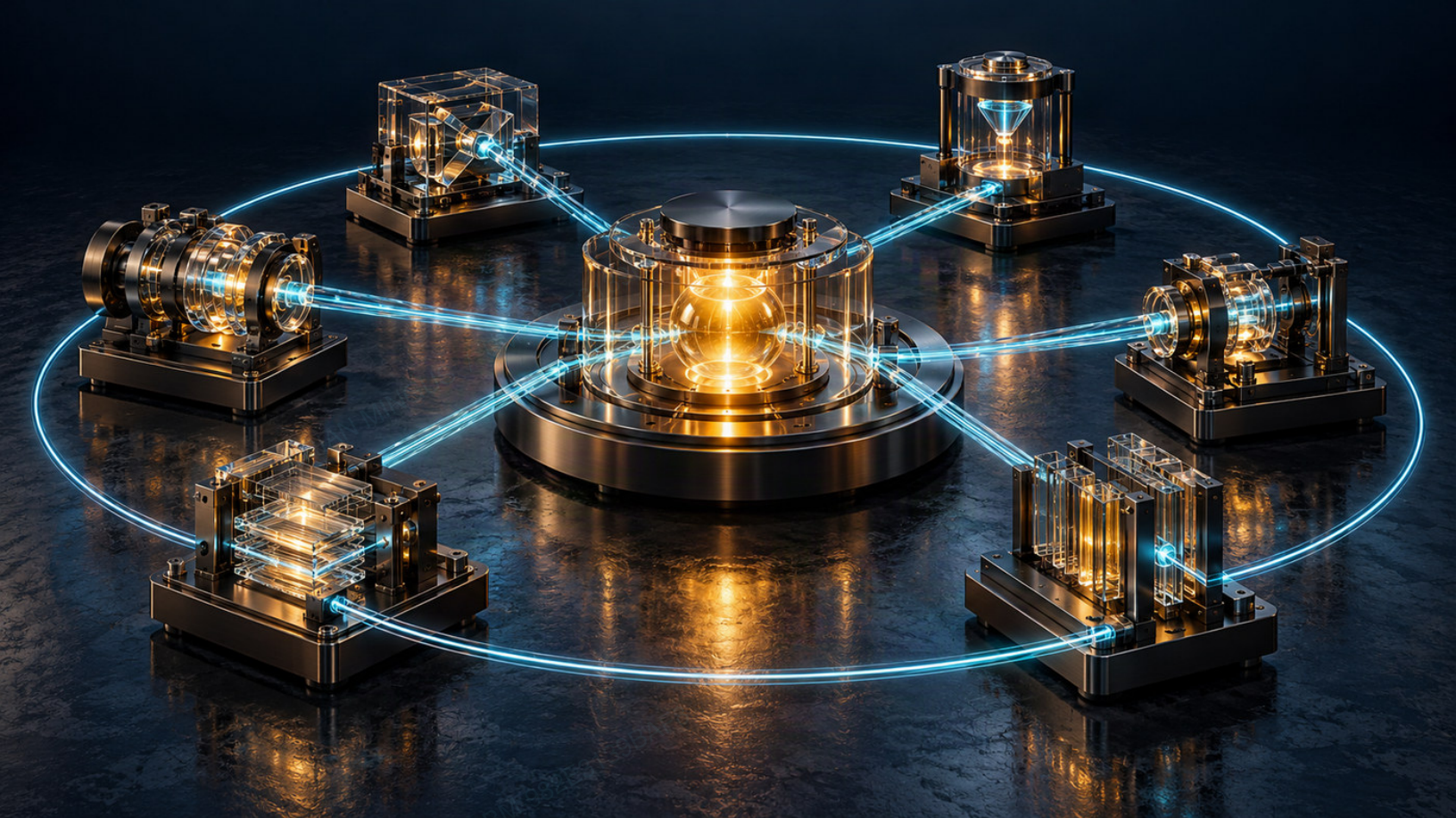




MACHINE LEARNING | U.S. PROVISIONAL PATENT APPLICATION

A Self-Optimizing Information and Software- Brokering Platform as a Service

June 6, 2017 | Application 62/515,837

Brian Mingus | Applicant: Latently

A marketplace for funding, composing and deploying machine-learning ensembles, with a feedback loop that uses its own activity to improve recommendations and guide further development.

Filed June 6, 2017. Complete 16-page application follows.

Collection edition: October 17, 2026. Historical filing; not a grant of patent.

AI-generated conceptual artwork; separate from the original figures.

brianmingus.com

TITLE OF THE INVENTION

A Self-Optimizing Information and Software-Brokering Platform as a Service

CROSS-REFERENCE TO RELATED APPLICATIONS

Not Applicable

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

Not Applicable

THE NAMES OF THE PARTIES TO A JOINT RESEARCH AGREEMENT

Not applicable

INCORPORATION-BY-REFERENCE OF MATERIAL SUBMITTED ON A COMPACT DISC

Not applicable

ABSTRACT

[0001]

