

Genetank platform Terams and Module Diagrams

[Document subtitle]

December 11, 2017

Genetank Inc

Cambridge, MA

Genetank Platform Terms and Module Diagrams

# Platform Overview

GeneTank GWAS AI Model Sharing Platform is a distributed blockchain based genetic and health data sharing and AI prediction service platform. The overview of the platform is shown in the below picture.

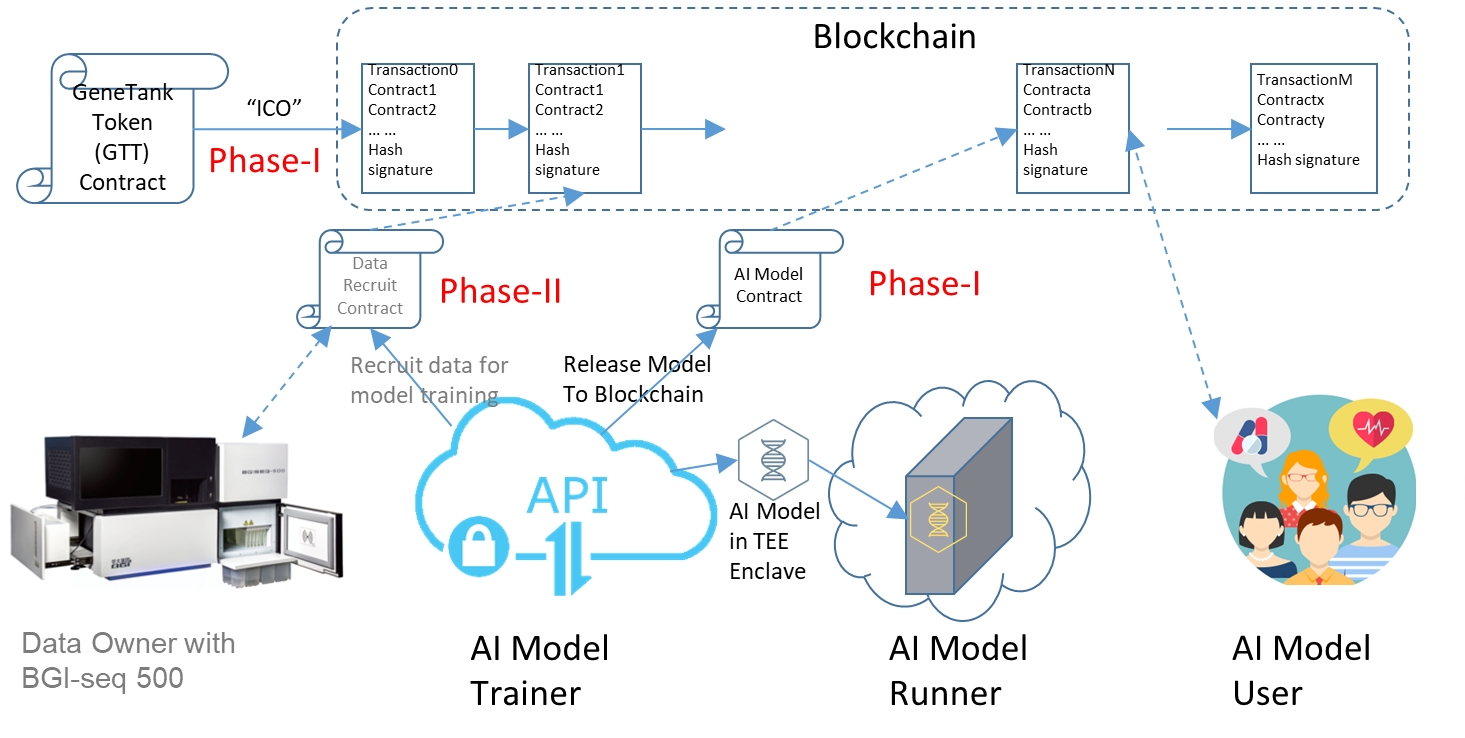


Figure. Overview of Genetank Platform

The platform uses Ethereum blockchain as the backbone to integrate all the data and services into an anonymous, secure, privacy protected, and open system.

The blockchain is a distributed and virtual database. The information on the blockchain is open to all the participants and is protected by cryptography and continuously verified by participants. Therefore, the information on the blockchain can be trusted.

The platform consists of a Genetank Token, a Model sharing, and a genetic/health data sharing subsystem. We will develop the Genetank Token and Model sharing subsystem in the phase-I. The genetic/health data sharing will be developed in phase-II.

From developer point of view, the platform has some subsystems which need to be developed and some others which are used for the development and platform running.

