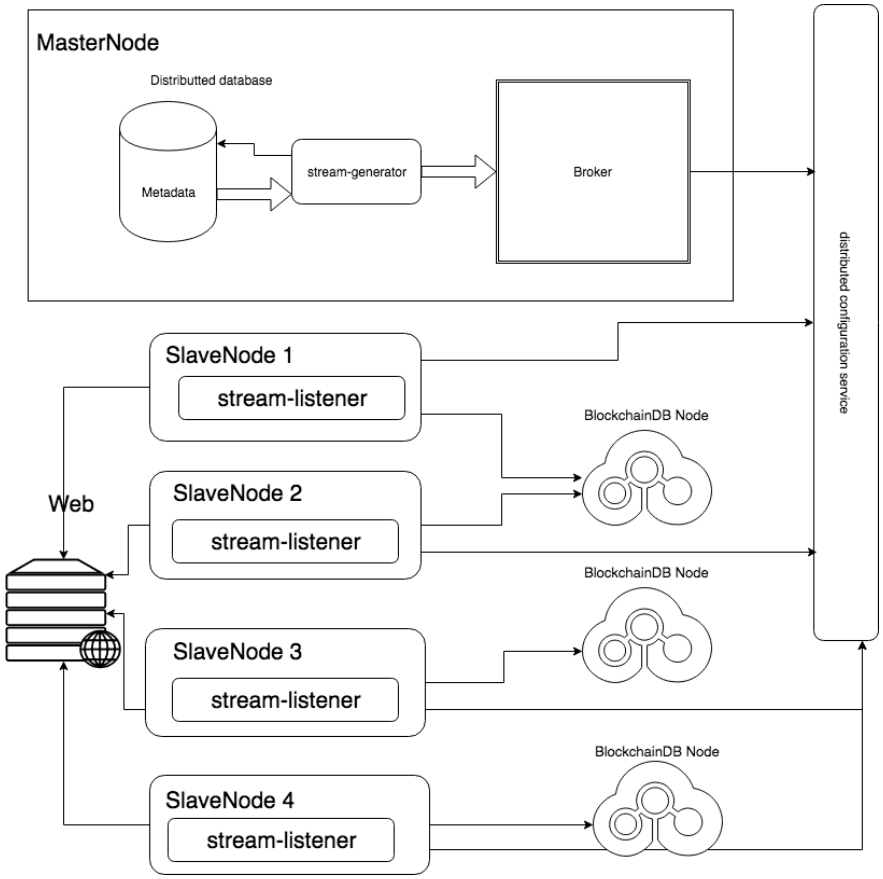


Figure 15: gCrawl crawling system

3.2.a. Crawling task as a transaction

We will abstract crawling task as a transaction on the blockchain ledger.

```
contract CrawlingTask {

    struct CrawlingMachine {
        bool crawled;
        address location;
    }

    struct Task {
        string url;
    }

    address submiter;
```

```
    mapping(address => CrawlingMachine) machines;
    Task[] tasks;

    function crawl() {
        //The crawling business logic
    }
}
```

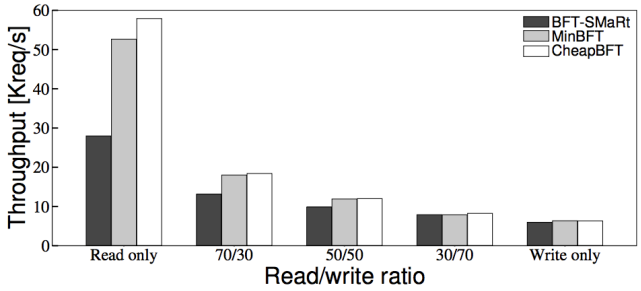
3.2.b. Crawled data as an asset

We can abstract crawled data as an asset on the blockchain

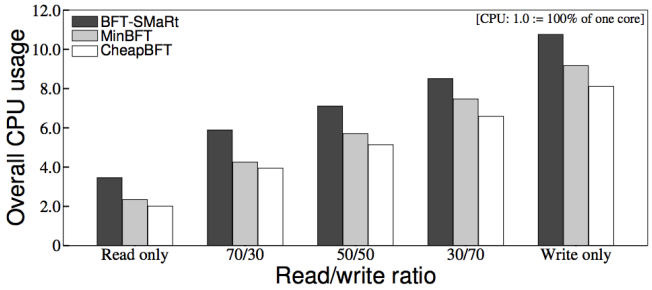
```
contract CrawledData {
    string[] metadata;
    string transcript;
    address audio_url;
}
```

3.2.c. Fault-tolerant distributed high performance crawling

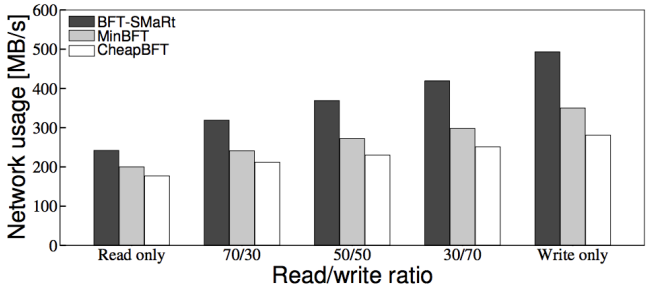
ZooKeeper is a crash-tolerant coordination service utilized as a part of large-scale distributed frameworks for essential tasks like pioneer decision, synchronization, and failure recognition. This section displays an assessment of a ZooKeeper-like BFT benefit that depend on BFT-SMaRt, MinBFT, and CheapBFT for request dissemination, separately. ZooKeeper enables customers to store and recover (generally little) chunks of data in information hubs, which are overseen in a various leveled tree structure. We assess the three executions for various blends of read and compose tasks. In all cases, 1,000 customers more than once get to various information hubs, perusing and composing information pieces of arbitrary sizes between one byte and two kilobytes. The outcomes demonstrate that with the execution stage(I. e., the ZooKeeper application) performing genuine work (and not simply sending answers), the effect of the assention convention on framework execution is diminished. In result, each of the three ZooKeeper usage give comparative throughput to compose overwhelming workloads. Be that as it may, the asset impressions fundamentally contrast between variations: in contrast with the MinBFT-based ZooKeeper, the reproductions in the CheapBFTbased variation spare 7-12% CPU and send 12-20% less information over the system. Contrasted with the BFT-SMaRt execution, the asset funds of the CheapBFT-based ZooKeeper signify 23-42% (CPU) and 27-43% (organize).



