The Dopamine Network

Crowdfunding Whitepaper

Dr. Uri Yerushalmi

Version 1. 0

Abstract

As scientists, we try to replicate nature in ways that benefit the greater cause of humanity. From medicine to energy, nature is the basis of technologies that sustain our life. In the field of artificial intelligence (AI), we use neuroscience, the study of the human brain, as an inspiration to develop machines that can think at a high level. Intelligent machines can solve problems, such as having to work, drive, or clean your home. Super-intelligent machines can potentially solve even bigger challenges, such as hunger or natural disasters.

A neural network is a sophisticated data processing paradigm that is inspired by biological brain. We believe that highly sophisticated data processors, individuals or organizations like Google with intellectual property (IP) for processing or holding data, can form networks that independently learn and grow much more rapidly than a simple neural network. The conduit to super-intelligence will be through the formation of Decentralized Collaborative Processing Networks (DCPNs) where highly sophisticated data processors work together to create a larger super mind. A great example of a naturally occurring simple DCPN is the collaboration in medical research done by virologists during the outbreak of Zika. Many scientists from around the world selflessly shared their findings, built on top of others' work, and collaborated tirelessly to identify Zika's root causes and implement preventative measures to control the outbreak. Within months, more research had been done on Zika than on any other outbreak in history over such a short period of time. Their incentive was to save humanity, and the urgency was high. But how can you bring sophisticated researchers together when the urgency is not as tangible as an epidemic? How do you encourage collaboration to find solutions to global warming, war, and pollution?

The Dopamine Network is the first ecosystem for Decentralized Collaborative Processing Networks (DCPN), protecting individuals' IP while providing an incentive structure designed to encourage collaboration and penalize bad actors. The world is dominated by centralized organizations who do not collaborate because societal and economic incentivizes encourage competition and IP protection. Through the use of Ethereum smart contracts and Dopamine tokens, we can economically incentivize individuals and organizations to use their IP collaboratively, reducing the need to trust each other or join a centralized organization.

Dopamine is committed to building an ecosystem of data suppliers and consumers of processed data who will supply data to and purchase results from the DCPN. The first consumer of DCPN output will be the Dopamine Fund, established from the crowdfunding sale of DOPA tokens, incentivizing data consumers and service providers to join the network. Dopamine Foundation will also actively seek data processors to join the network and create campaigns for individuals and organizations to publish their data to the Dopamine Network in exchange for DOPA. Our mission is to attract, through economic incentivization, the world's first DCPN structured in a way that allows for learning, growth, and general super intelligence.

1. Introduction

Highly competitive and innovative organizations are constantly looking for new markets and exploring new paths to success. As Big Data and new technological advancements continue to pave the road towards a new era of competitiveness, data continues to be one of the most valuable assets an organization can have. Analyzing and using customer data is not a new concept. In fact, data has been used to improve sales and marketing since the first mail-order catalogs. Today, new technologies and decreases in data storage costs have made it possible to collect massive amounts of data. From generating sales reports to developing predictive business intelligence, using customer data is the best way of enhancing efficiency and customer relationships, as well as opening new revenue streams. During the last decade, however, the sole nature of data and its environment started to change as companies started to realize that data offers additional potential when it is properly enhanced and analyzed.

According to PwC's Global CEO survey, 68% of CEOs see data as generating the greatest return for stakeholder engagement. Even so, many business owners are still finding it hard to monetize their data. For many C-suite executives it is still much easier to just consider and treat data as a valuable asset than it is to monetize it and derive value from it, not to mention the monetization difficulty of private data owners. While the benefits and potentials are numerous, big data comes with its share of risks, and while the Federal Trade Commission and National Telecommunication and Information Administration mostly deal with concerns regarding customer privacy, the greatest risk hides somewhere else. That risk is a threat to free market competition.

2. Big Data. Big Problems.

It's not an exaggeration to say that data is a new fuel of global economy. Much like oil a century ago, "data" is a commodity that drives business innovation and creates a competitive advantage for its owners. Big Data generated on the web, by mobile applications, and IoT (Internet of Things) drives AI research, business analytics, and strategic management. Along with that, there is a growing understanding of the symbiotic relationship between data acquisition, data processing, and the development of AI (Artificial Intelligence). Despite these positive trends, the emerging data and AI markets are ridden by structural contradictions. The most pressing problem today is a weak link between data acquisition, data processing, and AI innovation. As a result, individuals and companies who try to enter the emerging data and AI markets are constrained by:

- Monopolization of data by deep pocket companies
- Data acquisition issues
- Unrealized potential of small data providers
- Lack of decentralized market for AI

Above all, there exists no integrated and transparent market for data and AI where multiple players could benefit from cheaper and more available data and services of

data providers, data processors, and AI trainers needed to make Big Data work more efficiently to the benefit of AI innovation and research.

1.1. Data Monopolization

The marketplace of consumer data is under the rule of big data giants. These companies are monopolizing consumer data from social networks, cab companies, and search engines. Sometimes, these companies are themselves social networks or search engines but, more often, consumer data is sold by the primary mining company to a third party, a data broker.

Tech giants, such as Alphabet, Facebook, Amazon, Apple and Microsoft, de facto monopolized the market turning data into the expensive proprietary asset that drives their own profits. Centralized ownership of data has parallels with the oil rent in the oil rich countries. Despite being a public asset on paper, oil is controlled by a tiny fraction of the ruling class. Like the royal families who expropriate public assets from their rightful owners, modern data giants deny the right of real data producers to receive a reward for their contribution to emerging data-driven economy. Monopolization of publicly generated data poses a tangible problem to companies seeking to introduce data-driven innovation in their operations. It also creates a wide array of challenges for the entire data economy:

- Monopolization slows down the development of an integrated and transparent market where data providers, data processors, and end users meet each other.
- As any monopoly, data monopoly drives data prices up.
- Lacking access to the publicly available data, companies have to make large upfront investments into data infrastructure and data sets.
- Data monopoly creates market entry barriers for smaller companies and individuals. If data was cheaper, more companies and individuals would have an opportunity to participate in AI research and innovation.

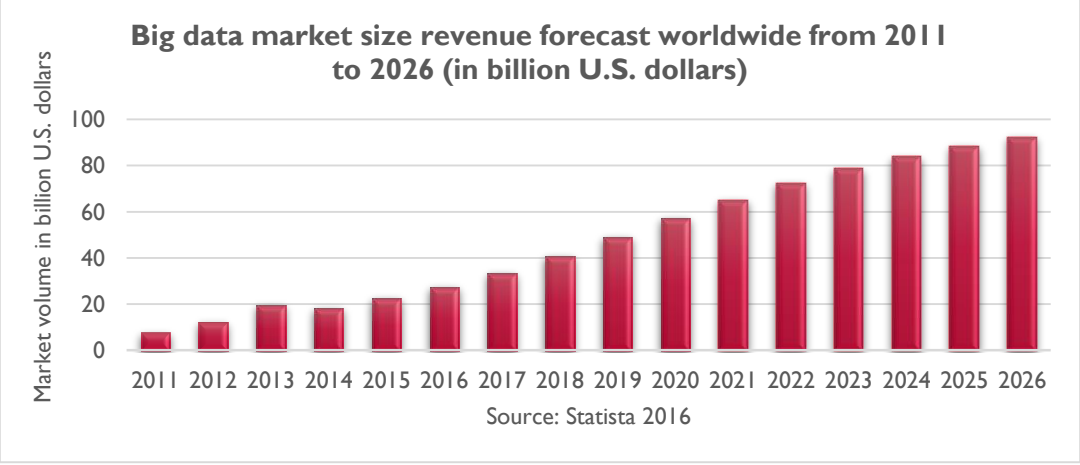


Diagram # 1. Big Data Market Size Revenue (2011-2026)1

[1] Statista. Forecast of Big Data Market Size, Based on Revenue, from 2011 to 2026 (in billion U.S. dollars). Retrieved from https://www.statista.com/statistics/254266/global-big-data-market-forecast/

Above all, monopolization stifles data "monetization" potential. If the data market becomes more democratic, "data monetization" will become one of the major sources of revenue in an economy that will create 180 zettabytes of data (or 180 trillion gigabytes) in 2025, up from less than 10 zettabytes in 2015[2].

The Dopamine network is a disruptive solution to the current data & AI monopoly of the big giants. It does so by allowing frictionless micropayments for exchanging data, by providing a decentralized marketplace for data providers, data processors, and end users to work together, and by removing entry barriers for small players in the data and AI industry.

1.2. Data Acquisition Issues

Even though it becomes easier for companies to obtain data, many businesses are still facing serious data acquisitions issues. While data monopolists, like Google and Facebook, acquire free data from their users who voluntarily leave millions of search queries, comments, ratings, and likes, smaller companies may struggle to find accurate and compelling data to kick-start their AI research and business intelligence strategies. Even though more data is now available via public data sets, it may be hard for them to find a publicly available data set tailored to their specific needs. As a result, companies turn to commercial data set providers that offer data via licensing agreements, cloud-based services, and APIs. The resultant data shortage slows down AI research and makes it impossible for many companies to catch up with tech giants. Under these circumstances, manual in-house work is sometimes the only available alternative for small and medium companies. The problem with in-house data collection and labeling, though, is that they are labor intensive (requiring additional employees) and do not easily scale.

The Dopamine network breaks up this never-ending data acquisition loop by opening a market for exchanging data and enabling also small and mid-sized data consumers to consume data without involving the big giants.

1.3. Unrealized Potential of Small Data Providers

Huge volumes of mission-critical data used in business analytics, advertising, and marketing, in fact, are produced by the anonymous army of Internet users. Most of them are not aware that they are sitting on a goldmine of data that can be monetized. The same is true about businesses and researchers who neglect the opportunities of data crowdsourcing and data acquisition which these anonymous users could provide. The data market, as it looks today, cannot secure fruitful cooperation between data providers, data processors, and end users.