TODO

FIGURES

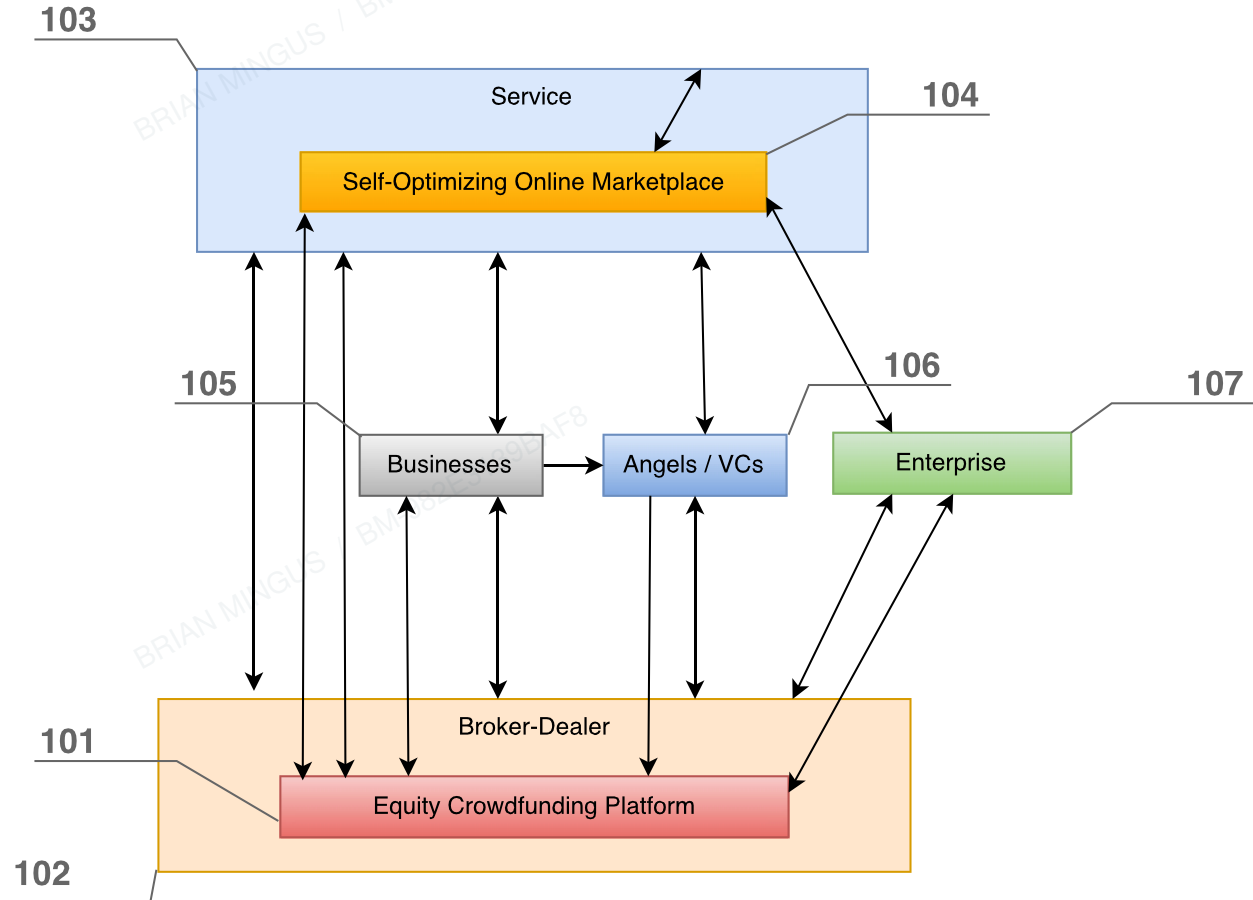


FIG. 1: Overall diagram of the system. The Equity Crowdfunding Platform (101) is developed by Service (103) and hosted by the Broker-Dealer (102). Service (103) and Enterprise (107) post challenges/bounties on the platform (101), resulting in Businesses (105) submitting proposals on the platform (101) that are reviewed by Angels / VCs (106). If Angels / VCs (106) invest, the proposal is developed on Service's (103) platform. The resulting product/ensemble is then hosted on the Self-Optimizing Online Marketplace (104) and purchased by Enterprise (107). The money that Enterprise (107) pays to Service (103) is distributed, according to contracts, to Businesses (105), and from businesses to Angels / VCs (106).