(a) Realized throughput for 1,000 clients.



(b) CPU usage per 10 Kreq/s normalized by throughput.



(c) Network transfer volume per 10 Kreq/s normalized by throughput.

Figure 16: Execution and resource utilization results for different BFT variants of our ZooKeeper service for workloads comprising distinctive mixes of read and write operations.

3.2.d. RAFT consensus algorithm

Raft is the consensus algorithm used in many NoSQL databases. We use NoSQL database to store all the metadata information of the crawling task. Below is the pseudo code for Leader Election for the RAFT consensus algorithm

```
# follower not receiving heartbeats from leader
if election_timeout?
  increment_term()
  change_state_to_candidate()
  vote_myself()
  send_request_vote_to_all_partners()
```

```
  while candidate?
    wait_feedback(random_time)
    process_feedback()
    foreach entry in append_entries_received()
      if self.term() < entry.term()
        mark_current_election_as_lost()
      end
    end
    if i_won_election?
      change_state_to_leader()
      start_heartbeat_signals()
    else if partner_won_election?
      change_state_to_follower()
    else if no_one_won_election?
      increment_term()
    end
  end
end

# continue as leader or follower...
```

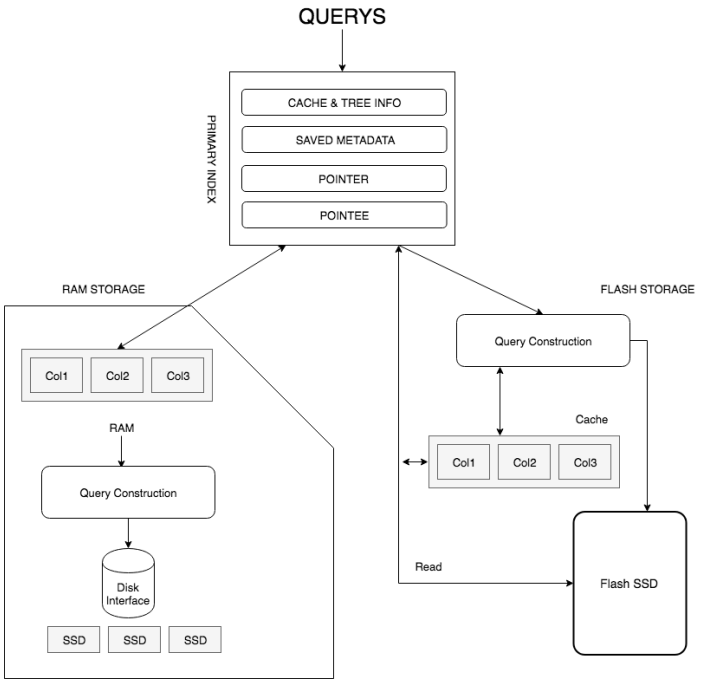


Figure 17: Distributted database storage

3.2.e. Decentralized data storage

We will use different decentralized data blockchain-based storage solution to store the crawled data. We will explain more about the storage in the section 3.1

3.2.f. DFSM decoding path

When we prepare the supervised learning data, we simplify our model decoding path and ask the model DFSM a simpler question. Instead of asking the model to predict what is the phenom of this utterance, we also present the hypothesis text and ask the model to tell us wether we can align the text into the audio segment or not. Usually the Language model of speech recognition system is unsupervised-learning,

Command to generate the full Kaldi decoding graph.

```
make_lexicon_fst.pl lexicon_disambig.txt 0.5 sil '#'$ndisambig | \
   fstcompile --isymbols=lexgraphs/phones_disambig.txt \
    --osymbols=lmgraphs/words.txt \
    --keep_isymbols=false --keep_osymbols=false |\
   fstaddselfloops  $phone_disambig_symbol $word_disambig_symbol | \
   fstarcsort --sort_type=olabel \
   > lexgraphs/L_disambig.fst
```

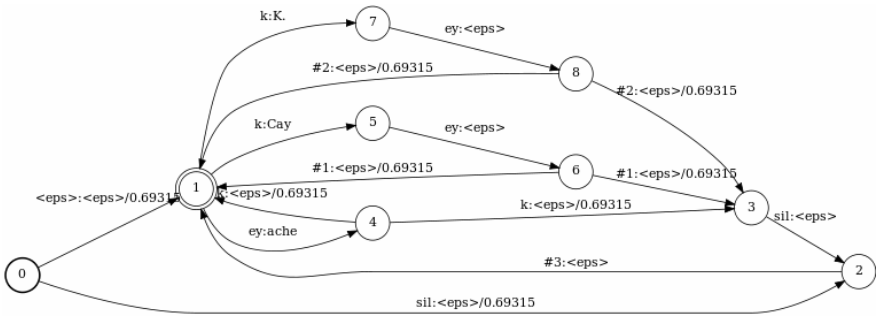


Figure 18: Lexicon full decoding graph

Once acoustic models have been created, Kaldi can also perform force-align on audio accompanied by a word-level transcript

1. Create data/train files according to the new transcript and audio
2. Extract MFCC features
3. Align data
4. Obtain CTM output from alignment files
5. Concatenate CTM files
6. Convert time marks and phone IDs
7. Split final_ali.txt by file