The Dopamine network seeks to make change by empowering individual data providers, both companies and individuals, who can monetize their personal or

[2] Turner, Vernon (2016). IDC Directions Conference 2016. http://techblog.comsoc.org/2016/03/09/idc-directions-2016-iot-internet-of-things-outlook-vs-current-market-assessment/

proprietary data. Following are just some of the user categories for which the Dopamine network will simplify entry to the data market:

- End users of the web who can monetize personal information to be eventually consumed by ad publishers, consumer web companies, and other stakeholders.
- Retail stores that can monetize data related to their consumers, products, or business revenues. At present, many retailers lack infrastructure to monetize their data.
- Users of Internet-connected devices, video cameras, and other equipment that can generate, trace or process data that may be interesting to data-driven businesses and AI researchers.

It is important to note that the potential of small data providers is not only in their ability to provide additional huge amounts of data, but also in the quality of the data, emerging from the expected variety of samples. While in many cases existing open source datasets have their own biases[3], datasets based on many different providers are expected to be less biased.

The Dopamine network will empower these potential data providers to monetize their data and content via bilateral transactions, and contracts accessible via the Dopamine directory and browser extensions/apps.

1.4. Decentralized AI market

Contemporary breakthroughs in AI (Artificial Intelligence) are fueled by the availability of very large datasets. Deep Learning (DL), a sub-set of AI, has recently emerged as the hottest tech trend, with Google, Facebook, Baidu, Amazon, IBM, Intel, and Microsoft introducing AI-first strategies in their software suites and cloud-based services. Many companies followed suit of tech giants introducing their own AI solutions. However, as it turns out, it might be hard for them to boost their AI strategy.

To kick-start AI innovation, companies need labeled, deduplicated, cleaned, and regularized data. These data transformations require more man-hours, which adds economic overhead even for the small datasets. An alternative option is the outsourcing of data services. The main issue with the current data crowdsourcing infrastructure, however, is a middle layer of deep pocket data providers who charge additional fees to data consumers and reduce rewards of data processors.

Also, even after an AI solution module is ready, the current market structure enables using it only once it is integrated in a final solution. It is very rare to see products that include AI solutions from different providers.

With its blockchain system of transparent contracts, the Dopamine network erases the unnecessary layer of data middlemen. Aside from enabling efficient data analysis and data cleanup services, we also expect various data processing services to be rewarded via the dopamine network, ranging from simple services to DL abstractions.

---

[3] FastCompany, Google, Mozilla, And The Race To Make Voice Data For Everyone (2017). https://www.fastcompany.com/40449278/google-mozilla-and-the-race-to-make-voice-data-for-everyone

Other benefits of the Dopamine network for data processors include the following:

- A broader range of data sources, types, and channels.
- On-demand data collection and generation by users from different locations and with diverse backgrounds.

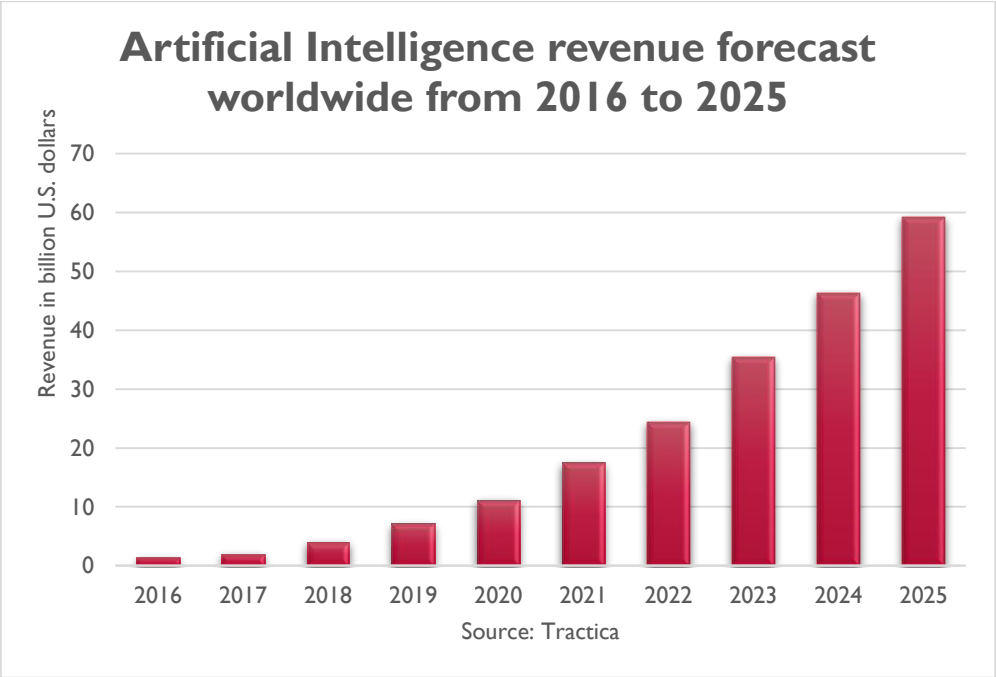


Diagram # 2. Revenue from application of AI software, World Markets (2016-2025)4

These features of the Dopamine network will drive revenues not only from the data industry. The Dopamine network is expected to disrupt the race for IP in the AI space among the leading tech companies, estimated between $26 billion and $39 billion only in 2016[5]. The expected synergies between data processors and data providers coupled with the growing number of market players will benefit the established companies and new businesses seeking to leverage the power of Big Data in their corporate strategy[6]. In its turn, the availability of the high-quality data and data services on the Dopamine network, as well as a vast population of AI trainers, prediction providers, and other data agents will boost AI research and innovation contributing to the widespread distribution of pre-trained models that solve a broad variety of business tasks.

[4] Tractica (2017). Artificial Intelligence Software Revenue to Reach $59.8 Billion Worldwide by 2025. Retrieved from https://www.tractica.com/newsroom/press-releases/artificial-intelligence-software-revenue-to-reach-59-8-billion-worldwide-by-2025/

[5] McKinsey (2017). McKinsey's State Of Machine Learning And AI, 2017. Retrieved from http://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx

[6] IDC (2015). Press Release. Retrieved from https://www.idc.com/getdoc.jsp?containerId=prUS41826116

2. Revolutionizing Big Data with Blockchain and AI

As data monopolies continue to mine billions of data points each month, they are creating a marketing industry valuated at $156 billion per year[7]. At the same time, consumers are not the owners of their data. Since most of data and IP is owned by these deep pocket firms, other companies must struggle with high-priced end-to-end services. While these companies are gradually initiating data science departments to gather and process data internally, independent developers or small IP related firms usually will not be able to keep their own IP. Instead, they will have to give services to enterprises, including handing off their IP.

2.1. Deep Learning

Deep Learning, a subfield of Machine Learning and AI inspired by biological computational models (brain and neural networks), has passed through some radical transformations over the past few years. Deep Neural Networks have enabled critical advances in computer vision, image recognition, bioinformatics and Natural Language Processing. Deep Learning is the bleeding edge of AI revolution with a market that is projected to grow up to $1.7 billion in 2022[8].

Deep neural networks have a few distinctive features that may be leveraged in a decentralized system such as the Dopamine network:

- Consist of a deep cascade of multiple layers that extract features from data and do transformations. Each layer of a deep neural network uses the output from the previous layer as its input.
- Use complex unsupervised learning techniques to produce a hierarchical representation of data. The network derives higher level features of data from lower level features. For instance, deep learning networks used for image recognition derive edges and corners as lower level features, geometric structures as the middle level, and objects (e.g. a car or a dog) as the higher-level features.
- Learn multiple levels of data representation that grasp different levels of abstraction.
- Can be broken into modules. A typical machine learning process involves forward-propagation of data and backpropagation of errors. During forward-propagation or

[7] DMA (2017), The Value of Data, http://ddminstitute.thedma.org/files/2013/10/The-Value-of-Data-Consequences-for-Insight-Innovation-and-Efficiency-in-the-US-Economy.pdf
[8] Market & Markets. Deep Learning Market Worth 1,722.9 Million USD by 2022. http://www.marketsandmarkets.com/PressReleases/deep-learning.asp

backpropagation, each successive layer of a neural network is locked until data or error propagates to it. A modular system may be used in DL networks to break this constraint and decouple the network's modules[9].

- Can be reused as pre-trained models (Transfer learning)[10]

These days, Deep Learning (DL) innovation happens mostly in a centralized way. Typically, one entity manages the entire DL process starting from data acquisition and ending with the training of its numerous hidden layers. The most obvious rationale behind this centralized approach is the need to protect the Intellectual Property (IP) from being taken over by other companies. This approach has several limitations, though. It requires hiring a full-fledged AI/ML team with the expertise in every single aspect pertinent to the Deep Learning. Similarly, centralized DL solutions are less flexible in Transfer Learning, which frequently involves third-party libraries and models hard to acquire under the tight control of IP. To top it all off, centralization of DL makes AI research less transparent and publicly controlled, which raises social concerns over the risks of Artificial Intelligence.

To address the above-mentioned problems, we need a decentralized network where AI innovators are incentivized to create decentralized solutions while keeping their IP safe. The Dopamine network moves DL research precisely in that direction. Its decentralized infrastructure allows splitting the learning process into separate stages where various hierarchies of data (low-dimensional vs. high-dimensional representations) may be used by different entities depending on their expertise and needs. The Dopamine network also solves the problem of data availability for Deep Neural Networks by creating an easy to use interface for data providers and data processors to exchange data.

[9] Market & Markets. Deep Learning Market Worth 1,722.9 Million USD by 2022. http://www.marketsandmarkets.com/PressReleases/deep-learning.asp

[10] Bengio, Yoshua (2017) Deep Learning of Representations for Unsupervised and Transfer Learning. Proceedings of ICML Workshop on Unsupervised and Transfer Learning, PMLR 27:17-36 Retrieved from http://proceedings.mlr.press/v27/bengio12a.html (page 30)

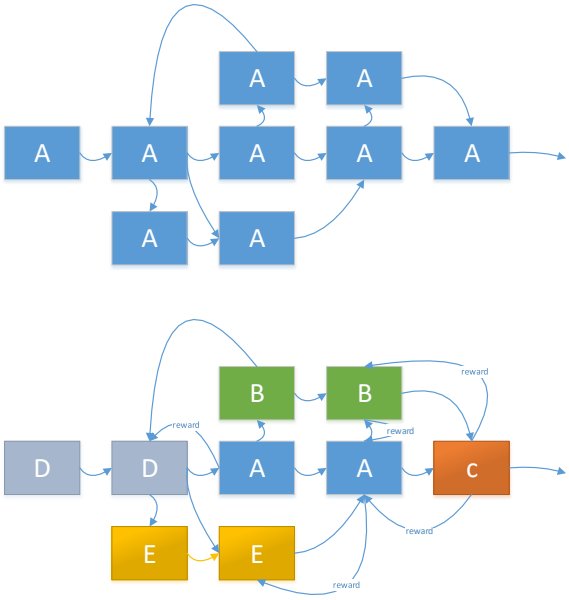


Image #1 Top: A Centralized type of DL solution, where all components belong to the same entity 'A'. Bottom: a decentralized solution using the Dopamine network. Every box in the figure above represents a DL module; every color represents a different entity ('A' to 'E').