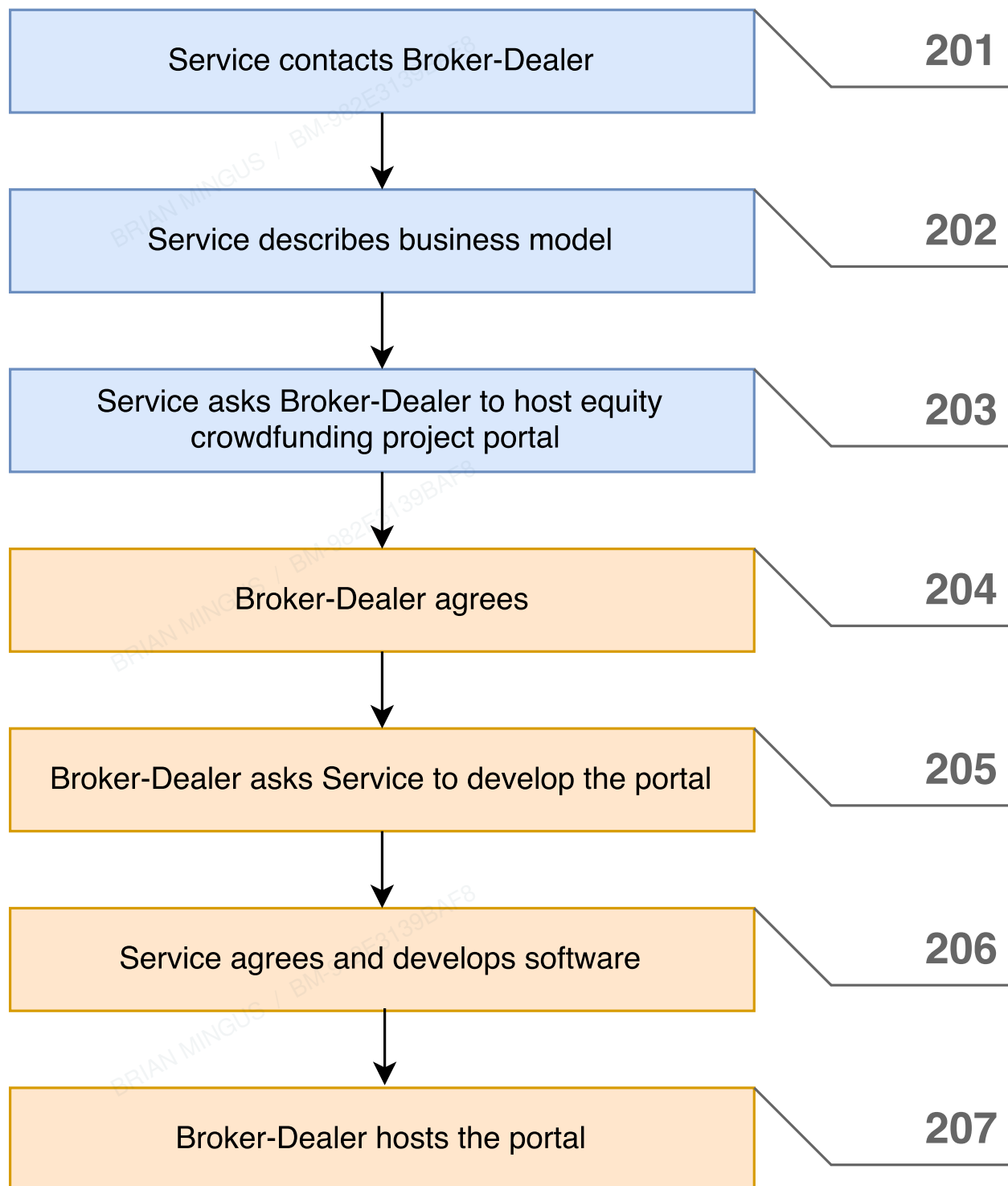


FIG. 2: Overall diagram in which Broker-Dealer (102) contracts software development out to Service (103) and hosts project portal (101). Service first identifies a Broker-Dealer (201) and describes our business model, including the benefit to the Broker-Dealer (201) who receives fees pro-rata the size of the investment by Investors (106) in addition to other incentives. The Broker-Dealer (201) agrees to host the portal (101) (204) and asks Service (103) to develop the portal (101). Service (103) agrees and develops the portal (101) which the Broker-Dealer (102) then hosts (207).

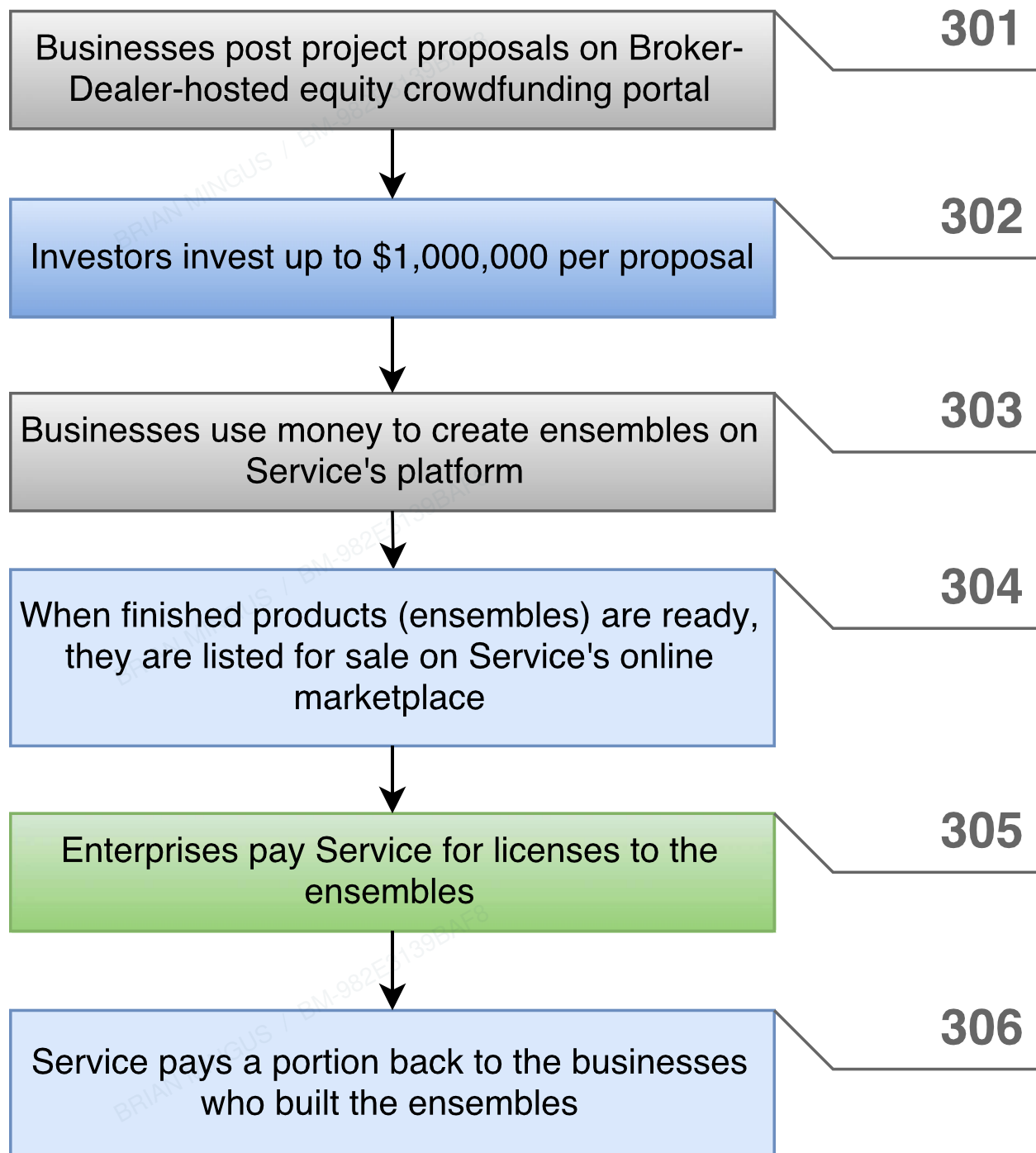


FIG. 3: One process through which Businesses (105) post project proposals on the portal (101), Investors (106) invest, Enterprises buy and the Businesses and Investors get paid. First, Businesses post project proposals on the Equity Crowdfunding Portal (101) (optionally, Enterprises, Service or Investors or anyone else can post the proposals). Next, investors invest up to \$1M per proposal (302) and then Businesses spend the money as agreed in the investment contract on implementing the proposal (303). After the ensembles are implemented they are listed for sale on the Self-Optimizing Online Marketplace (104), Enterprises buy them (305) and Service returns a portion to Businesses as contractually agreed (306).