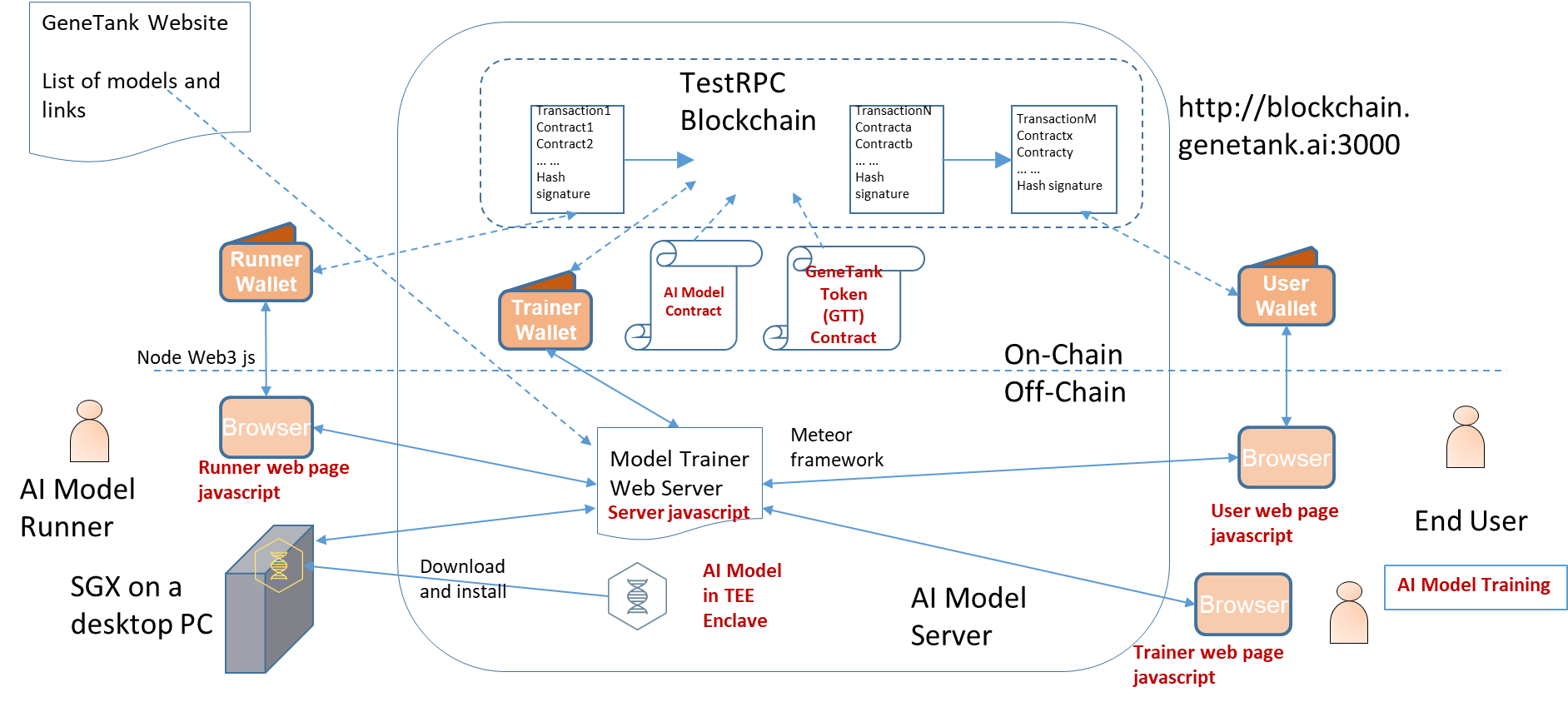


Figure. Development Framework

The subsystems to be developed include the AI Model, Genetank Token, Model sharing. The components that we are going to use are the TestRPC (private test Ethereum blockchain), wallet software (metamask --TBD), Genetank website, and browser (Chrome).

The Model sharing subsystem includes a model trainer web server and the backend and frontend html, javascripts, and other data for the model trainer, model runner, and end user.

# Term Definitions

Terms:

|  |  |  |  |
| --- | --- | --- | --- |
| Terms | Abbreviation | Descriptions |  |
| Model Trainer | Trainer | Genetank network participants who develop, train and share AI models on the Genetank platform |  |
| Model Runner | Runner | Genetank network participants who download, install and run the AI models developed by the model trainers. The runners provide service to the end users. |  |
| End User | User | Genetank network participants who use the AI model services provided by the runners. |  |
| Enclave Type | EnclaveT | A serial of enclaves which are developed and released by one model trainer. An enclave is a secure software which can run on a secure CPU. It is not possible to access an enclave’s code and data when it is running even with the system privileges. The trainer integrates an AI model into the enclave to ensure the safety of the data used by the model. |  |
| Enclave Version | EnclaveVer | A version of an enclave type. Each enclave version has a main version number and a subversion number. The main version number increases when there are feature changes. The subversion number increases when there some bugs fixes without new feature added.  Each enclave version must be registered and audited before it can be installed to a computer. |  |
| Enclave Instance | Enclave | An enclave instance is an enclave installed and run on a computer.  Each enclave instance must be attested and registered before it can provide any service to the users. |  |
| AI Model | Model | An AI model is a software which can predict certain phenotypes of a user from the user’s genetic and healthcare data.  An AI Model software includes some codes and data. |  |
| AI Model codes | ModelCode | The codes are used to execute the AI algorithms. The codes are integrated into an enclave version, released and installed along with the enclave version. |  |
| AI Model data | ModelData | The data are the parameters/coefficients of the AI model. The data are the results of the model training and are used in the prediction services provided by the enclave. |  |
| AI Model data version | ModelVer | The model data version (model version in short) is a set of data ready to be loaded into an enclave. An enclave instance must have one model version loaded before it can provide any service.  A new model version can be generated by more model training. A model version may support only a sub set of enclave versions. |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Module Diagrams