2.2. Blockchain technology and Dopamine

Blockchain technology is a distributed ledger that keeps records of digital transactions. Instead of having a central administration like a database, server, or bank, a distributed ledger is implemented as a decentralized network or database replica synchronized across users via the Internet and controllable by anyone within the system. All blockchain transactions are protected by cryptographic blocks. The system is protected from the fraud by cryptographic puzzles that are difficult to solve intentionally.

Blockchain design has several advantages as an interface for online transactions:

- No middlemen involved. Blockchain transactions are cleared algorithmically and there is no need for banks or trustees to oversee them.
- Strong security. The blockchain is an automated transaction system with a bullet-proof security protocol.
- High-quality data. Blockchain data is consistent, timely, accurate, and widely available.
- Faster transactions. Interbank transactions can take days for final settlement. In contrast, blockchain transactions are processed 24/7 and are cleared in a matter of minutes.
- Easy market entry. Joining blockchain does not require any legal procedures, such as business registration.

These advantages of blockchain technology can power data acquisition, data processing and AI in the Dopamine network. Above all, blockchain technology solves the problem of data monopolization and limited access to data.

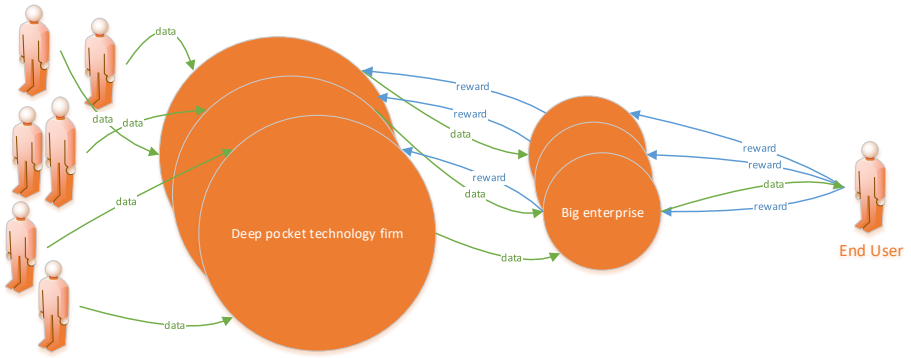


Image #2 Hierarchy of Current Data Economy

Currently, deep pocket technology firms control the flow of data and rewards (see Image #2 above). Having a privileged access to Big Data, they supply it to other big enterprises and end users. This hierarchy drives data prices up at each stage of the data acquisition process.

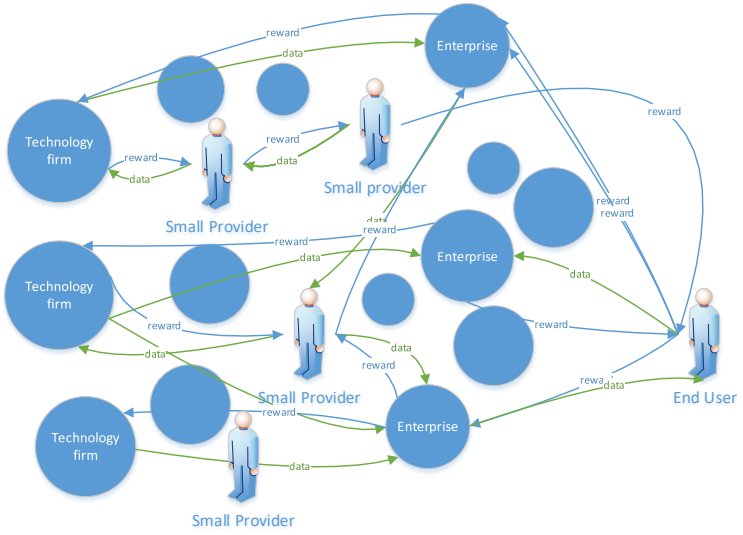


Image #3 Decentralized Data Economy

An alternative option is the decentralized data and AI economy where small providers of data, technology firms, and enterprises enter in mutually beneficial relationships and where strong competition leads to fair prices and easier market entry. This economy benefits both data providers, processors, and what is more important, end users who take advantage of cheaper AI/ML solutions and data.

Leveraging the power of blockchain the Dopamine network provides the following advantages:

- More data. Since more individuals and entities can become data providers, data supply finally meets demand and data consumers can harness data in their business analytics, advertising, and research.
- Better expertise. The Dopamine network offers broad opportunities for data scientists, AI/ML researchers and experts to offer their services to a wide range of data companies and AI startups.
- Cheaper data labeling and data cleansing services. Blockchain removes the middle layer of crowdsourcing platforms that connect data workers with data processors. As a result, data services become cheaper and more available.
- An automatic reward system for data and data services. The Dopamine network makes it easy for a wide variety of users to monetize their data. With access to the Dopamine network provided to users via web browser extensions/apps, any Internet user can become a provider of valuable data for AdTech companies, AI researchers, and data-driven businesses.

Smart contracts in the Dopamine network enable all types of operations with data and support monetization of ML services and models. Below are key benefits of the dopamine smart contracts for data providers and processors:

- Facilitating complex relationships among parties. Dopamine smart contracts may be used to manage complex relationships among various parties in data processing, data provision/acquisition and AI servicing. Scripts can automatically structure these contracts ensuring that contract obligations are properly written out, beneficiaries are rewarded, and disputes resolved.
- Ensuring reliability and trustworthiness. Dopamine smart contracts may involve intermediaries (trustees) or witnesses that guarantee execution of contract obligations for a reward.
- Securing Efficient Division of Labor. Smart contracts can connect data providers (including individuals who know nothing about ML but have access to data via video cameras, mobile phones, etc) to data and AI experts with narrow expertise. This system produces an efficient contract-based division of labor that reduces transaction costs.
- Meeting the requirements of Deep Learning. Dopamine smart contracts can serve different layers of Deep Neural Networks and stages of the learning process. For example, data may be passed from experts in low-dimensional tensors to AI researchers specializing in high-level abstractions. The same division of labor that fuels the dopamine data acquisition may, thus, be used to power Deep Learning.
- Supporting the reuse of pre-trained models. ML models may be leased, shared, or sold to interested parties. Contracts may be also used to train new models for a reward. Each trainer is rewarded per his or her contribution to the learning process.

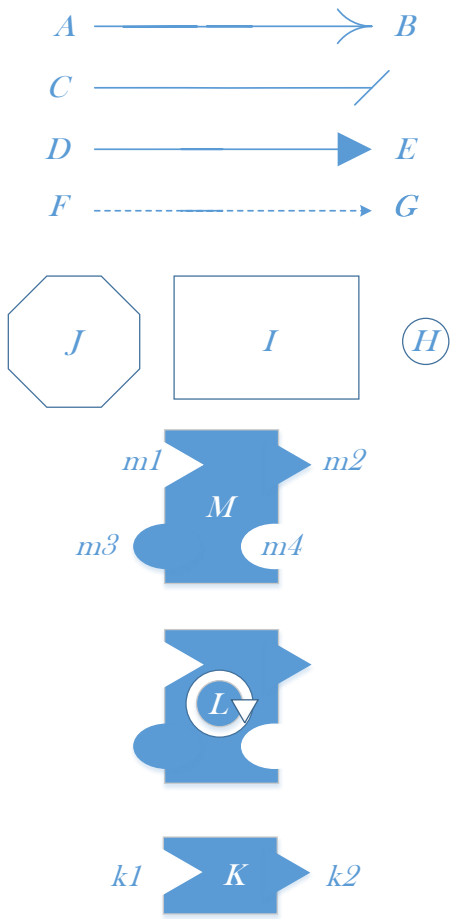
3. The Dopamine Network Project: Towards a Winning Data Monetization Strategy

To address the problem of data monopolization, data availability, and facilitate efficient market relations among data providers, data processors and AI/ML companies and experts, the Dopamine network uses blockchain technology and smart contracts. In combination, these features of blockchain and smart contracts enable the efficient market for data processors and data providers and an environment favorable for the AI innovation and research.

The Dopamine Network is intended for a broad class of decentralized applications with particular emphasis on contexts where the interaction of multiple parties is required and where the service or channel involves multiple successive steps which might require the division of labor. The Dopamine Network does this by combining a blockchain technology with a set of development tools, allowing the Dopamine users to create smart contracts with arbitrary rules of ownership, transaction formats, and state transition functions. Leveraging the power of blockchain, the Dopamine Network is specifically tuned to the requirements of transactions involving data and AI/ML technologies.

3.1. Dopamine Modeling Diagrams

In the following paragraphs, a developmental modeling diagram language is presented, offering a way to visualize a dopamine system's architectural blueprints in a diagram. Including elements as follows (some of the elements are to be explained in following sections):



(A->B) Reward flow from A to B

(C) Flow of contribution from C

(D->E) flow of data from D to E

(F->G) Other association from F to G

(J) Dopamine Certificate

(I) An IP holding entity

(H) Contract

(M) A Learning process, where m1 is its input leg, m2 is output leg, m3 is feedback about input (in deep learning case it would be the gradients tensor for m1), m4 is feedback about m2. Process M does not update its internal parameters.

(L) Like M, where L updates its internal parameters.

(K) A computational process, where k1 is input and k2 is output

3.2. The Dopamine Token (DPMN)

The Dopamine Token would be the only token circulating on the Dopamine network as a reward. Once issued, the DPMN tokens will be enabling transactions, contracts, and distributing rewards.