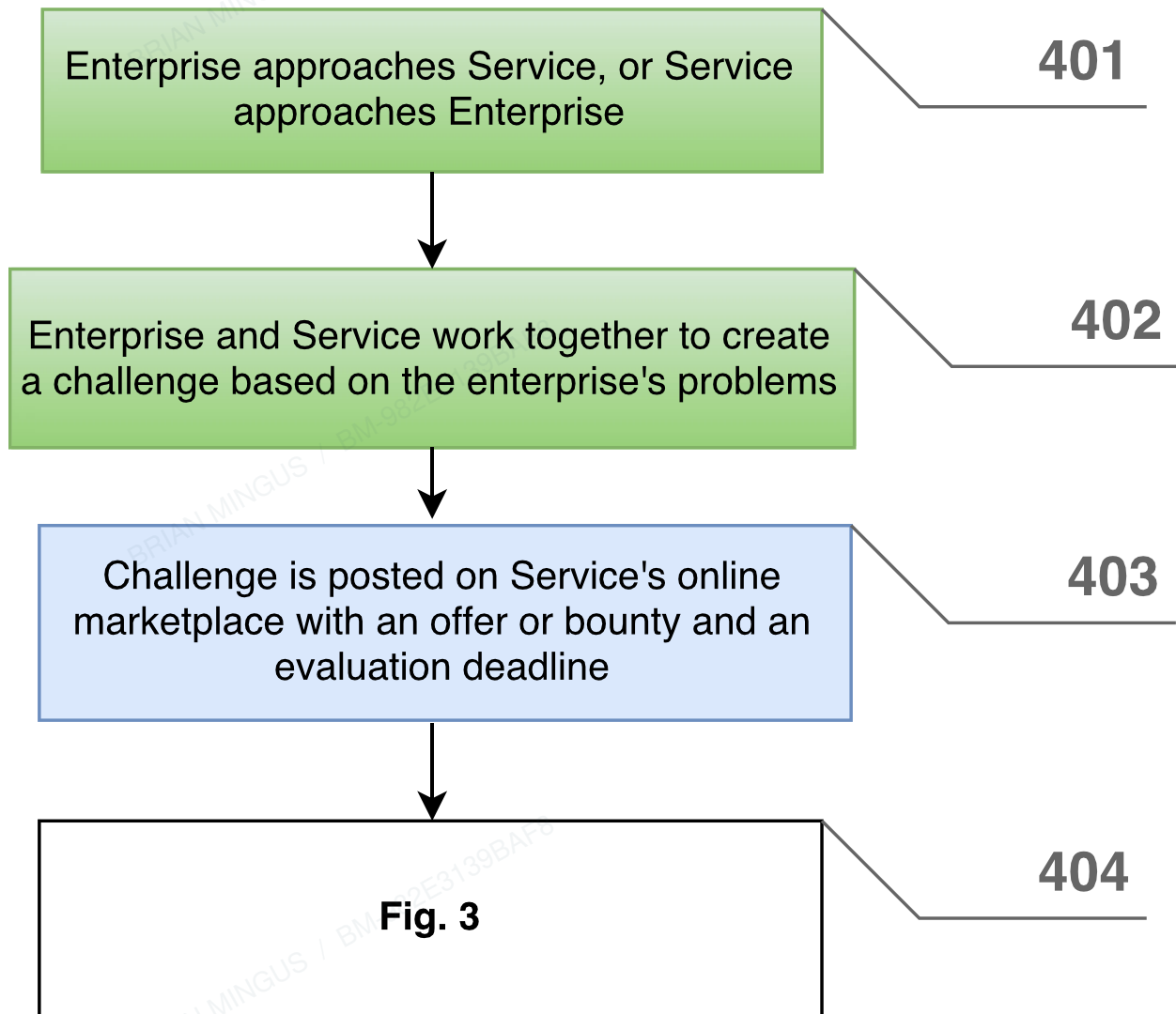


FIG. 4: Challenges are posted as project proposals on the Equity Crowdfunding Portal (101) or on the Self-Optimizing Online Marketplace (104) and completed as in Fig 3.