8. Create word alignments from phone endings
9. Append header to each of the text files for Praat
10. Make Praat TextGrids of phone alignments from .txt files
11. Make Praat TextGrids for word alignments from word_alignment.txt
12. Stack phone and word TextGrids

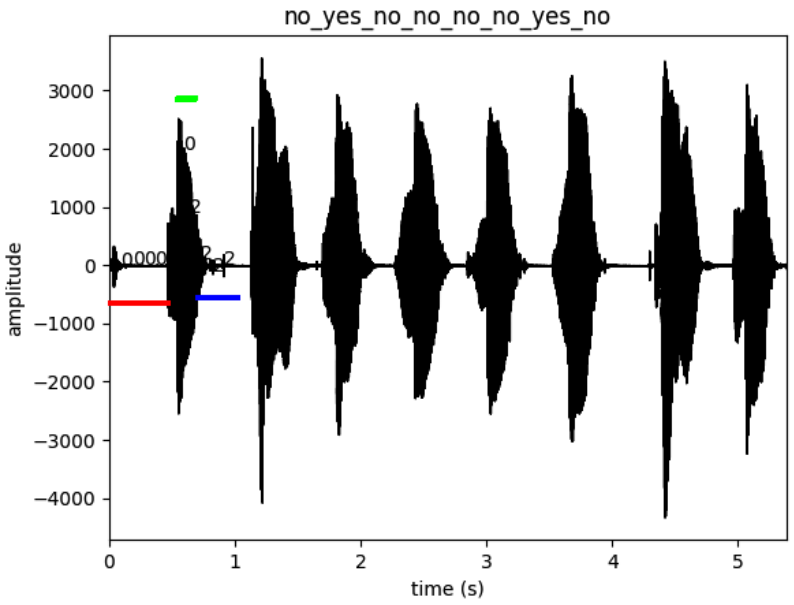


Figure 19: Simplified audio force-align flow

3.2.g. Transfer learning to update the model parameters

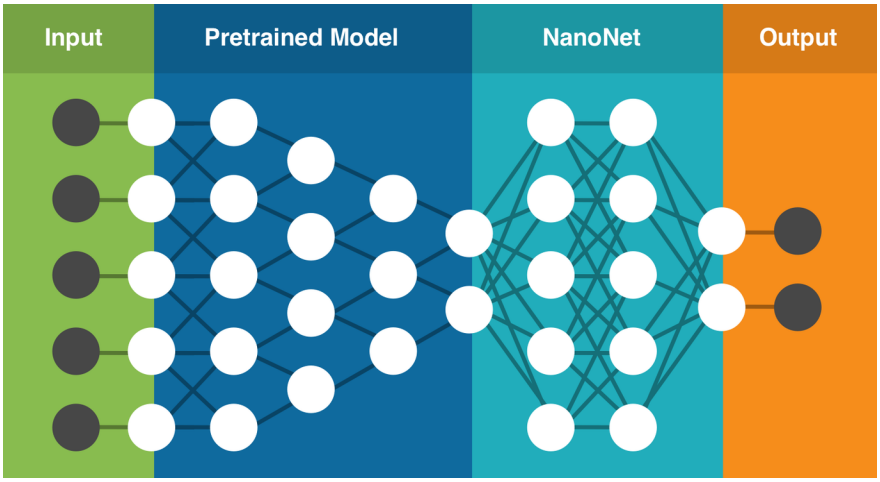


Figure 20: Transfer Learning diagram

(Transfer Learning) Given a source domain D_S and learning task T_S, a target domain D_T and learning task T_T , transfer learning aims to help improve the learning of the target predictive function f_T $(\cdot)$ in D_T using the knowledge in D_S and T_S, where D_S != D_T , or T_S != T_T . In the above definition, a domain is a pair D = $\{$X , P(X)$\}$. Thus the condition D_S != D_T implies that either X_S != X_T or P_S(X) != P_T (X).

3.3. gPredict: Decentralized gradient learning framework

In order to improve the performance of AI algorithm, we need to increase the precision/recall, decrease the MSE. In this section, we present a set of ways to reduce the AI model training turnaround time on the blockchain.

Estimator: is just a simple function to model some meaningful characteristics of data sample. For example, $Y = g(S)$, $S = (x(1), , , , x(m))$, where x(i) is a random variable drawn from distribution D, i.e. $x(i) \sim D$. **Estimator Bias** measures how good our estimator is in estimating the real number. $Bias(Y_S, Y) = E_{S\sim D^m}[Y_S] - Y$ **Estimator Variance** measure how 'jumpy' our estimator is to sampling. $Var(Y) = Var_{S\sim D^m}[Y]$. **Bias-variance decomposition for estimators:** combines those two properties in one formula. $MSE = E[(Y_S - Y)^2] = Bias^2(Y_S, Y) + Var(Y_S)$ where the expectations are taken with respect to **S** random variable.

Formal proof for the above conclusion:

$$E[(Y_S - Y)^2] = E[Y_S^2] + Y^2 - 2E[Y_S]Y$$

$$Bias^2(Y_S, Y) = (E[Y_S] - Y)^2$$

$$= E^2[Y_S] + Y - 2E[Y_S]Y$$

$$Var(Y_S) = E[Y_S^2] - E^2[Y_S]$$

3.3.a. Training task as a transaction

```
contract TrainingTask {
    TrainedModel previousModel;
    function training();
    TrainedModel trainedModel;
}
```

3.3.b. Trained model as an asset

```
contract TrainedModel {
    string[][] configuration;
    double[] weights;
    state[] optimizerState;
}
```

3.3.c. Floating Point mapping

According to our study of the limited number precision within the low-precision fixed-point context, we find that by controlling the rounding scheme using only 16-bit representation when using stochastic gradient descent, we can train the model even faster without hurting any accuracy.