3.2.1. DPMN Tokens

The Dopamine tokens (DPMN) are standard ERC20 tokens to be tradable in the cryptocurrency exchanges[11].

Diagram #3. A scenario wherein entity B rewards an entity A with Dopamine tokens in exchange for data

3.2.2. Dopamine Smart Contracts

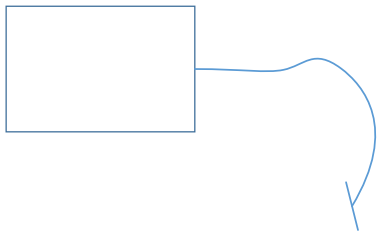
The Dopamine network would include a variety of smart contracts to support different interactions among data consumers, providers and processors.

In many of these contracts, Dopamine tokens are sent based on a 'success' event. When sending tokens conditionally, a consumer will need to declare:

- The number of tokens to be sent in case of "success".
- Deadline for success /failure decision
- Contract type & addresses of witnesses (see below)

3.2.3. Contribution of Dopamine tokens

In certain Dopamine related contracts, it is essential that no party will gain in a given scenario, for such cases, a "Contribution" action is possible, where the tokens are simply "burned", raising the value of (contributing to) all other DPMN token holders.



[11] ERC20 is a token standard developed by Ethereum in 2015 to allow dapps (Decentralized Apps) and wallets to handle tokens across multiple interfaces/dapps.

3.2.4. Conditional deal contract types

3.2.4.1. Single witness

In this case there would be a single witness to have the right of declaring "success" or not. A witness will normally be a third party trusted by both participants of the contract (C):

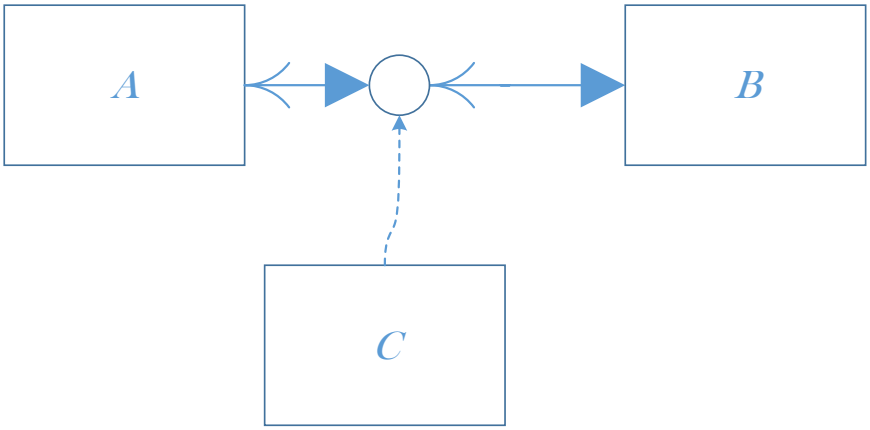


Diagram #4. A scenario wherein an entity B rewards an entity A for data if contract obligations are met. Entity C is a single witness of the contract.

However, the same contract can be cleared by the payer (B) himself without any witnesses at all (see the Diagram # 5).

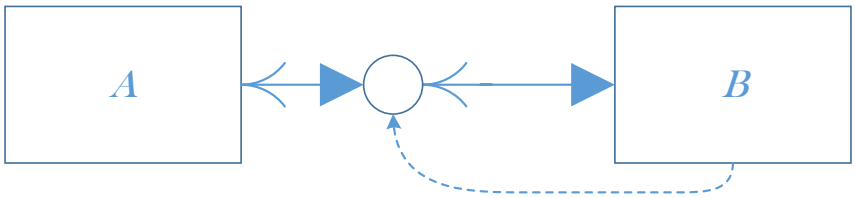
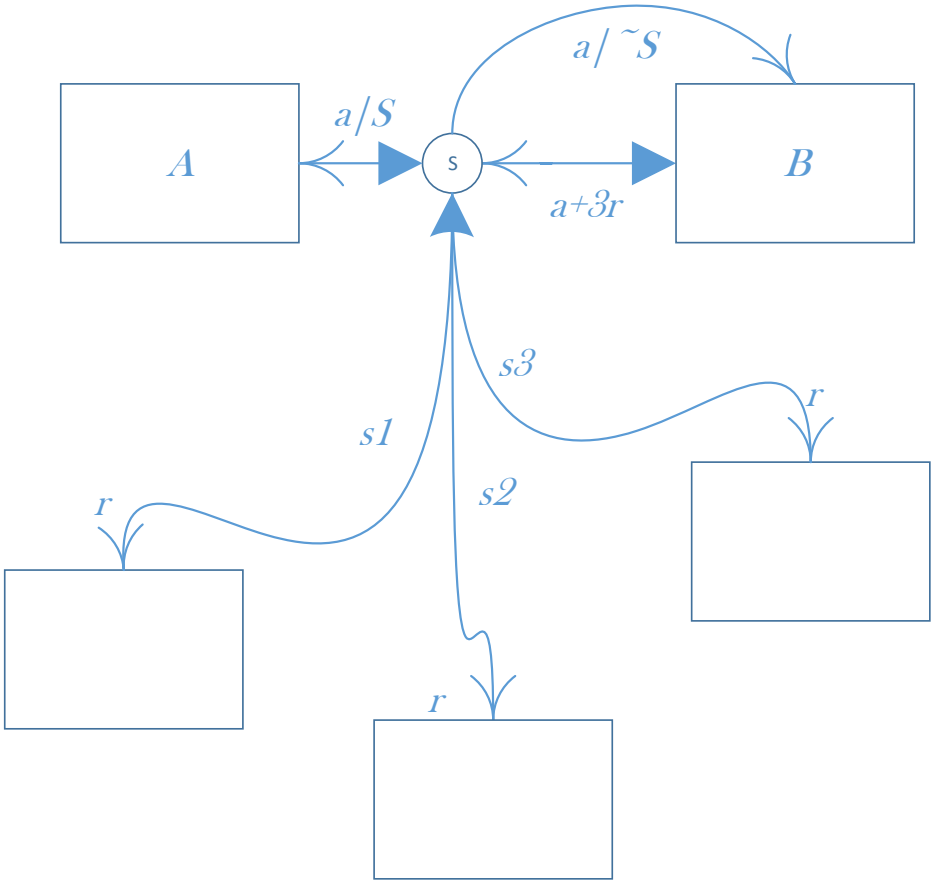


Diagram #5. A scenario wherein an entity B rewards A without any 'success' confirmation by the third party.

3.2.4.2. Multiple Witnesses with the Majority Vote

Under this scenario, success(S)/failure(F) of the contract will be decided by the majority vote of pre-declared witnesses who will be rewarded with r DPMN tokens for their participation. Voting rewards will attract more Dopamine users to become witnesses in this type of contract.

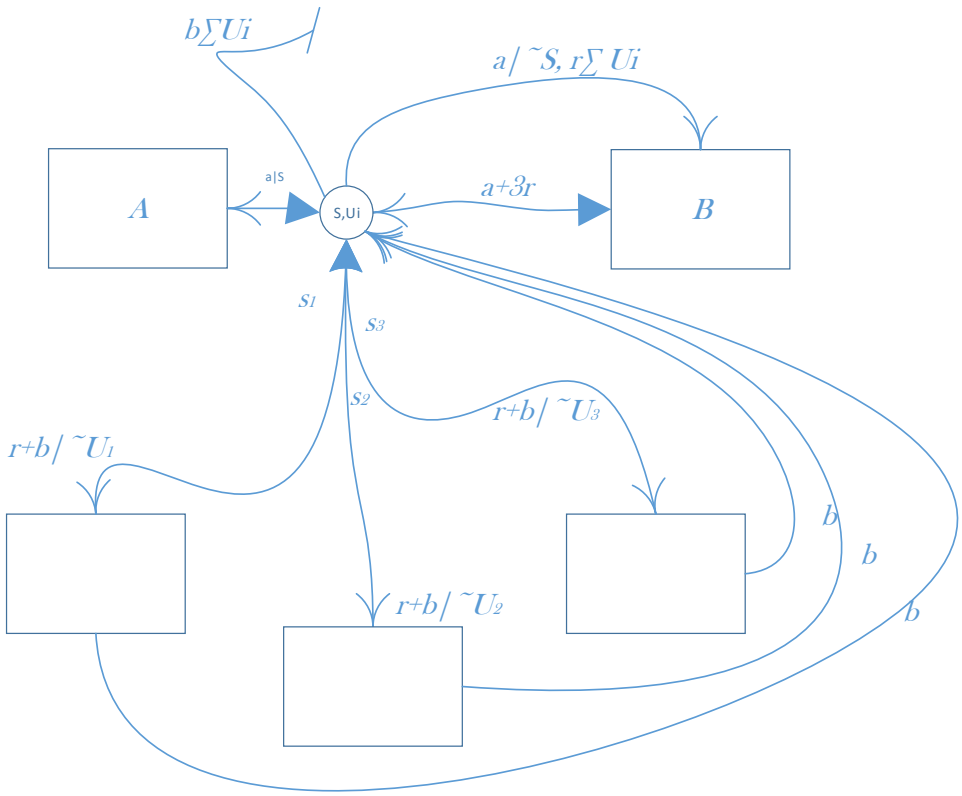


$$s1, s2, s3 \in \{0,1\}$$
$$S = s1 + s2 + s3 > 1½$$

Diagram #6. A scenario wherein three witnesses decide the success (S) of the contract by the majority vote. Each vote is either a 0 (complete failure) or a 1 (success). The 'success' event occurs if the total sum of all votes is greater than 1.5.

3.2.4.3. Multiple witnesses with majority vote and outlier penalizing

In this type of contract, the final decision will be based on the majority vote of several pre-declared players, where each player might be penalized if he was an outlier (U), or rewarded otherwise. An 'outlier' is defined by the contract as any player who does not agree with the majority of the witnesses. To be able to vote, each player will need to put some amount (b) of the DPMN tokens into an escrow fund.



$$U_1, U_2, U_3 \in \{0,1\}$$

Diagram #7. A contract with three witnesses where an outlier witness (U) is penalized by b DPMN tokens and non-outlier witness is paid r DPMN tokens, not including the b tokens he has put in the escrow fund.

