

ОПИСАНИЕ И УСЛОВИЯ ПРЕДОСТАВЛЕНИЯ УСЛУГ

УСЛУГА «MODEL TRAINING»

НАИМЕНОВАНИЕ УСЛУГИ

1. Наименование Услуги: «MODEL TRAINING».
2. Настоящий документ содержит описание состава Услуги, ее базовой функциональности, возможных сопутствующих и дополнительных услуг, общего порядка подключения, изменения и отключения Услуги, условий предоставления и ограничений.

ИНФОРМАЦИЯ ОБ УСЛУГЕ

1. Краткое описание Услуги

Услуга предоставляется на базе инфраструктуры облачной платформы «AI Cloud».

Услуга предоставляет Заказчику среду разработки Jupyter Notebook, набор инструментов для хранения данных в Объектном хранилище S3, а также набор инструментов и библиотек для запуска задач по исполнению кода обучения моделей Машинного и Глубокого обучения на ресурсах суперкомпьютера “Кристофари” и мониторинга процесса обучения.

Для оказания услуги Заказчику необходимым условием является наличие на его площадке подключения к сети интернет, достаточного для эффективной загрузки данных на сервер, а также наличия собственных данных для обучения модели.

С помощью услуги Заказчик может вести разработку моделей и производить ускоренное обучение моделей на больших объемах данных, благодаря мощностям суперкомпьютера и высокопроизводительным графическим ускорителям.

Заказчику для успешной реализации задачи обучения моделей на больших объемах данных предоставляется возможность загрузки и хранения данных в Объектное хранилище S3, а также возможность подключения к этому хранилищу как из Jupyter Notebook’а, так и из кластера, на котором будет вычисляться задача обучения модели.

На Рисунке 1 приведена общая упрощенная схема взаимодействия с сервисом Model Training с удаленной площадки Заказчика (с указанием зон ответственности):

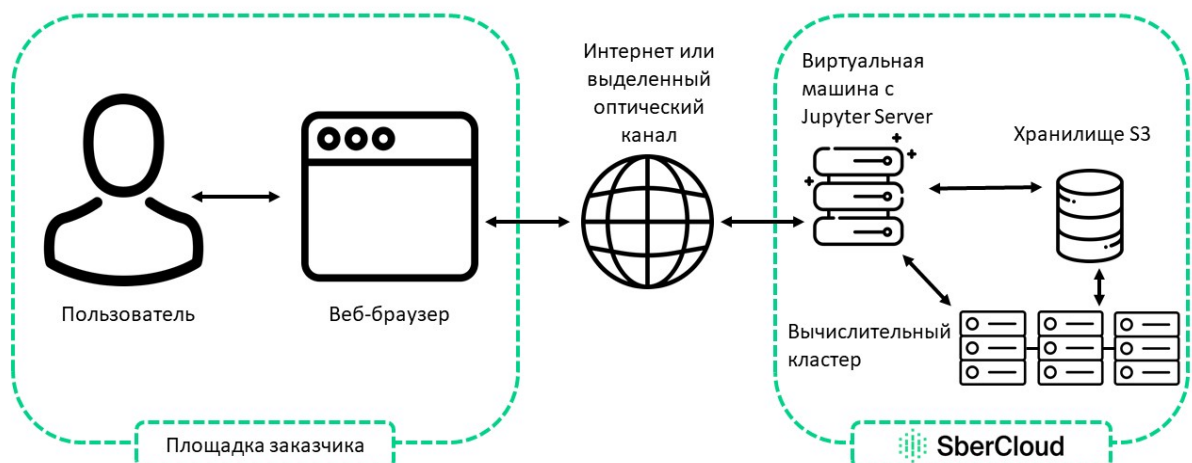


Рисунок 1 Схема взаимодействия заказчика с Сервисом обучения моделей машинного и глубокого обучения на мощностях кластера SberCloud – Model Training.

В зоне ответственности SberCloud – функционирование Серверов с развернутым Jupyter Server, функционирование вычислительного кластера, Объектного хранилища S3, а также всей прочей инфраструктуры.