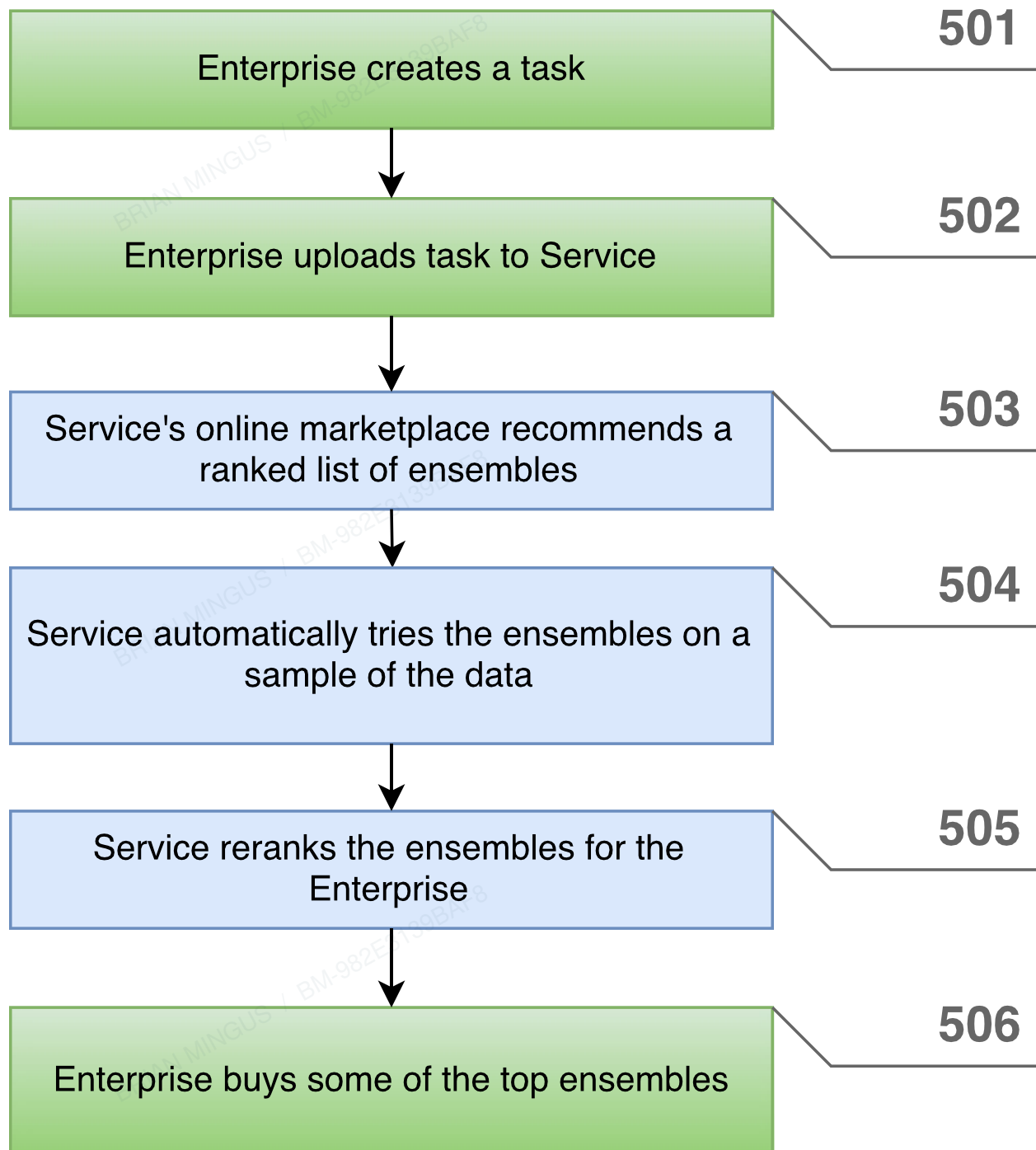


FIG. 5: In this scenario Enterprise creates a challenge, project proposal or task and uploads it to the Self-Optimizing Online Marketplace (104). The marketplace, which is a machine learning / artificial intelligence algorithm, analyzes the data and intelligently recommends a ranked list of ensembles (505) for Enterprise to buy (506).

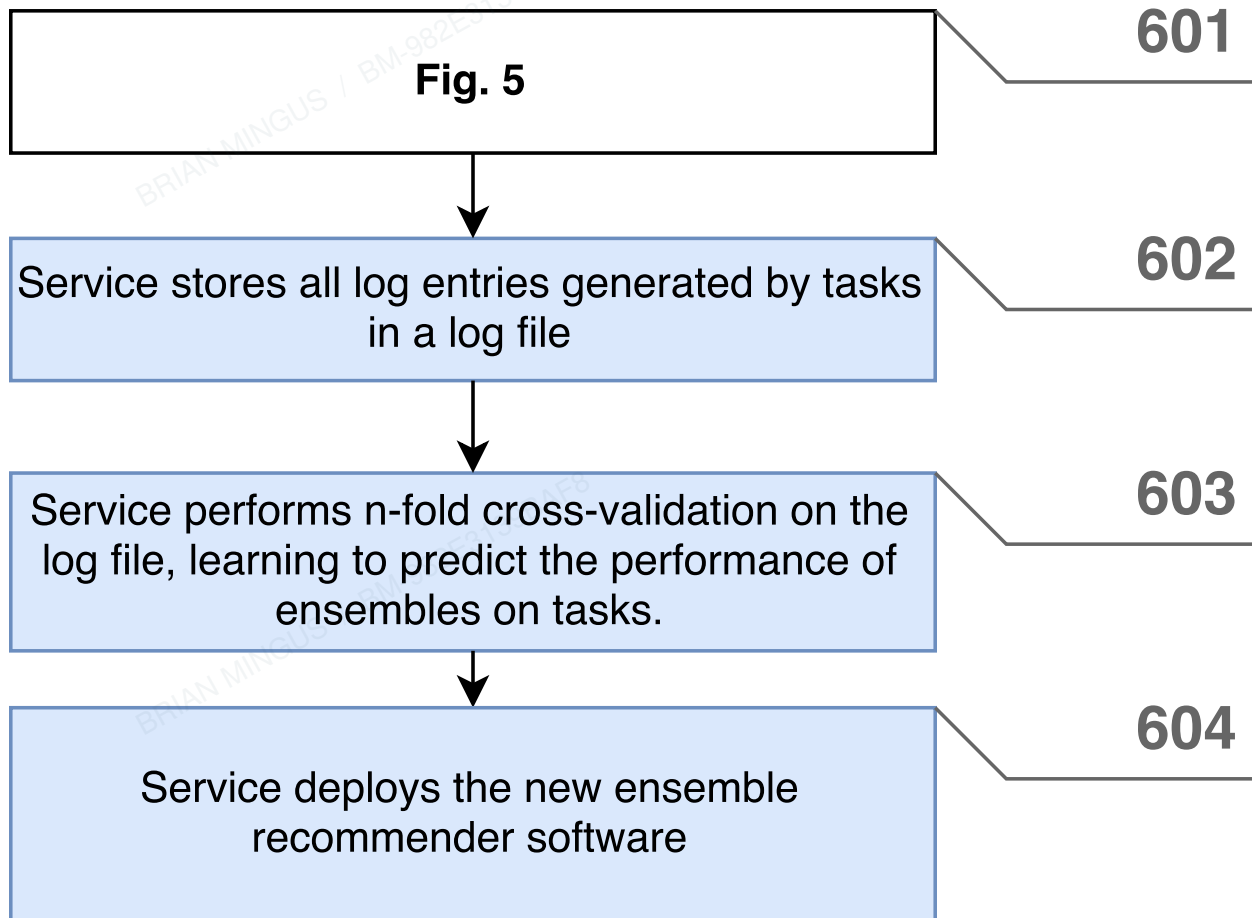


FIG. 6: Here Service (103), runs a Self-Optimizing Online Marketplace (104). (104) is a recommendation algorithm (a subset of machine learning / artificial intelligence algorithms that use so-called "implicit" information to generate recommendations for users, potentially in the context of "explicit" information such as a query). (103) stores all data (log entries) generated by all requests to process any data through any ensemble on (104) (602) and trains a recommendation algorithm on this data (603). The recommendation algorithm learns to predict which ensembles are appropriate for which task, or conversely, which tasks are appropriate for which ensemble, and in both cases, to what extent. The recommendation algorithm is trained "on-line", which means that as new log entries are generated the recommendation algorithm is re-deployed, having adapted to the new information, in near-real-time. (604)