3.3. The Dopamine Directory

All services presented on the dopamine network will be listed in the Dopamine directory. The directory will be in form of a website, it will be easily searchable, and will include:

- Description of the provided interface, provided by the service provider.
- Technical interfaces, provided by the service provider.
- Dopamine Certificates (see below)
- Contracts & costs. The dopamine network will apply a transaction only if the terms provided by both sides match.

3.4. Trust mechanisms

When working with data in a decentralized system, in many cases a trust mechanism is required to handle fraud attempts. In such a case, in addition to the agreed reward r that an entity A pays to B, both parties put an agreed amount of DPMN tokens (tA & tB respectively) in a trust fund. After the deal is closed, each party has a time window to declare the deal "unsatisfactory" (F = fa|fb) and erase (contribute) the DPMN amounts

in the trust fund. The described trust mechanism will penalize fraud users and incentivize honest users while increasing transparency and trustworthiness of the Dopamine ecosystem.

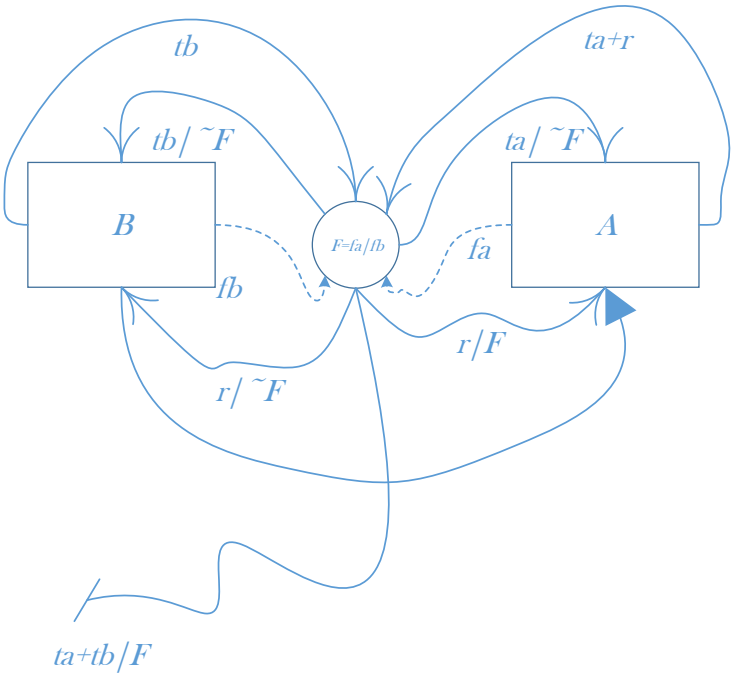


Diagram #8 A Dopamine Network Trust Mechanism

3.5. Dopamine Certificate

Prediction markets where predictions are made about the outcomes of future events are gathering momentum. Unfortunately, there is no good way now to ensure the quality of a prediction unless one holds the prediction algorithm or follows it for a long time. Therefore, enterprises usually either purchase the prediction software or develop it in-house. Addressing this shortcoming of prediction markets, the Dopamine network allows ensuring the quality of a given prediction algorithm without exposing the IP (Intellectual Property), nor by requesting future predictions.

This improvement is achieved in the following way:

A prediction data provider "P" can ask the Dopamine network for a Dopamine Certificate "C". Once the Certificate is issued, the provider can choose a number of DPMN tokens to put as a deposit in the Dopamine Certificate (Pd, which is paid back whenever the certificate is destroyed) and a number of DPMNs to add to the Dopamine Certificate as a contribution (Pc, which is just removed from the DPMN supply, raising the value of all other DPMNs). From that point on, a prediction provider "P" can log all his/her predictions onto the DPMN network. The DPMN network will use cryptographic checks to ensure that all predictions are transmitted as declared.

In their turn, consumers of prediction data will be able to define templates for such certificates. Customized templates will make it easier for prediction consumers to compare predictors and score/rank them.

For example, a hedge fund can request all its potential prediction data suppliers to show ownership of a Dopamine Certificate that:

- Includes a deposit of 1 DPMN
- Includes a contribution of 0.1 DPMN
- Has prediction entries for expected closing prices for all listed stock symbols and for all business days. Entries need to be transmitted before 4 AM ETC.

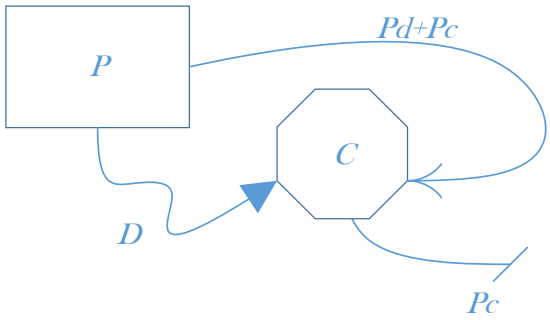


Diagram #9a. The Dopamine Certificate for Prediction Markets. Pd is the deposit in the certificate, Pc is the contribution. D is the prediction data logged in the certificate.

Prediction providers can maintain their Dopamine Certificate without revealing their future predictions. The Dopamine Certificates require holders to deposit funds and contributions, which discourages fraud players trying to predict many outcome combinations in advance to participate in the prediction market. Similarly, since data "D" which represents predictions cannot be removed from the Certificate, prediction providers are encouraged to make quality predictions about future events.

Likewise, Dopamine certificates can be maintained for copyright compliance certification, where the data provider gives writing access to his consumers so that all copyright related complains are available and unremovable:

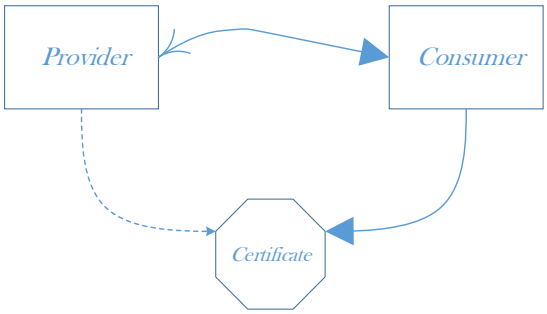


Diagram #9b. Dopamine Certificate provided by the data provider, where the data consumers have writing access.

Dopamine certificates are also part of the solution for bot fueled ad fraud in digital advertising, used as "I'm not a bot" certificate.

3.6. AI Interfaces

The Dopamine network supports passing data to AI/ML interfaces and algorithms for a reward. In this scenario, rewards may be collected either by the owners of data or owners of the AI/ML interfaces.

- Entity A provides prediction services to entity B. In this scenario, entity B passes data to the input layer (m1) of the training algorithm and receives the prediction data from its output layer (m2). In its turn, an entity A receives an agreed reward.

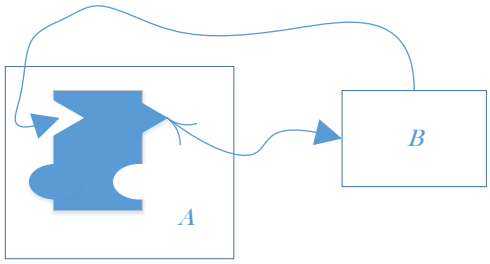


Diagram #10. Prediction Services using AI/ML Interfaces

- Entity B provides training services to entity A: In this case, entity B provides the recurrent training for A's algorithms in the form of data labeling, classification, evaluation of the verisimilitude of generated model samples etc. For these services, B receives an agreed reward from A.

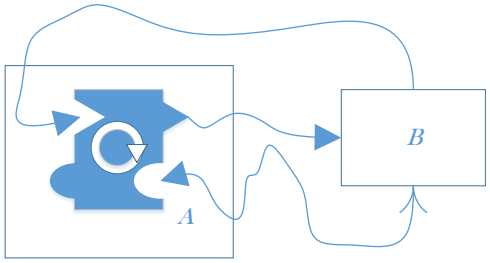


Diagram #11. Training Services for AI/ML Algorithms/Systems

4. Use Cases

The Dopamine network supports many use cases and scenarios that cover a broad range of activities in data acquisition, data processing, and AI/ML application. All use cases below are based on the same tools in the Dopamine pipeline infrastructure described above.

4.1. Targeted Advertising

John browses the web on a regular basis. Ad publishers that buy advertising space on the sites John visits would love to know John's consumer and brand preferences, interests, and demographic data to adjust ads right for him. In his turn, John would not mind receiving a reward from advertisers for sharing his personal data. The question is: what is a way for John and ad publishers to enter a mutually beneficial relationship?

The Dopamine network infrastructure can enable John to become a data provider and help an ad Data Management Platform (DMP) to become a consumer of his data. All John needs to do is to install an open source Dopamine extension/app to his browser. In its turn, the ad DMP needs to integrate the open source Dopamine scripts into its system and agree to reward John at the price John has set. The Dopamine network will take care of each step of the process: clearing the transaction, encrypting data, and delivering it to the data consumer.

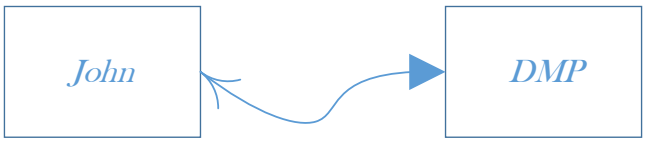


Diagram #12. Targeted Advertising Use Case in Dopamine network

In this way, using the Dopamine network for targeted advertising will enable millions of Internet users to monetize their personal data and increase the ROI for the ad publishers. The opportunity to monetize personal data and show fine-tuned ads will also decrease the ad blockers usage which limits ad sellers revenues. The Dopamine network will incentivize users to stop blocking ads, or at least the ones that reward them back.

Whenever a deal described above occurs, the Dopamine network takes control over the data flow. The data is initially locked by the Dopamine network and transferred to the data consumer while being locked. The data is unlocked by the Dopamine network only after the reward is received by the data provider. All steps of this process are made algorithmically and have a strong encryption protection. The diagram below describes the simplest transaction without trust mechanisms, conditional rewards or other features discussed in other sections of this document.