В рамках Услуги Заказчик может самостоятельно отслеживать состояние заданий обучения модели.

Для подключения к Услуге Заказчик может выбрать один или несколько типов подключения:

- Подключение через общий канал Интернет (shared) предполагает логическое подключение к общему для всех Заказчиков Услуги каналу передачи данных. Скорость сетевого соединения для каждого Заказчика не является гарантированной и зависит от загрузки общего канала передачи данных (Услуга предоставляется по умолчанию).
- Подключение через прямой канал связи. Данный способ подключения позволяет обеспечить взаимодействие сетей Заказчика с сетью в облаке с помощью выделенных каналов связи стороннего провайдера. Опционально, с помощью данного сценария, к Услуге Заказчика может быть подключен альтернативный канал в сеть Интернет. Для данного подключения могут быть использованы выделенные каналы Заказчика, организованные с использованием «темной оптики» (Услуга оплачивается отдельно).

2. Обеспечение защиты инфраструктуры облачной платформы «AI Cloud»

В целях обеспечения кибербезопасности инфраструктуры облачной платформы «AI Cloud» реализовываются следующие меры и механизмы защиты:

- защита инфраструктуры облачной платформы и средств ее управления;
- защита консоли управления «AI Cloud»;
- очистка пользовательских данных .

Защита инфраструктуры облачной платформы и средств ее управления

Защита инфраструктуры облачной платформы «AI Cloud» и средств ее управления обеспечивается на следующих уровнях:

- на физическом уровне обеспечивается
 - размещение всего оборудования инфраструктуры в ЦОД, соответствующих требованиям надежности по категории Tier 3;
 - контроль и управление доступом к оборудованию;
 - наличие системы видеонаблюдения на объектах информатизации ЦОД.
- на сетевом уровне обеспечивается защита периметра инфраструктуры и ее сегментирование с использованием межсетевых экранов нового поколения (NGFW), осуществляющих в том числе выявление и предотвращение компьютерных атак;
- на инфраструктурном уровне обеспечивается:
 - антивирусная защита инфраструктуры с использованием антивирусных средств для облачных сред;
 - управление доступом к инфраструктуре с использованием средств двухфакторной аутентификации подключающихся к ней администраторов;
 - контроль действий привилегированных пользователей с использованием специализированных средств;
 - регулярный контроль и анализ защищенности инфраструктуры с использованием специализированных средств по выявлению уязвимостей в используемом ПО и его

некорректной конфигурации, влияющей на уровень защищенности ПО, с устранением выявленных уязвимостей и/или недостатков;

- сбор и анализ событий информационной безопасности.

Помимо этого, осуществляются периодические тестирования на проникновение и аудит информационной безопасности инфраструктуры облачной платформы «AI Cloud» с привлечением сторонних организаций. Выявленные в ходе соответствующего тестирования и/или аудита недостатки устраняются по факту их выявления.

Защита консоли управления «AI Cloud»

Защита консоли управления «AI Cloud» обеспечивается на уровне приложений с использованием специализированного межсетевого экрана уровня приложений (Web Application Firewall). Помимо этого, осуществляются регулярные сканирования консоли на наличие актуальных уязвимостей и его периодические тестирования на проникновение с привлечением сторонних организаций. Выявленные уязвимости и/или недостатки устраняются по факту их выявления.

Очистка пользовательских данных

Перед выделением и предоставлением доступа в «AI Cloud» к местам памяти для временного хранения и обработки данных под очередную задачу (произведения вычислений, обучения модели и т.п.) осуществляется полная очистка пользовательских данных, ранее хранимых в указанных областях памяти в ходе выполнения предыдущих задач.

Пользователям «AI Cloud» предоставляется доступ только к выделенным для них областям памяти контейнера и объектного хранилища (S3). При этом на время пользования Услугой доступ к указанным областям памяти других субъектов запрещен.