Limited Precision Arithmetic: Standard executions of back-propagation powered deep learning system commonly utilize 32-bit floating-point representation of real numbers for data storage and control. We abstract this problem by representing [IN.FR], where IN and FR relate to the integer and fractional part of the number, individually. The number of number bits(NB) plus the quantity of fractional bits(FB) yields the aggregate number of bits used to represent the whole number. The sum of NB + FB is referred as the word length WL.

Stochastic rounding: Given a number x and the target fixed-point representation <NB, FB>. We define $\lfloor x \rfloor$ as the largest integer multiple of $\theta \left(= 2^{-FB}\right)$ <= x. Stochastic rounding is an unbiased rounding scheme and possess the desirable property that the expected rounding error is zero, i.e. E(Round(x, <NB,FB>)) = x

So, P(rounding x to $\lfloor x \rfloor$) is relative to the approximation of x to $\lfloor x \rfloor$:

$$Round(x, < NB, FB >) = \begin{cases} \lfloor x \rfloor, & amp; \text{if } 1 - \dfrac{x - \lfloor x \rfloor}{\theta}, \\ \lfloor x \rfloor + \theta, & amp; \dfrac{x - \lfloor x \rfloor}{\theta}. \end{cases}$$

Then,

$$Convert(x, < NB, FB >) = \begin{cases} -2^{NB-1}, & amp; \text{if } x <= -2^{NB-1}, \\ 2^{NB-1} - 2^{-FB}, & amp; \text{if } x >= 2^{NB-1} - 2^{-FB}. \\ Round(x, < NB, FB >), & amp; otherwise. \end{cases}$$

Multiply and accumulate operation: Consider two dimensional vectors **d** and **h** such that every segment is represented in the settled fixed-point <NB,FB>, and define **e** = d.h as the inner product of d and h. Then we can split the computation into the following steps:

$z = \sum_{i=1}^{d} a_i b_i$

$c_0 = Convert(z, < NB, FB >)$

3.3.d. Hyper-Parameters routing

gPredict's activity is to locate the best estimation of a scalar-esteemed, perhaps stochastic function over an arrangement of possible arguments to that function. While numerous bundles will expect that these data sources are drawn from a vector space, gPredict is distinctive in that it urges developer to depict the search space in more detail. By giving more data about where the function

is characterized, and where the best qualities are, gPredict can search the best parameter more efficiently.

3.3.e. Results reducing

Finally with the help of floating point mapping and hyper-parameters routing, we can get the results by reducing from blockchain type of distributed system.

Reducer algorithm

```
class REDUCER
    def REDUCE(string t, pairs [(s1,c1),(s2,c2)...])
        sum <-- 0
        cnt <-- 0
        for all pair (s,c) belongs to pairs:
            sum += s
            cnt += c
        rAvg = sum / cnt
        EMIT(string t, integer rAvg)
```

3.4. gCompute: Internet-scale AI dApps solution

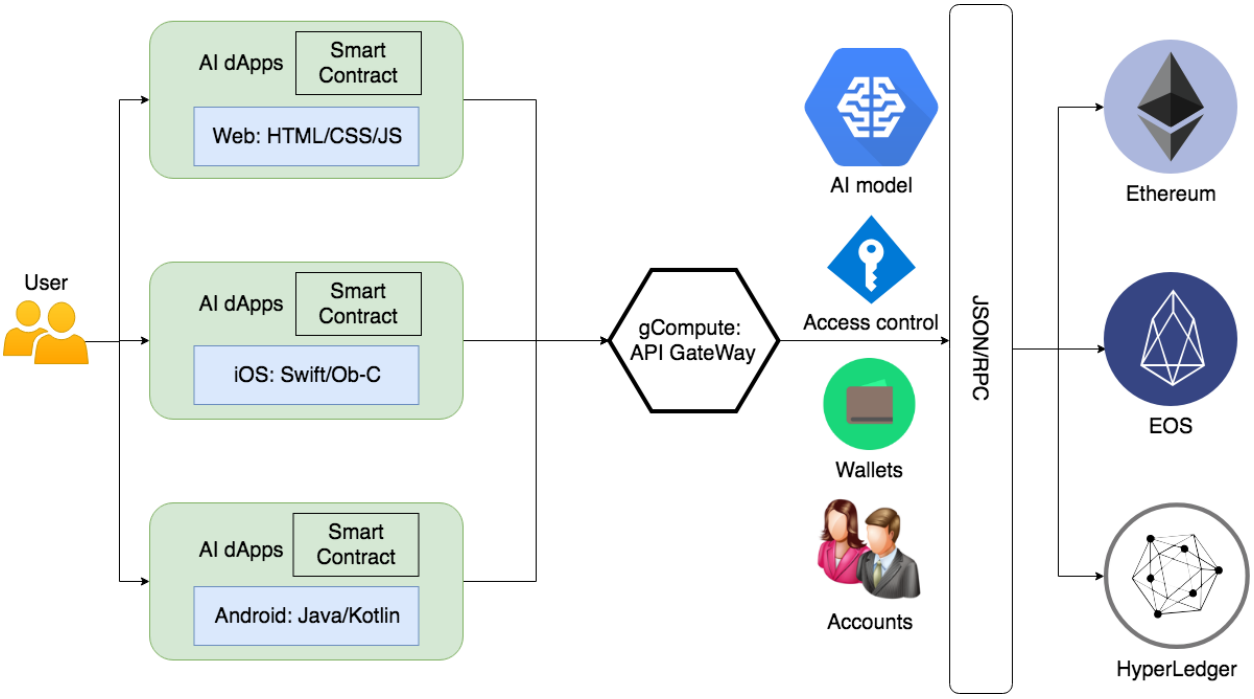


Figure 21: gCompute API GateWay

3.4.a. AI model as an asset

Similar to the our definition of TrainedModel above, we can represent an AI model as an asset on the blockchain.