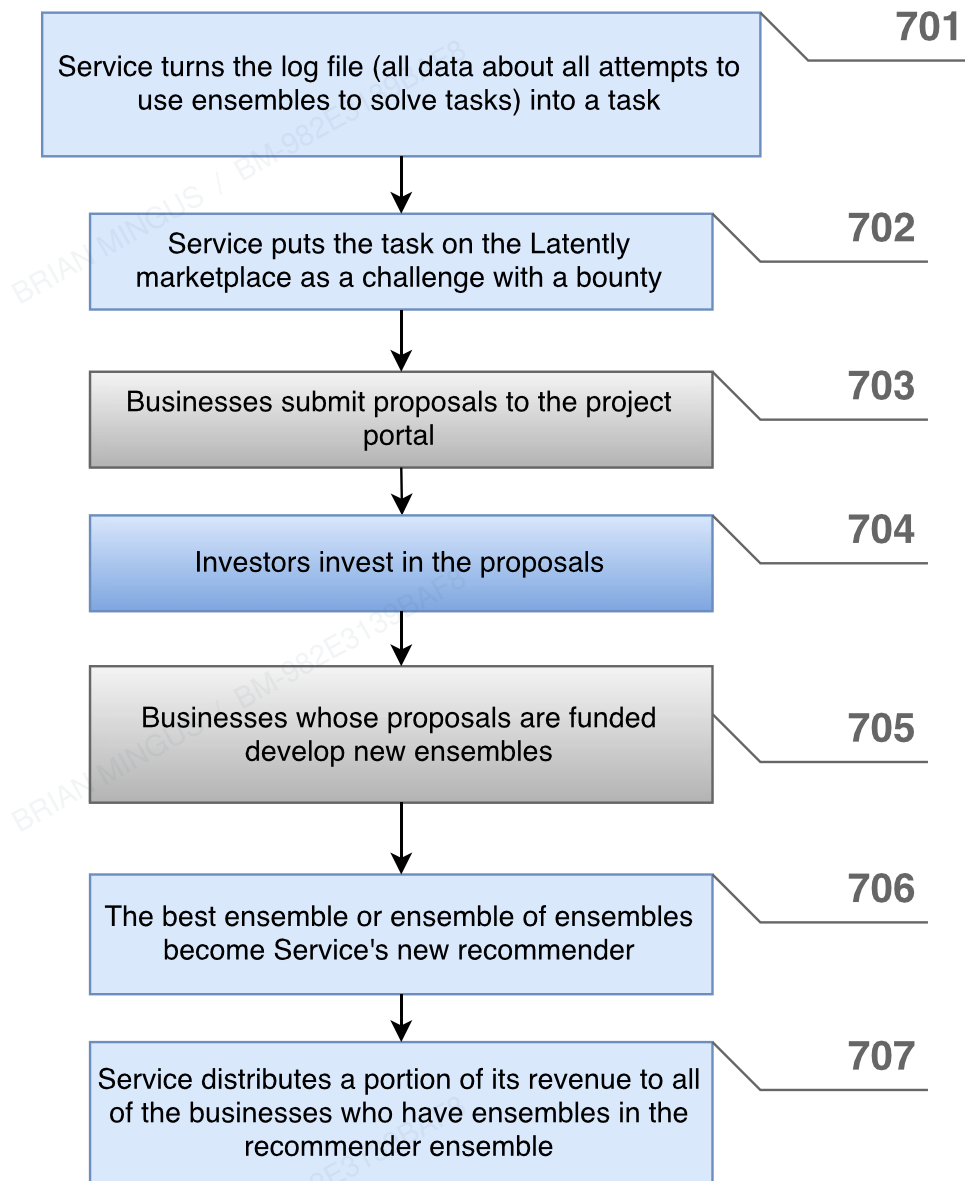


FIG. 7: Here Service "dogfoods" its recommendation algorithm that is its Self-Optimizing Online Marketplace, which is what makes the "Online Marketplace" Self-Optimizing. In order to accomplish this Service first turns the log file, which is a collection of all log entries, each of which corresponds to a task or challenge (701) and puts it on the Self-Optimizing Online Marketplace as a challenge (702). Businesses then submit proposals to the Equity Crowdfunding Platform (101) which propose to best solve this challenge. The procedure for interacting with the Service as indicated in Fig 3. then ensues, and assuming the top-ranked ensemble on the Self-Optimizing Online Marketplace (which could be the current ensemble) is chosen to be the new ensemble that runs the recommendation component of the Self-Optimizing Online Marketplace. (706) Finally, as usual, revenue is distributed as contractually agreed (707).