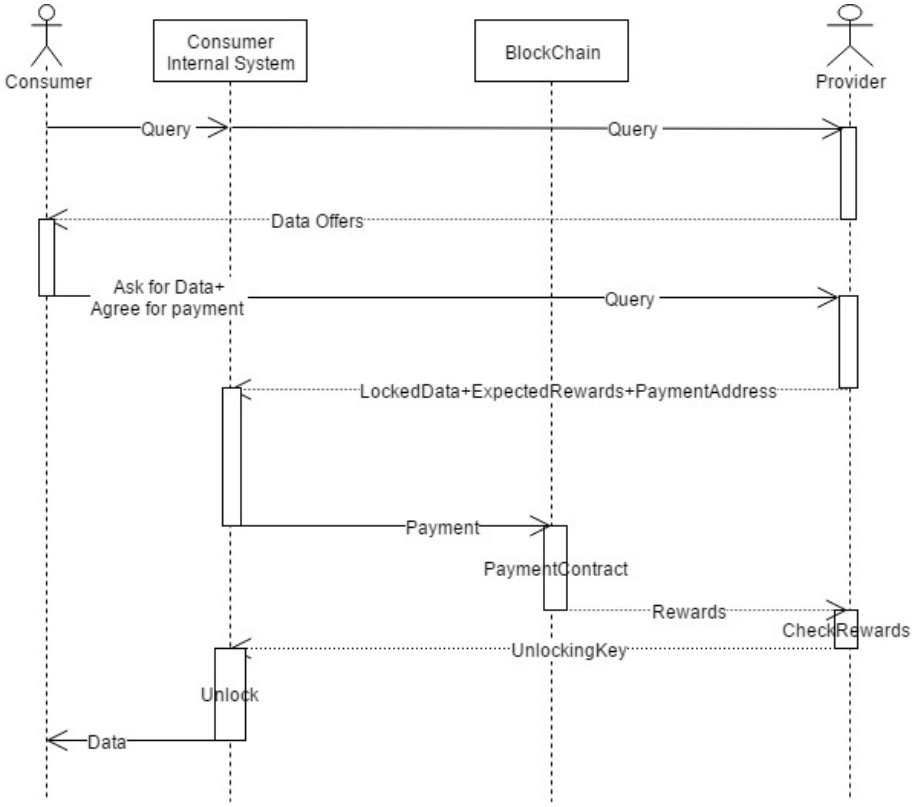


Diagram #13 A Diagram of Data Acquisition Mechanism of the Dopamine Network

4.2. Simple Data Processing

Often, it may be more feasible for a data consumer to acquire the aggregated data instead of collecting the individual data like in the Targeting Advertising scenario discussed above. Let's assume that some firm "A" is willing to know a total number of passengers passing through a certain road. In this case, "A" needs a collective data about all drivers rather than the individual data of each driver. For that reason, the firm "A" can use the services of a firm "P" that collects GPS data from many data providers (drivers) and aggregates it for several firms like "A" and "B"

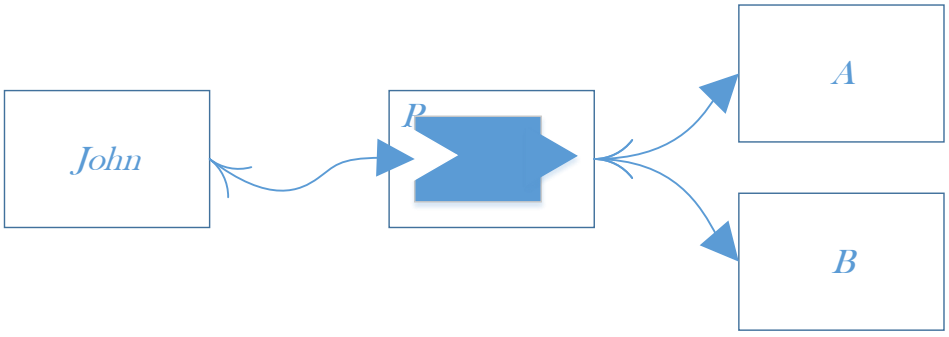


Diagram #14. Simple Data Processing of the Aggregated Data

4.3. Connecting Data Providers, Data Processors, and Data Consumers at Each Stage of the ML (Machine Learning) Process

XHF is a hedge fund trading in crude oil and sugar futures. Data scientists and business intelligence experts at XHF found out that several oil trucks on certain roads is a good predictor of the short-term oil price movements. The hedge fund is motivated to acquire oil trucks data from these roads.

Rick owns a house close to the road. The house is equipped with a camera that can film the truck movement. Even though Rick uses the camera for security reasons, he believes that someone might be interested in acquiring the live stream it creates. For that reason, Rick has installed the Dopamine extension on his PC to which the camera is connected.

Joe is an AI researcher with an expertise in video processing and image recognition. He is specializing in feeding low-dimensional tensors to the deeper layers of neural networks that recognize higher level abstractions, such as objects. Joe offers his AI/ML services on the Dopamine network.

Finally, Jack is an AI expert who can process low-dimensional tensors provided by Joe and train ML algorithms to recognize predefined objects (e.g trucks) like the ones XHF is looking for.

The Dopamine network can help XHF managers find Rick, Joe & Jack, arrange a multi-party contract between them, and get fast access to a streaming feed of the oil truck count in the target road. In his turn, Rick can monetize his camera data that would have gone unnoticed without the Dopamine network. Similarly, with the Dopamine network, Joe and Jack can find an appropriate application of their narrow expertise in AI.

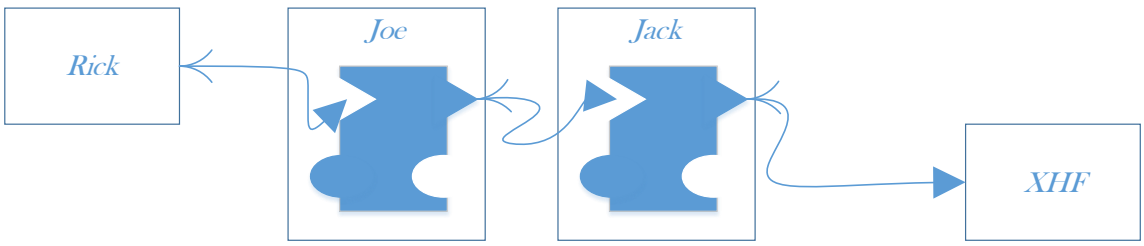


Diagram #15. Connecting Data Providers, Data Processors, and AI experts via the Dopamine Network

4.4. Training Services

These days, access to the high-quality training data is one of the main prerequisites for successful training of AI/ML models. Many companies and researchers spend literally months to collect a good training set for their algorithms. The Dopamine Network saves their time by directly connecting AI/ML innovators to users who offer training and data services for their AI/ML algorithms.

To illustrate this use case, let's assume that Dan is an AI developer who, given a voice sample, needs to classify whether the speaker's age is above or below 30. Even though Dan has some data, it's clearly not enough to train his Deep Learning (DL) model. That's why Dan decides to use all his data as a validation set, and asks the public to train his model on the Dopamine network in exchange for DPMN tokens. Dan is willing to pay 10 DPMN tokens for raising his score from 50 percent to 98 percent. Once Dan has placed all needed tools in the Dopamine Directory and API, users can train his model and be paid according to their contribution. As a result, Dan gains access to the needed data while ensuring that his IP (Intellectual Property) is not exposed.

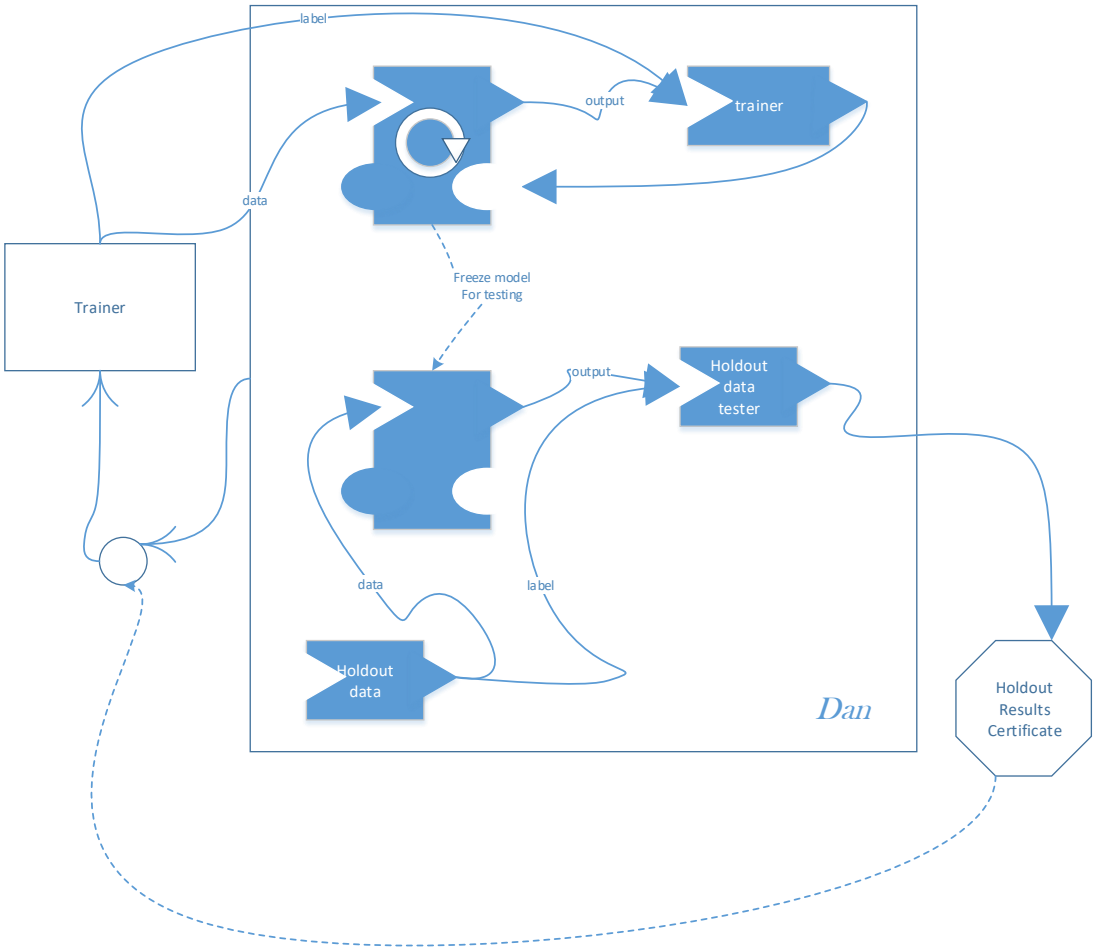


Diagram #16 Training AI&ML model in the Dopamine Network. The Dopamine Network saves time by connecting AI/ML innovators with users. Here, Dan is being trained by "Trainer" who, in turn, is rewarded based on the quality of the trained model on a holdout data. Quality of trained model can be logged into a certificate.