3.4.b. Scalable

Technical scalability: In order to compete with traditional internet giants like Google, Facebook, Alibaba, Tencent, our underlying platform blockchain technology has to horizontally scalable. If we have a AI dApp signup and use our platform for easy integration into all the existing blockchain backends, our system is able to handle tens of millions of daily active users. For this type of functional requirement, we will deploy multiple instances, abstract the core business logic into a docker and put load balancer in front of our instances to make sure every API/JSON calls went smoothly.

Business model scalability: Traditional internet company has proven to us that the best business model is to offer free service first and then offer a variety of revenue models (freemium, premium, subscription). So our underlying blockchain technology platform should have the flexibility to let

application offer creative business revenue model. It is the free to use for users that causes the internet to gain more widespread adotion.

3.4.c. Controllable

Constant upgrades: Either big enterprises or small startups, even individual developers should have the flexibility to upgrade their application with new features constantly. Our platform will support both model software as well as smart contract upgrades. Even Google chrome, Facebook, Tencent WeChat, Apple iOS are subject to to different bugs or system failure, our platform and the underlying platform should be robust enough for bug fixes and automatic program call stack generation.

Continuous integration: When developers add new functionality to their existing softwares, they do not want to brake any old features. That is the where the continuous integration solution comes in. If certain commits break the old feature(can't pass certain testcases), the code hosting platform will reject that commit and generate a report about the error. Our system as well as the underlying blockchain platform will also support this kind of good practice.

Failure recovery: Worst case scenario: if we have network failure, system failure, underlying blockchain platform DDOS attack or any user key credential being compromised, our platform as well as underlying blockchain platform will have a failure recovery system. By versioning different release of the AI models as well as the database state, we will be able to provide this kind of features for our end user as well as developers.

3.4.d. Difficult

PoW: PoW is the most adopted consensus algorithm for the unpermissioned public blockchain. Pioneered by bitcoin, PoW is the most known consensus algorithm on the blockchain ecosystem. PoW is hard to break and requires more than 50% attack to destroy the system. We will be able to safely integrates all PoW based public blockchain and finish most of our AI dApp business logic on the corresponding blockchain.

PoS: For the unpermionsoed public blockchain, PoS/BFT-DPOS/DPOS are the new challenger. PoS is proven to be more energy friendly. Under those algorithms, users who hold tokens on the corresponding blockchain platform are able to select block producers and produce blocks. Our AI dApp platform will be able to safely communicates with those public blockchain and implements all of our business logics.

SIEVE: For the permissioned private blockchain, there are numerous innovation around the consensus algorithm. One of the most interesting one is SIEVE. In short, SIEVE expands the first PBFT calculation by including theoretical execution and check stages to: 1) identify and sift through conceivable non-deterministic requests and set up the determinism of exchanges entering the PBFT 3-stage assertion protocol, and 2) enable agreement to be keep running on yield condition of validators, notwithstanding the accord on their info state offered by Classic PBFT. SIEVE

is gotten from PBFT separately (motivated by thoughts portrayed in [Aublin et al., TOCS'15]) by reusing the PBFT see change convention to bring down its multifaceted nature and abstain from executing another accord convention sans preparation.By the nature of permissioned private blockchain, our system is able to securely connect those blockchain with access control implemented.

Quantum resistance: Post-quantum cryptography, otherwise called quantum-safe cryptography, can stand up to the assaults by quantum PCs. The improvement of such encryption innovation takes a more customary way, in view of troublesome issues in particular arithmetic fields. Through investigating and creating calculations, the post-quantum secure encryption innovation can be connected in the system, what's more, to give the most abnormal amount of information security. We will be looking forward to see exciting use case of such algorithm being put into practice on various public/private blockchain.

3.4.e. Maintainable

Extensibility: The power of modern software is built on top of each other. With vibrant support of third-party libraries and packages, developer is able to develop the software at much faster speed. We will provide a set of library dependency management tools to allow developers easily call and manage the already published smart contracts or import other contracts as a library dependency.

Readability: We will convert/parse all the transactions on the given blockchain and allow user to understand all our of systems AI dApps. User will be able to see who create the AI smart contract, the primary interactions of that smart contract, automatic smart contract abi generations, etc.

Explorbility: We should be able to explore the entire blockchain. We will provide insights into the recent mined blocks on the blockchain. We will also provide insights into any transactions in any block that has already been mined and is currently attached to the blockchain network. User can also check the history of any public address and audit the balances, transaction history etc.

3.4.f. Responsive

Fast response: A great customer experience requests instantaneous feedback within couple seconds. User will get fractured if the application requires longer time to respond and the unresponsive AI dApp results in less competitive situation than traditional solutions. Our gCompute platform will try to optimize as much as we can to reduce the application response time.

Networking optimization: We will try to minimize the communication latency within the network. Network optimization ought to have the capacity to guarantee ideal use for system resources, enhance profitability and also effectiveness for the community. Network optimization takes a look at the individual workstation up to the server and the tools and connections related with it. Our system could consist of traffic shaping, redundant data elimination, data caching and data compression and streamlining of data protocols.

3.4.g. Concurrent

Synchronous execution: If the dApp itself requires sequentially independent steps, we should support fast sequential execution as well.

Asynchronous execution: Even though computers run in sequential mode, and uses the combination of locks & context switch to implement the pseudo asynchronous execution, modern applications all support concurrent execution by default. With the full internet-scale AI dApp solution in mind, we will try to optimize our code for the asynchronous execution as we can.