3. Условия хранения данных в Объектном хранилище S3

Хранение, использование и тарификация хранения и использования данных в Объектном хранилище S3 осуществляется в рамках соответствующей Услуги по Договору.*

*В связи с тем, что Услуга Объектное хранилище S3 в настоящий момент Исполнителем не введена в качестве самостоятельной Услуги по Договору, то Стороны руководствуются следующим:

- Заказчику для потребления Услуги Model Training предоставляется доступ к Объектному хранилищу S3 в размере, необходимом для хранения данных обучаемой модели.
- Для того, чтобы воспользоваться услугой доступа к Объектному хранилищу S3 Заказчику необходимо в Личном кабинете активировать доступ к ней посредством проставления «галочки» в соответствующем поле «Требуется S3 хранилище».
- Объём потреблённой в течение Отчётного периода услуги Объектное хранилище S3 рассчитывается Исполнителем в соответствии с данными АСИ.
- Стоимость Услуги Объектное хранилище S3 определен в Приложении № 7 к Договору.
- Оплата Услуги Объектное хранилище S3 осуществляется Заказчиком в порядке постоплаты на основании выставленного Исполнителем счёта и при условии подписанного Сторонами Акта. Акт и счёт выставляются в порядке, установленном ст. 4 Договора.
- Указанные выше условия доступа к Услуге Объектное хранилище S3 действует до момента ввода Объектного хранилища S3 в качестве самостоятельной Услуги по Договору. С указанного момента порядок потребления такой Услуги определяется в соответствующем Приложении к Договору.

4. Использование сервиса Model Training

Создание, конфигурация и запуск задач на обучение осуществляется напрямую Заказчиком.

5. Техническое описание решения

2.4.1. Программная платформа

Услуга реализуется средствами Jupyter Server и Jupyter Notebook. Посредством него и программных библиотек пользователь имеет возможность запускать задачи на вычисление на кластере.

2.4.2. Аппаратная платформа

Вычисления и обсчет задач осуществляется на предоставляемой Заказчику в рамках услуги области кластера (суперкомпьютера) Кристофари.

2.4.3 Технические особенности и ограничения:

Скорость загрузки данных на площадку Исполнителя ограничена пропускной способностью канала доступа в Интернет из инфраструктуры Заказчика до облака SberCloud, а также скоростью чтения данных с СХД Исполнителя.

Общие значения параметров услуги «Model Training»

Описание	Мин. значение	Макс.значение
Размер предоставляемого хранилища S3	1 Гб.	20 Тб.
Количество утилизируемых в рамках вычисления задачи GPU	1 GPU	В соответствии с количеством свободных GPU, отображаемом в Личном кабинете

ТАРИФИКАЦИЯ УСЛУГИ

3.1. Возможные виды тарификации Услуги:

3.1.1. Динамическая тарификация (Pay as you go) (п. 2.4.2. Договора);

3.2. Стоимость Услуги формируется в зависимости от количества GPU, на которых происходило вычисление задачи, времени, в течение которого вычислялась задача, объема зарезервированного Заказчиком Объектного хранилища S3.

3.3. Момент начала списания денежных средств – с момента запуска обучения модели.

ИНЫЕ УСЛОВИЯ, ПРИМЕНИМЫЕ К УСЛУГЕ

4.1. Возможные виды подключения / изменения / удаления Услуг:

4.1.1. Подключение Услуги посредством подписания Заказа;

4.1.2. Подключение Услуг посредством совершения действий на Портале.

4.2. Возможный порядок расчётов по услуге:

4.2.1. Предварительная оплата;

4.2.2. Постоплата (на основании отдельно заключенного письменного бланка Заказа).

4.3. Возможные способы оплаты / порядок пополнения баланса:

4.3.1. Оплата в безналичном порядке на основании выставленного Исполнителем счёта;

4.3.2. Оплата посредством электронных средств платежа.

4.4. В связи с характером потребления Услуги, а также объёмов, которых она может достигнуть в Отчётный период, Стороны установили, что в случае заключения с Заказчиком соглашения о применении Постоплаты (пп. 4.2.2. настоящего Приложения) Заказчик выбирает лимит в пределах Отчётного периода, по достижении которого Услуги оказываются на основании соответствующего обращения уполномоченного лица в Контактный Центр и, по требованию Исполнителя, предоставления гарантийного письма.