The top level of genetank AI model sharing platform module diagram is shown below.

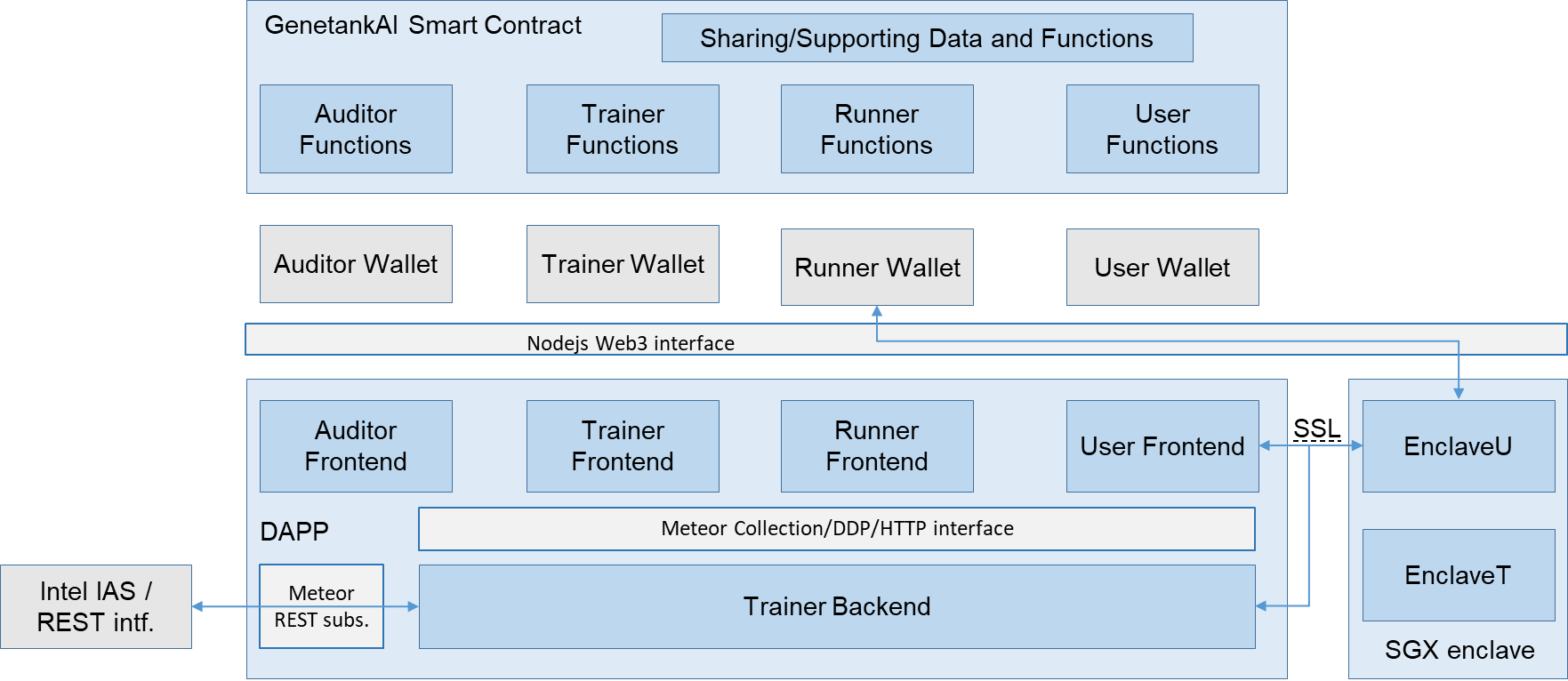


Figure. Genetank AI Model Sharing Module Top Level Diagram

The blue boxes are the modules we need to develop. The grey boxes are the tools or software we are going to use.

## 3.1 Trainer Module Diagram

## 3.2 Auditor Module Diagram

## 3.3 Runner Module Diagram

## 3.4 User Module Diagram

## 3.5 Enclave Module Diagram