3.5 Consensus algorithm

Gradient will maintain two different types of network, one using DPoS as the consensus algorithm. Another one using Vectority - Convex hull powered ASIC friendly PoW mining algorithms. gStorage, gCrawl and gCompute will rely on the DPoS networks, since those system have fairly trivial computing needs. The gPredict onchain training network needs to have fairly large amount of GPU for it to perform well, so gPredict will rely on the Vectority consensus algorithm.

3.5.a. DPoS

For high performance and functional requirements, we would like to use DPoS consensus mining algorithm for gStorage, gCrawl and gCompute systems. By downloading our mining software, users are able to write transactions into the ledger and get token as a reward.

> Delegated Proof of Stake (DPoS) is the fastest, most efficient, most decentralized, and most flexible consensus model available. DPoS leverages the power of stakeholder approval voting to resolve consensus issues in a fair and democratic way. All network parameters, from fee schedules to block intervals and transaction sizes, can be tuned via elected delegates. Deterministic selection of block producers allows transactions to be confirmed in an average of just 1 second. Perhaps most importantly, the consensus protocol is designed to protect all participants against unwanted regulatory interference.

3.5.b. Vectority - Convex hull powered ASIC friendly PoW

Cryptocurrency is being criticized for not doing meaningful computation on the network. In other words, the mining process does not generate any useful calculation to the humanity. Gradient introduces a brand new mining consensus algorithm in which miners perform the convex hull algorithm computation, which is the training hypothesis for all the deep learning frameworks. In the greater part of our ML jobs, we will likely limit the loss function. To do that, we prepare our model utilizing some iterative optimization calculation, for example: Stochastic Gradient Descent. Strict Convex function ensures global minimum and so while training our optimization algorithm reaches

there and tells you that your model is now optimized. Training deep learning model is all about those hyperparameters that our function has to be optimized over

Similar to how tensority(proposed by Bytom) works, Vectority is also a ASIC friendly PoW mining algorithm. Vectority can not only support tensor(mainly used by tensorflow) but also support the general vector used by many other deep learning frameworks.

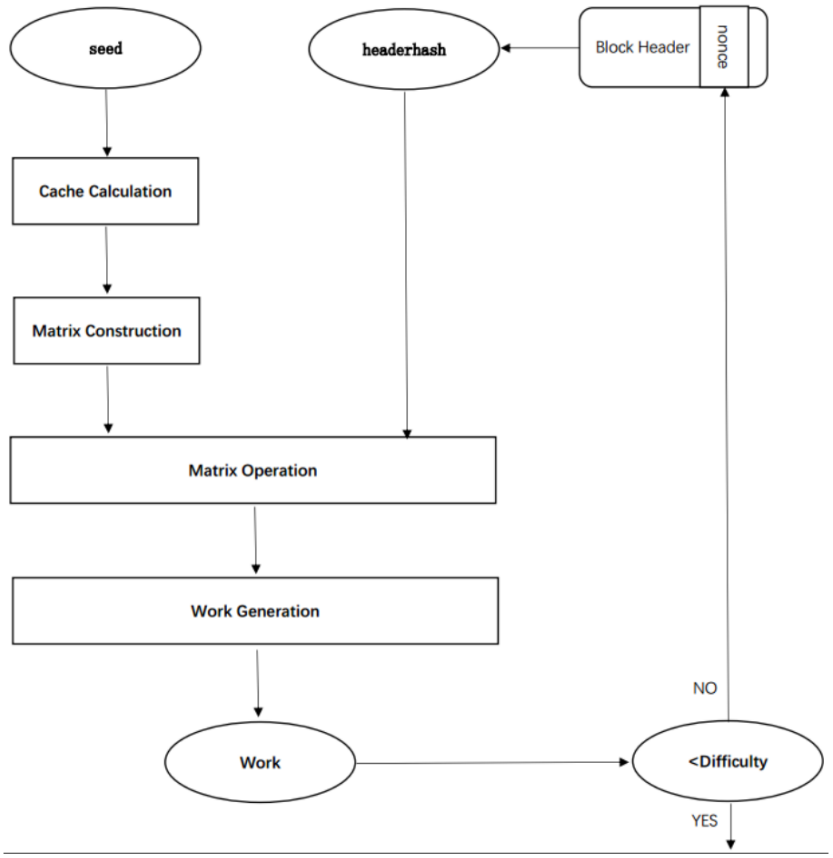


Figure 22: Cite: Tensority algorithm data flow

4. Applications and Future Work

4.1. Aviation

AI Assistants: Responding to the client request and reacting to voice orders for residential carrier flight information and ticket accessibility through connections utilizing normal dialect

Brilliant Logistics: Machine learning calculations are being connected to information to help computerize carrier activities.

Facial Recognition: Facial acknowledgment innovation is being utilized to perform client personality confirmation and to coordinate travelers to their baggage through stands

4.2. Education

Artificial intelligence can automate essential exercises in teaching, as grading. In school, evaluating homework and tests for substantial address courses can be repetitive work, notwithstanding when TAs split it between them. Indeed, even in bring down evaluations, educators regularly find that reviewing takes up a lot of time, time that could be utilized to cooperate with understudies, plan for class, or work on proficient advancement.

Education software can be adjusted to student needs.From kindergarten to graduate school, one of the key ways manmade brainpower will affect teaching is through the utilization of more prominent levels of individualized learning. A portion of this is now occurring through developing numbers of adaptive learning programs, recreations, and programming. These frameworks react to the necessities of the student, putting more prominent emphasis on specific subjects, rehashing things that understudies haven't aced, and for the most part helping understudies to work at their own particular pace, whatever that might be.

It can call attention to places where courses need to improve.Teachers may not generally know about holes in their addresses and instructive materials that can leave students confused about specific ideas. Artificial intelligence offers an approach to take care of that issue. Coursera, a gigantic open online course supplier, is as of now placing this into practice. At the point when countless are found to present the wrong response to a homework task, the framework cautions the educator and gives future understudies a redid message that offers indications to the right answer.