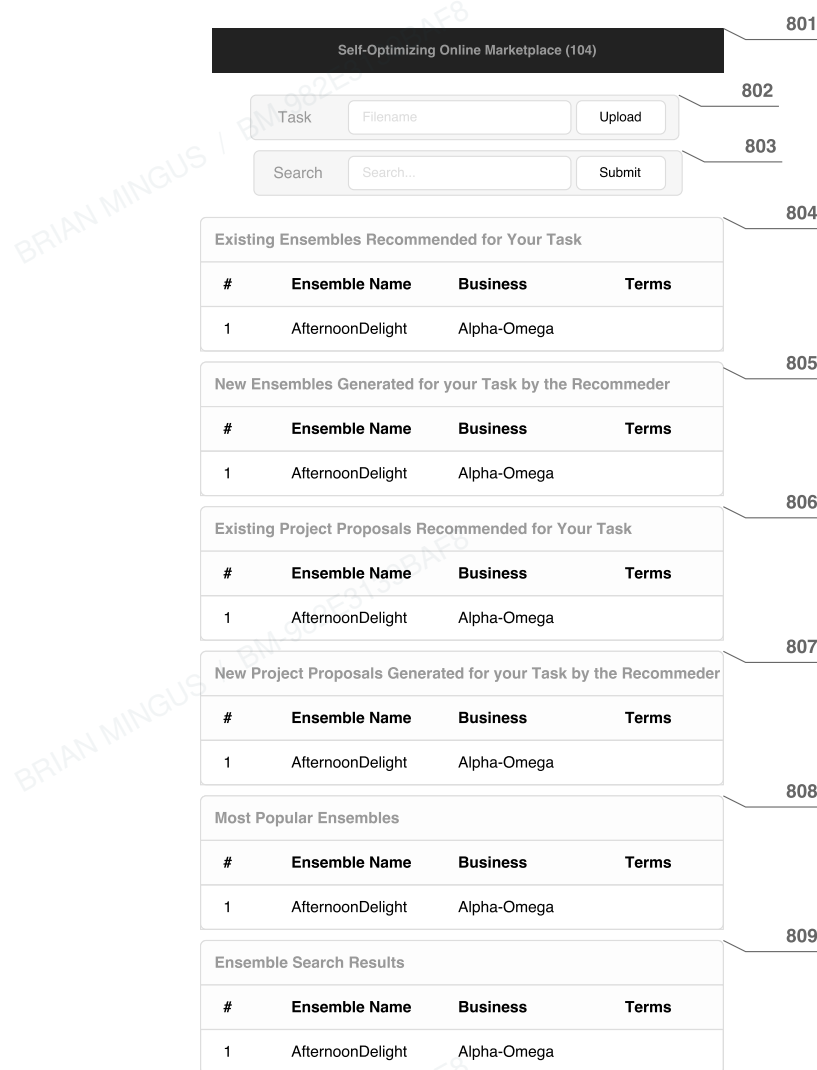


FIG. 8: UI for the marketplace recommender. Enterprises (107) upload tasks / datasets / challenges and the recommender: 1) Suggests existing ensembles for this task in order of how well they perform (804) 2) Uses the recommendation engine to suggest new ensembles that may perform well for your task, and sorts them by how well they perform (805) 3) Uses information on past project proposals and how well their ensembles have performed on tasks like this one to suggest existing project proposals which, if funded, might perform well for this proposal. (806) 4) Recommends new project proposals (i.e. Businesses, their investors and ensembles they might create (807) 5) Shows a list of the most popular ensembles and 6) If the search engine feature (803) is used, return corresponding search results (809).

BACKGROUND AND SUMMARY

[0002]

Today it is difficult for entrepreneurial software engineers, recreational programmers, data scientists and machine learning engineers to get paid for writing either specific or general purpose software (herein called models or ensembles). In the case of machine learning the problem is particularly difficult as the training of a model has a variable cost from tens to thousands of dollars or more. Without institutional support it is difficult or impossible to scale to the size of real-world datasets (such as the web), meaning that only large institutions can afford to work with the state of the art. These limitations keep the talent pool for the state of the art small and prevent entrepreneurial dreamers from leveling up their skills.

[0003] For enterprises who have lots of data and sufficient money, it can be difficult to hire the requisite talent needed to build out the infrastructure and supporting software needed to train and deploy the models needed to solve their machine learning and other problems. This is partly because these hires, before hire, did not have access to scalable infrastructure at a reasonable cost, preventing them from developing their skills and requiring costly on the job training. Further, once hired, it can be difficult for these engineers to ensure that the state of the art is continually deployed in-house. Even in the case when leading companies such as Google release whitepapers concerning recently developed in-house technology, it can often be years before they release implementations, if they are released at all. Thus, despite patent grants and business-friendly licenses the rest of the market has a significant lag.

[0004]

An additional hurdle preventing entrepreneurs from solving problems for the enterprise is that entrepreneurs typically lack access to knowledge about the enterprise environment in addition to knowledge of and access to the software required to integrate into the enterprise. Also, when

software is developed outside the context of a typical employment contract, it is often released under a license that allows for it to be incorporated into enterprise software without any compensation being returned to the person that developed the software. We solve these problems by allowing entrepreneurs to incorporate other software on our platform into their software ensembles and for all of the entrepreneurs who developed any of the software in the final ensemble of ensembles to get paid for its use. The amount paid can either be determined through terms stated in software licenses, or automatically by Service. For example, in the case of a machine learning ensemble (or mixture of experts) it is sometimes possible to come to an approximate guess as to how much a particular ensemble contributes to the final performance of a classifier. Thus, the share of proceeds that are owed to entrepreneurs can be determined automatically, and entrepreneurs can distribute portions of their shares back to their investors according to how much equity in their companies the investors own.

[0005]

Also, for both entrepreneurs and enterprises it can be difficult and costly to determine what software ensembles are most suitable for solving a particular enterprise challenge. The knowledge required to determine this is typically not specified, but rather exists only as enterprise software implementations and as know-how in the brains of the engineers that have created the software. We solve this problem by making our marketplace a recommendation engine. Enterprises upload challenges in the form of data to the marketplace and the marketplace recommends ensembles that best solve the challenge. For example, in the case of a typical classification or regression problem the marketplace can take into consideration all previous attempts at using ensembles on the marketplace to solve other classification and regression problems in order to recommend an ensemble for this particular problem. This solution naturally extends to allowing entrepreneurs to create ensembles that best solve enterprise challenges; if they are successful, the marketplace will learn from their success and will generalize to generating useful recommendations to enterprises in the future. Additionally, the marketplace learns from the lessons of the enterprise by training on the data contained in their challenges,

which could include information about the enterprise operating environment such as infrastructure, running software, data and meta-data thereof such as log files.

[0006]

The previous solution naturally gives rise to a meta problem which the marketplace itself faces: what is the best ensemble on the marketplace for creating the ensemble that recommends ensembles?

[0007]

This problem is solved making the machine learning and infrastructure challenges that the marketplace itself faces available as challenges on the marketplace. As the ensembles on the marketplace become better at solving problems in general, this will naturally extend to better solving the problems the marketplace faces. And the problems the marketplace faces are also indicative of problems that the rest of industry faces, resulting in an virtuous, value-creating learning cycle. And in doing so we allow a global community of entrepreneurs and investors to develop and invest in software solutions for the operating the marketplace itself and to thus receive a share of the profit and success of the overall marketplace as it brings entrepreneurs, investors and enterprises to the same table to solve enterprise problems on behalf of consumers.