4.4.1. Training Contribution Calculation

Assuming Dan has used the Dopamine API to indicate that the reward is to rise linearly as scores progress, then, internally, the total reward r (10 in our example) for raising the score from S_0 (0.50) to S_1 (0.98) is calculated as follows:

$$r = \int_{S_0}^{S_1} f(x)dx$$

where Dan has set the values for a, b, S_1, S_0

For every trainer *j*, the reward will be calculated according to his/her samples *i*'s contribution to the holdout sample improvement:

Where S_a^i is the score before the sample *i* was learned and S_b^i is the score afterwards.

Assuming training has reached score S_f, Let $S_t = \min(S_1, S_f)$

By being able to filter out non-contributing teachers, sum of all rewards matches:

$$\sum_j r_j = \sum_j \sum_{i \in j} \int_{S_a^i}^{S_b^i} f(x)dx = \int_{S_0}^{S_t} f(x)dx \leq r$$

The Dopamine network can support several types of such f functions, motivating many "trainers" to train such models and be rewarded accordingly.

4.5. Bitcoin Trading Firm Looking for Signal Providers

Pb is a prediction provider that has just entered the prediction market. The company creates a Dopamine certificate and starts logging its predictions about the future Bitcoin prices into it on a daily basis hoping that someday a trading firm *T* will be convinced to use *Pb*'s prediction services. In its turn, *T* has already examined *Pa*'s Dopamine Certificate and found that its predictions are accurate. *T* signs a contract with *Pa* according to which rewards are released to *Pa* based on the quality of its predictions. Both parties agree to use the service of an oracle "O" who declares the actual Bitcoin prices which are used as the benchmark to evaluate *Pa* predictions.

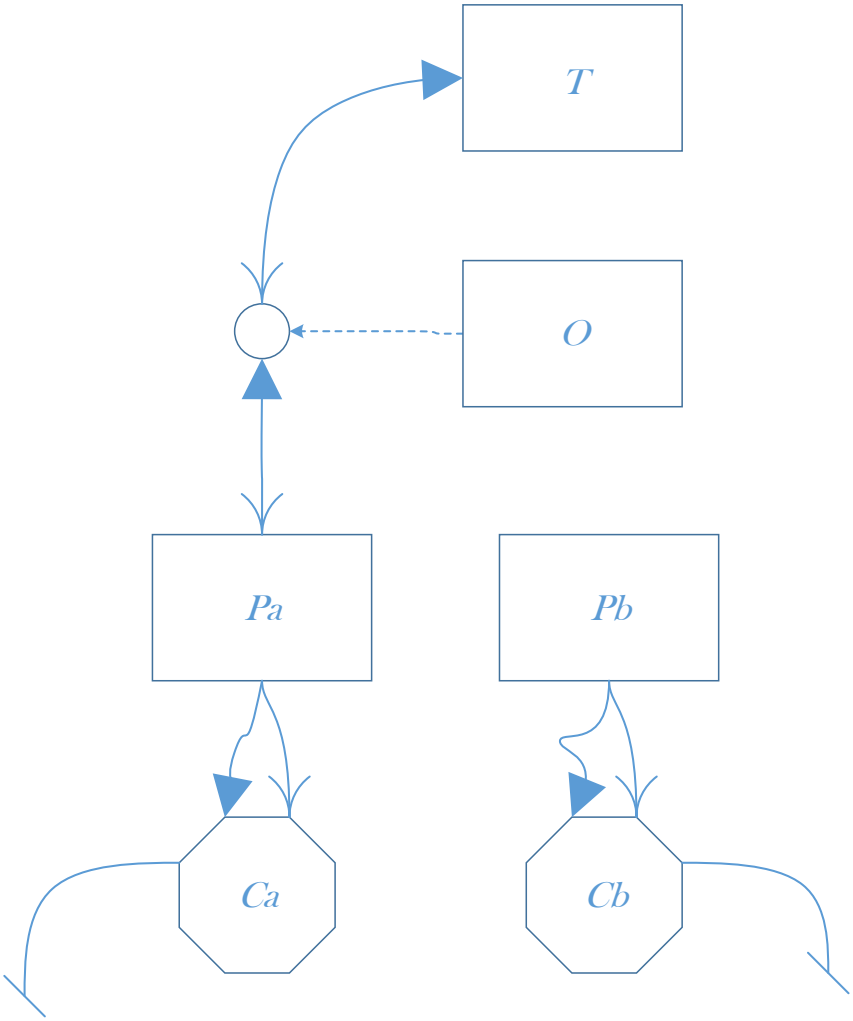


Diagram #17 Bitcoin Prediction Market with the Dopamine Certificate and Independent Oracle. Pa & Pb are two prediction providers with Dopamine Certificates Ca & Cb respectively. Based on the Dopamine Certificate, T enters a contract with Pa, who gets rewarded according to its predictions and the Oracle "O"s input.

4.6. Crowdfunded AI

The residents (*Ri)* of one neighborhood have been suffering from a wave of car burglaries in the parking lots and the local police was not able to handle the issue. Therefore, the residents decided to estimate the likelihood that a car parked in their neighborhood will be broken into, based on different features (model, place of parking, security items, etc). Lacking the necessary skills to produce such an estimation, the residents organized a contest among AI/ML providers (*Pj)* on the Dopamine network. Each of 5000 residents has contributed r = 50 DPMN tokens which amounted to 250,000 DPMNs in crowdfunding. A contest participant with the best estimation model will take all these rewards. In return, the residents will know the way to minimize the likelihood of their cars being burgled in their parking lots.

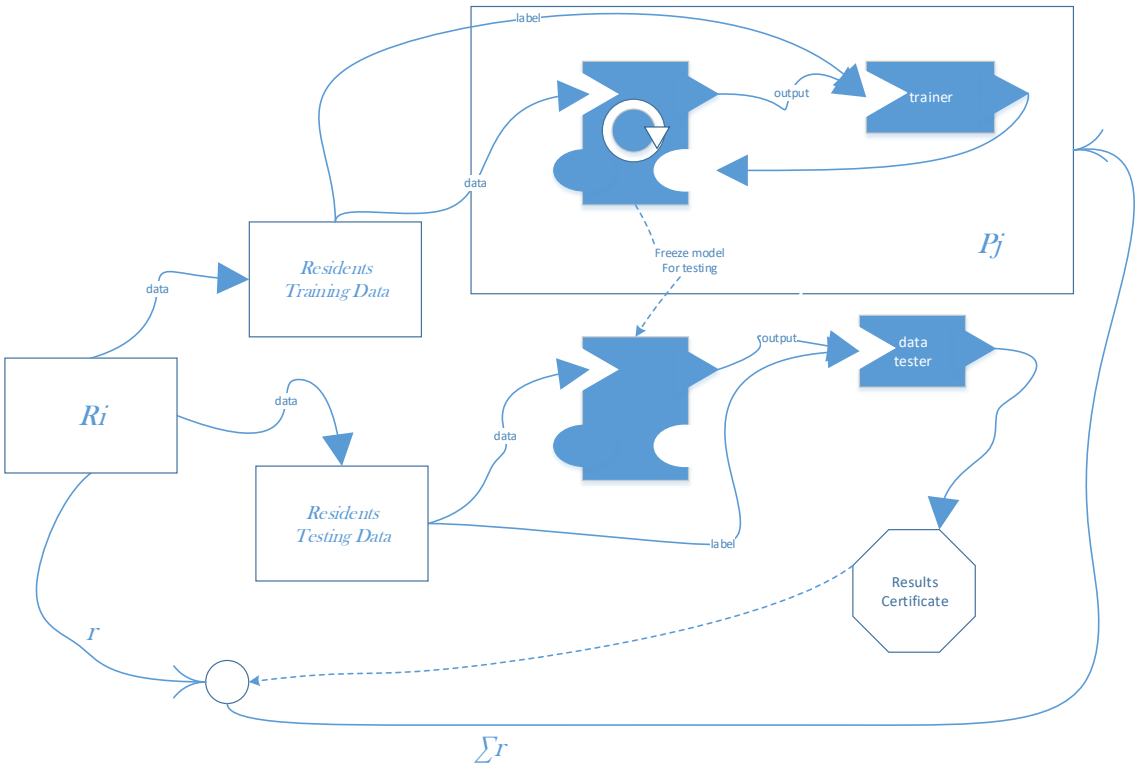


Diagram #18. A Diagram of AI Crowdfunding by the Neighborhood Residents. The Residents Ri train the AI provider Pj and once a model is ready, it is examined, results are logged and a reward is given to the best provider Pj

4.7. Deep Transfer Learning Services

Transfer Learning is an AI technique that allows applying pre-trained models to other tasks with mild changes. Instead of building a new model from scratch, one may use the model trained by others to solve a similar problem.

For example, if an AI developer "A" has trained a neural net "N" to solve some task, the developer "B" can use the knowledge acquired by "A" using the following methods:

- Train the model from scratch, using the topology of the model pre-trained by "A" with randomly initialized weights.
- Fine-tune the pre-trained model using one's own data.
- Fine-tune only some of the layers while freezing the others.
- Retrain some of the layers while freezing the others.

To enable these training options, "A" can place several packages with different pricing in the Dopamine Directory.