Students could get extra help from AI guides. While there are clearly things that human coaches can offer that machines can't, at any rate not yet, the future could see more students being mentored by guides that lone exist in ones. Some mentoring programs in view of computerized reasoning as of now exist and can help understudies through essential arithmetic, composition, and different subjects.

4.3. Finance

Portfolio Management / Algorithmic Trading / Fraud Detection: Join more available processing power, web winding up more regularly utilized, and an expanding measure of significant organization information being put away on the web, and you have an "impeccable tempest" for information security chance. While past money related misrepresentation recognition frameworks depended intensely on perplexing and powerful arrangements of tenets, current extortion identification goes past after an agenda of hazard factors – it effectively learns and aligns to new potential (or genuine) security dangers.

5. Development RoadMap

2018 Q4 First version of gCrawl ready to test

2019 Q1 First version of gStorage, gPredict ready to test

2019 Q2 First version of gCompute ready to test.

6. Deuro's Economics and Token economy

What makes Deuro one of a kind is that it's not only an arrangement of AI services or libraries—it's a secure system with a full set of features. Likewise, the service runs on top of its own native token economy, which is driven by most recent blockchain innovation.

6.1. Functional utility token

The native computerized cryptographically-secured protocol token of the Deuro(DUR) is a noteworthy segment of the Deuro and is outlined to be utilized exclusively on the system. DUR is a non-refundable utility token which will be utilized as the platform currency in the Deuro ecosystem. For activities did on the Deuro System, the expenses are to be evaluated in DUR also, paid to the Deuro System or the exchange counter-party. DUR does not in any way speak to any shareholding, support, right, title, or interest in the Foundation, its partners, or some other organization, endeavor or undertaking, nor will DUR qualifies token holders for any guarantee of expenses, income, benefits or venture returns, what's more, are not expected to constitute securities. DUR may just be used on the Deuro System, and the ownership of DUR conveys no rights, express or inferred, other than the privilege to utilize DUR as a way to empower the use of and collaboration with the Deuro System.

Specifically, you comprehend and acknowledge that DUR : (a) isn't a credit to the Foundation or any of its members, isn't proposed to speak to an obligation owed by the Foundation or any of its associates, and there is no desire for benefit; and (b) does not furnish the token holder with any possession or other enthusiasm for the Foundation or any of its subsidiaries. (c) is non-refundable

and can't be traded for money (or its comparable incentive in some other virtual cash) or any installment commitment by the Foundation or any partner. (d) does not speak to or present on the token holder any privilege of any shape as for the Foundation (or any of its offshoots) or its incomes or resources, including without impediment any privilege to get future income, shares, proprietorship right or stake, offer or security, any voting, dissemination, reclamation, liquidation, restrictive (counting all types of protected innovation), or other monetary or lawful rights or proportional rights, or licensed innovation rights or some other type of support in or identifying with the Deuro System, the Foundation, the Distributor and additionally their specialist co-ops; (e) isn't planned to be a portrayal of cash (counting electronic cash), security, ware, bond, obligation instrument or some other sort of money related instrument or venture;

6.2. The Deuro's Marketplace

The Deuro's Marketplace is a decentralized blockchain application based on Deuro system itself. The application servers as a marketplace center for posting, seeking, assessing, utilizing, and positioning reusable parts, for example, gCrawl solution, gStorage service, gPredict training and gCompute instance. The marketplace center is additionally in charge of keeping up feedbacks about the platforms furthermore, the ranking of their developers, and also other data. This data will help users assess reusable segments.

6.3. The Miners

Deuro allows communities("miners") to run Deuro system solutions and join the main network. However, the task of miners within the Bitcoin and Ethereum is quite different from the role of miners within the Deuro System. In Deuro's plan, Miners may give the computing assets, storage capacity and package resources to their own particular utilization, share them cross different gatherings, or offer them with anybody.

DPoS Token MINERS For gStorage, gCrawl, gCompute networks, we will adopt DPoS voting mechanism, since all those networks are not very computationally intensive. Normal desktop is able to maintain a Deuro node and maintain the ledger globally.

Vectority PoW COMPUTING MINERS Computing miners contribute their computing assets to the network. They may give distributed computing assets and keep running as a cloud node, or contribute self-facilitated hosted assets. The decision of mining is altogether up to the owner and can be changed as they have to.

PACKAGE MINERS Package miners contribute software packages to the network, for example, new crawling scripts for image recognition or natural language processing, an optimized model for quantitative trading, or prepared AI dApps. A smart contract characterizes how the miner's expense will be dispersed if the segment is created by various people. It will likewise characterize the rules for how the code for the part is forked.

7. Founding Team Members

Gray Chen Project Lead

EOS early developer & investor. Apple Inc. awards him 3 times for world-class top developer, inviting him twice as special guests to attend Apple WWDC(World-Wide-Developer-Conference). Apple WWDC 2015 Scholarship Winner/ Apple WWDC 2016 Apple Guest. He was Chief Data Scientist/ Mobile Engineering tech lead@AISense, founding engineer@AISense. Successfully helped AISense boost the speech recognition(AI) system accuracy for more than 30%, which he implements the high-performance and distributed big data crawling system. Venture partner at Skylight Investment. Bachelor of Science with Highest Distinction from Purdue's Computer Science Department, Dean's List. Graduate from the colleague within 3 years.