[0008]

When bringing entrepreneurs and enterprises to the same table an issue that will inevitably arise concerns the security of the enterprise data. Here, we allow for a gradation of enterprise security. In the most secure case all code produced by entrepreneurs and their businesses is audited by professionals at Service and then downloaded by the Enterprise. The code is then carried into a secure airgapped room and run on a *copy* of the enterprise data. The results are then audited by the Enterprise, returned to the Service and then returned to the Entrepreneur. Note that in this case the Enterprise' data is a black box that the Entrepreneur never sees. This requires the community of Entrepreneurs developing ensembles to develop ensembles that are effective at reverse engineering a description of the data that is sufficient for solving challenges. Less secure methods are also possible, but to our knowledge the most secure method has never been tried.

[0009]

After being created ensembles are offered for sale on Service's online marketplace. The marketplace assures, among other things, that the businesses behind an ensemble (which could be numerous) are paid as contractually agreed. Additionally, there may be questions regarding intellectual property ownership. For example an ensemble may be forked without giving credit to the originator of the fork. In order to solve these problems we may employ blockchain technology. With the blockchain all elements of an ensemble, from the project proposal to the details of investments in the project to the binary and source code of all the ensembles being ensembled and more can be stored in the order that they occurred. The blockchain, while originally hosted by Service (and thus subject to tampering), can eventually be decentralized among the community of businesses who own ensembles on the marketplace and even the general public.

[0010]

The field of the invention is TODO

[0011]

A metric refers to any measure of a system and could include but is not limited to height, weight, size, speed or efficiency along an arbitrary dimension.

[0012]

Meta-data refers to data about data and may include, but is not limited to, information such as provenance, licenses, descriptions in natural language, descriptive statistics, version information, etc.

[0013]

A dataset refers to a collection of information. The information may be structured or unstructured. The dataset may also contain meta-data.

[0014]

A data lake refers to a collection of one or more datasets. It may optionally have meta-data.

[0015]

A model refers to a function that takes input and produces outputs. It may optionally have

modifiable internal state in addition to meta-data.

[0016]

An ensemble refers to a model that takes one or more models as inputs and produces outputs. It may optionally have modifiable internal state in addition to meta-data.

[0017]

An objective function refers to a function that takes a desired metric, the current value of a metric and optionally other information as inputs and produces a measure of the distance between the current and desired metrics as an output.

[0018]

A target refers to a pointer to a column or columns in a data lake that are to be used to indicate the desired outputs of an ensemble.

[0019]

A task refers to the combination of: 1) An optional description in either natural language or computer code 2) A pointer to an ensemble 3) A pointer to a data lake 4) A target in the data lake 5) An objective function 6) An optional set of parameters and hyperparameters to configure the ensemble 6) Optional meta-data.

[0020]

An optimization problem refers to a function that tries to minimize or maximize an objective function so as to change the modifiable parameters of an ensemble such that the ensemble performs better at a task or tasks.

[0021]

A log entry refers to a dataset that is composed of 1) A completed attempt to solve an optimization problem 2) Any optional snapshots of the modifiable state of the ensemble taken during the task 3) Any optional snapshots of the transient state of the ensemble taken during the task, such as activations in a neural network or the representation of the ensemble in memory and on disk 4) Any optional periodic metrics recorded during the task, including but not limited to performance metrics, metrics about the computing environment such as CPU, GPU or memory

utilization, etc. 5) Any optional meta-data.

[0022]

A log file is a data lake and refers to the collection of all log entries.

[0023]

A project proposal is an offer by a business to an investor to fund the creation of an ensemble and to purchase equity in said ensemble.

[0024]

An equity crowdfunding platform refers to a user interface where businesses post project proposals and one or more investors can review the proposals and decide which ones to invest in. Said transactions are optionally mediated by a broker-dealer. The platform results in investors funding project proposals resulting in the creation of ensembles by businesses in exchange for equity in their business.

[0025]

An online marketplace refers to a user interface in which ensembles are made available for sale and in which competitions can be run.

[0026]

A competition refers to an optimization problem on an online marketplace that is paired with a reward.

[0027]

The recommendations of ensembles are produced by an ensemble that is itself also on the marketplace, enabling businesses to create project proposals for improving the recommendation of ensembles by an ensemble.

[0028]

The method proposed in this invention allows an equity crowdfunding platform to fund the creation of IP that is both automatically and manually composed into ensembles for the purpose of metric optimization. The money paid by clients for access to ensembles that improve their metrics is distributed to the online marketplace. The online marketplace charges a fee for training

ensembles and for making them available on the online marketplace. The remainder of money paid by clients is distributed to all business who own IP in the ensemble. The businesses distribute said funds to their investors.

[0029]

Benefits to businesses include reducing the burden of creating and selling ensembles by 1) facilitating the presentation of project proposals to one or more potential investors simultaneously, 2) allowing investors to fund just the training of ensembles, 3) providing the ability to purchase computing power for building and training funded project proposals, 3) providing an online marketplace for selling training ensembles, 4) facilitating collaboration and innovation opportunity for ensemble refinements and improvements to meet enterprise needs 4) allowing for the continued receipt of returns for as long as the ensembles created are being sold on the marketplace.

[0030]

Benefits to investors include 1) simplifying the search, identification, evaluation, and comparison of project proposals, 2) lowering the funding thresholds per project proposal created from benefits to businesses 3) increasing the opportunity diversify investments by allowing more business investments.

[0031]

Benefits to enterprises include improving the optimization of business metrics by 1) allowing the opportunity to create challenges for the development of new ensembles 2) reducing or eliminating the burden of creating ensembles that improve the optimization of business metrics in-house, 3) increasing the availability and variety of ensembles for addressing business needs across the enterprise 3) facilitating the selection of ensembles through recommendations on the online marketplace 3) providing the option to continuously update ensembles as refinements and improvements become available on the marketplace.