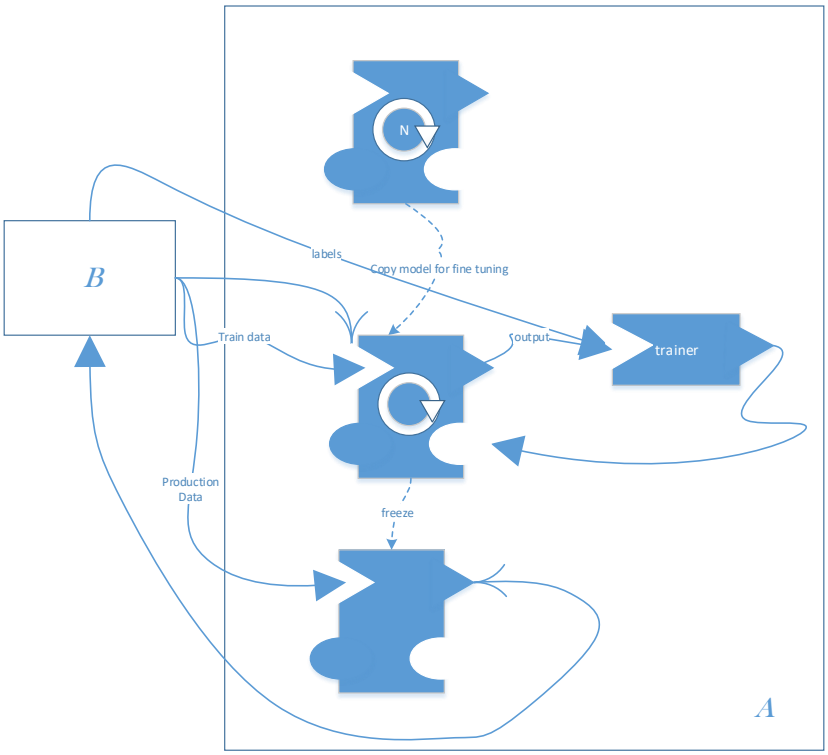


Diagram #19. Transfer Learning with the Dopamine Network

4.8. The Dopamine Network as a Basis of Global AGI (Artificial General Intelligence).

Unlike "Weak AI" approach, where a model is built solving a specific task, an "Artificial General Intelligence" (also referred to as "Strong AI") is aimed to solve a variety of tasks, as a human brain does. Recent development in AGI research shows the following:

- For artificial general intelligence (AGI) it's efficient to apply parameters reuse while having multiple users train the same network.

- Agents have shown a good performance in learning new tasks by discovering which parts of a giant neural network to re-use for these new tasks.

For being general, one probably needs giant neural networks, since the capacity of a neural network to absorb information is limited by its size.

In a paper published in January 2017, AI researchers from Google's DeepMind presented a new model named "PathNet". Each building block of the "PathNet" could be represented as a Dopamine node. For example, two blocks (e.g. "A" and "B") can be implemented as two data processors in the Dopamine network exchanging and passing data to each other. In this way, the Dopamine Network can serve as the pipeline among different neural networks like the one implemented in "PathNet". Below are several sample use cases of this approach.

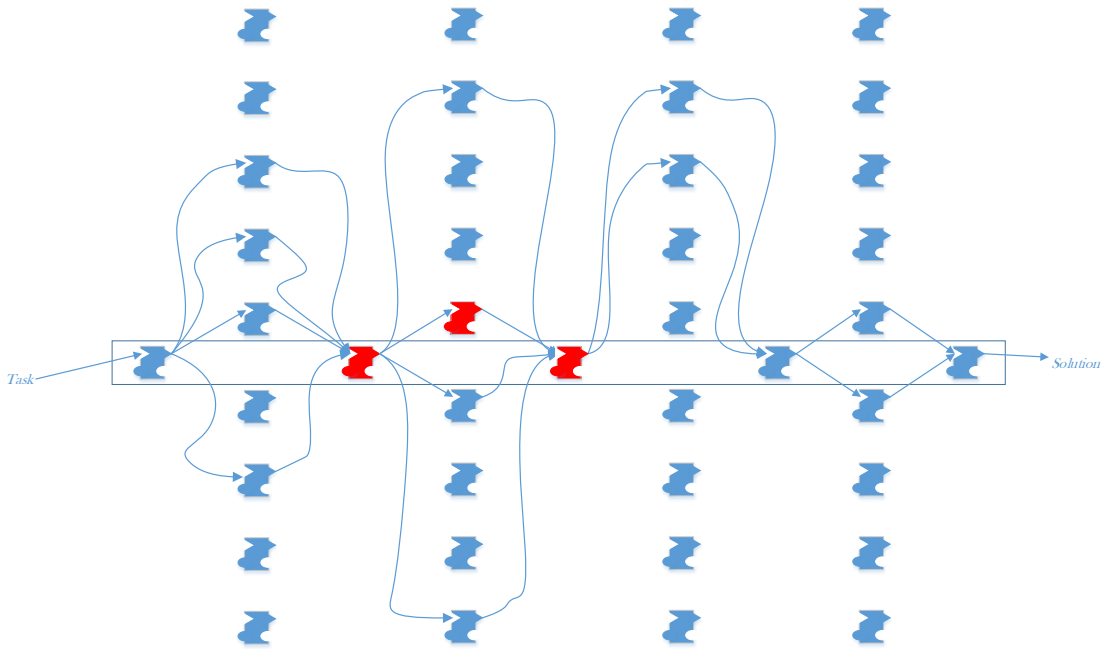


Diagram #20 PathNet like architecture implemented on dopamine network with a building block marked by red items12. The central rectangle is the entity that needs to resolve the task, it does so by cooperating with other layers that may belong to other entities. The red components are zoomed in in the following Images for different use cases

In the diagram below "A" is a neural network in its learning phase. It has not yet arrived at any established weights, so its owner seeks training from another neural network "B" that has made some advances. The owner of the network B will be awarded by the entity A with an agreed amount of DPMNs for the B's training services.

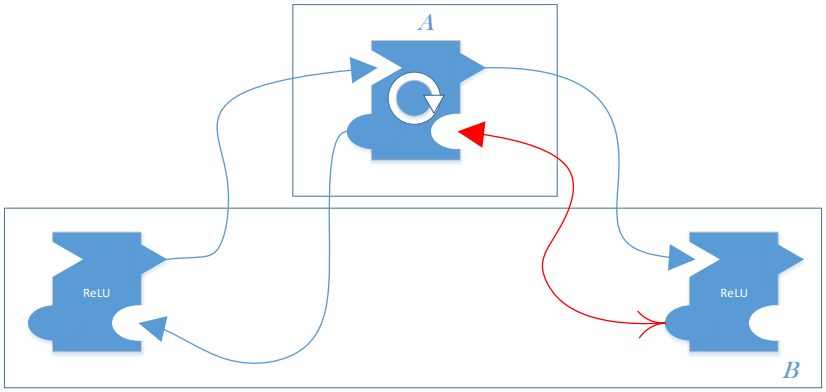


Diagram #21. Entity A awards B for AI Training Services

In the use case below, both A and B have well established weights. However, B has a task that only A can help solve. Thus, the entity B rewards A for solving B's task.

[12] PathNet (2017), A modular Deep Learning Architecture for AGI, https://medium.com/intuitionmachine/pathnet-a-modular-deep-learning-architecture-for-agi-5302fcf53273

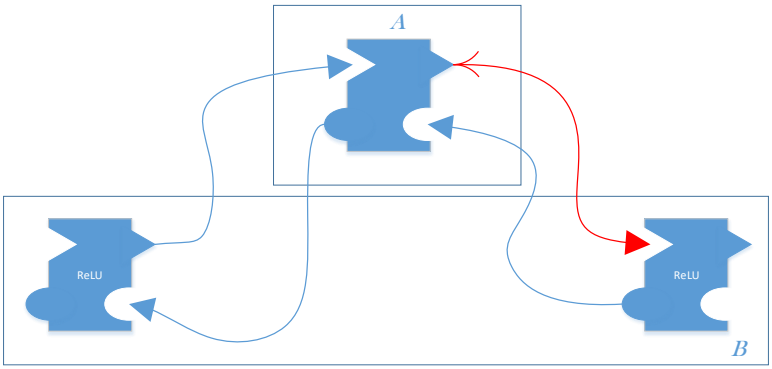


Diagram #22. Two Neural Nets Cooperating in Solving a Specific Task

In the last case, B is learning how to solve its task while A has well-established weights that can help B. B rewards A for training.

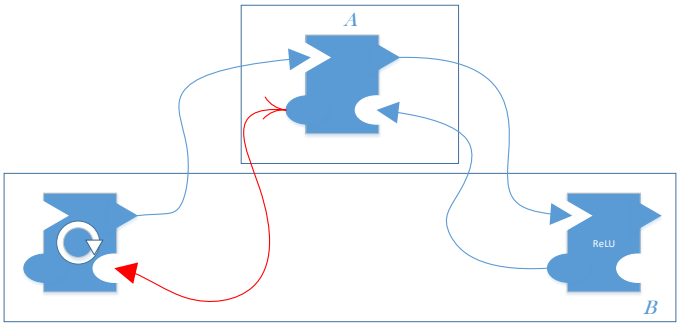


Diagram #23. Solving the Task with the Help of another NN (Neural Network)

5. Conclusion

Highly competitive and innovative organizations are constantly looking for new markets and exploring new paths to success. As Big Data and new technological advancements continue to pave the road towards a new era of competitiveness, data continues to be one of the most valuable assets. The marketplace of consumer data is under the rule of big data giants. These companies are monopolizing consumer data from social networks, cab companies, and search engines.

Tech giants have monopolized the market and turned data into the expensive proprietary asset that drives their own profits. Monopolization of publicly generated data poses a tangible problem to companies seeking to introduce data-driven innovation in their operations. It also creates a wide array of challenges for the entire data economy.

Envisioning the future data industry as a global, cooperative and decentralized industry, the Dopamine network is designed to facilitate exchange among data agents, meet growing demand for quality data sets and data labeling/cleansing services, and to promote deeper synergies between Big Data and AI innovation creating an essential infrastructure for data providers, data processors, and holders of AI and ML solutions to monetize their private and proprietary data, ML algorithms and models via a secure blockchain technology and a system of smart contracts.