Ziteng Zhang Head of Marketing

President of AmiaUnion that offers cross-border branding, marketing and PR service for Chinese and US companies. Clients included ETS, Macy's, UnionPay, Alibaba, BMW China

Prof. He Wang Chief Research Scientist

Professor He is Assistant Professor in the School of Computer Science Department at Purdue University. The University of Illinois at Urbana-Champaign, USA Ph.D., Electrical and Computer Engineering, Aug. 2016. Duke University, USA M.S., Electrical and Computer Engineering, Sept. 2013. Tsinghua University, China B.E., Electronic Engineering, Jul. 2011

Professor he is the reviewer for ACM IMWUT/CHI 2017/UbiComp 2016, IEEE Transactions on Mobile Computing/IEEE/ACM Transactions on Networking/ACM Transactions on Sensor, Networks/IEEE Internet of Things, IEEE Pervasive Computing/IEEE Transactions on Parallel and Distributed Systems/IEEE Transactions on Emerging Topics in Computing, IEEE Transactions on Industrial Informatics/IEEE Transactions on Human-Machine Systems/IEEE International Conference on Communications 2013, IEEE Journal on Selected Areas in Communications/Sensors/Journal of Ambient Intelligence and Humanized/ComputingComputers in Biology and Medicine

8. Advisors

Prof. Hong Wan

Dr.Hong is an Associate Professor in the School of Industrial Engineering at Purdue University. She also directs the Purdue Blockchain Lab, co-directs the Smart Design Lab, and is affiliated with the SEED Center for Data Farming at the Naval Postgraduate School. Her Operations Research interests include design and analysis of computer simulation experiments, stochastic process, quality improvement and quality control, applied statistical methods, and healthcare system engineering.PhD in Industrial Engineering and Management Sciences, Northwestern University,

Dec. 2004, advised by Professors Barry L. Nelson and Bruce Ankenman, Thesis Title: *"Simulation factor screening with controlled sequential bifurcation"*. MS in Industrial Engineering and Management Sciences, Northwestern University, June 2002. MS in Materials Science, Northwestern University, June 2001. BS in Chemistry, with a minor in Economic, Peking University, July 1998

Ryan Xu

Founding chairman at ColinStar Capital. Over 7 years' experience in fintech. Investor in bitcoin mining and other related projects. Initiatives include the Melbourne Bitcoin Technology Centre , Bitcoin Boulevard Australia and Bitcoin Buskers Awarded "Blockchain opinion leader" in 2016.

Sky Yu

Founding general partner at Skylight Investment. Sky was previously an investment manager at Taiyou Fund, during which he invested in Easy Transfer and successfully yielded more than 5 times return on the investment.

Alex Rong

Founding partner of MX Capital(Focusing on U.S. second market investment, total AUM is more than $200M, got 10x in 5 years, holding FB, TSLA, APPL,NTES)

Appendix

A1. Product

Figure 23: Otter.ai Product

TechCrunch: Otters new app lets you record transcribe search and share your voice conversations
NVdia: Otter App Aims to Use Power of AI to Set Gold Standard for Note Taking
ZDNet: AI breakthrough: Otter.ai app can transcribe your meetings in real time, for free
Mashable: Otter app transcribes conversations like it's no big deal

A2. Portfolios

Figure 24: HyperLoop One

Hyperloop One is an American company in Los Angeles, California, that is working to commercialize the Hyperloop for moving passengers and/or cargo at airline speeds at a fraction of the cost of air travel.

Figure 25: RaidenNetwork

The Raiden Network is an off-chain scaling solution, enabling near-instant, low-fee and scalable payments. It's complementary to the Ethereum blockchain and works with any ERC20 compatible token. The Raiden project is work in progress. Its goal is to research state channel technology, define protocols and develop reference implementations.

Figure 26: 0x Protocol

0x is an open, permissionless protocol allowing for ERC20 tokens to be traded on the Ethereum blockchain.

Figure 27: Kyber Network

Kyber Network is a new system which allows the exchange and conversion of digital assets. We provide rich payment APIs and a new contract wallet that allow anyone to seamlessly receive payments from any tokens.

Figure 28: Orchid Protocol

The Orchid Protocol is the decentralized, open-source technology for an Internet free from surveillance and censorship

Figure 29: Open Zeppelin

OpenZeppelin is an open framework of reusable and secure smart contracts in the Solidity language.

References

Bridgewater Associates. https://www.bridgewater.com/. URL https://www.bridgewater.com/. Accessed on Wed, April 25, 2018.

Mastering the game of Go with deep neural networks and tree search. http://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf. URL http://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf. Accessed on Wed, April 25, 2018.

Artificial Intelligence Startups. https://angel.co/artificial-intelligence, a. URL https://angel.co/artificial-intelligence. Accessed on Tue, April 24, 2018.

Global AI market size 2016-2025 — Statistic. https://www.statista.com/statistics/607716/worldwide-artificial-intelligence-market-revenues/, b. URL https://www.statista.com/statistics/607716/worldwide-artificial-intelligence-market-revenues/. Accessed on Mon, April 23, 2018.

Bridgewater's Ray Dalio Spreads His Gospel of 'Radical Transparency'. https://www.nytimes.com/2017/09/08/business/dealbook/bridgewaters-ray-dalio-spreads-his-gospel-of-radical-transparency.html. URL https://www.nytimes.com/2017/09/08/business/dealbook/bridgewaters-ray-dalio-spreads-his-gospel-of-radical-transparency.html. Accessed on Thu, April 26, 2018.

Zoom video conferencing service raises $100 million from Sequoia on billion-dollar valuation. https://techcrunch.com/2017/01/17/sequoia-invests-100-million-in-zoom-video-conferencing-service/. URL. Accessed on Tue, April 24, 2018.

Artificial general intelligence - Wikipedia. https://en.wikipedia.org/wiki/Artificial_general_intelligence. URL. Accessed on Wed, April 25, 2018.

China's AI startups scored more funding than America's last year. https://www.technologyreview.com/the-download/610271/chinas-ai-startups-scored-more-funding-than-americas-last-year/. URL https://www.technologyreview.com/the-download/610271/chinas-ai-startups-scored-more-funding-than-americas-last-year/. Accessed on Tue, April 24, 2018.