



**AMBER  
DATA**

**ОПИСАНИЕ ФУНКЦИОНАЛЬНЫХ  
ХАРАКТЕРИСТИК МОДУЛЯ CDP ДЛЯ  
ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ  
ПО «ВИ КРУТИЛКА» (AmberData)**

Сведения о правообладателе

Общество с ограниченной ответственностью «Амбердата»  
127521, г. Москва, 12-й проезд Марьиной рощи, д.9, корп. 1, пом. VI, ком. 1  
ИНН/КПП: 7703096683/771701001  
тел.: (499) 938-43-78  
e-mail: [info@amberdata.ru](mailto:info@amberdata.ru)

## Оглавление

1. Термины и определения .....	3
2. Введение.....	3
3. Основные функции модуля .....	3
4. Компоненты и структура модуля CDP.....	4
5. Описание компонентов модуля CDP.....	4
5.1. CDP-UI.....	4
5.2. CDP-API.....	5
5.3. CDP-Syncer .....	5
5.4. CDP-Mango.....	5
5.5. CDP-Flow-constructor .....	5
5.6. CDP-EASY (External Amberdata Syncers).....	5
6. Технологическая основа .....	6
7. Требования к техническим специалистам для разработки и эксплуатации платформы	6
8. Информация, необходимая для установки Системы .....	7
9. Вход в систему .....	7

## 1. Термины и определения

DMP	Data Management Platform, программно-аппаратный комплекс для сбора, анализа и управления данными.
Аудитории	Подготовка данных о профилях пользователей для медиа-партнеров в роли потребителей данных, управление выгрузкой и лицензиями сегментов во внешние системы медиа-партнеров (DSP, Exchange и т.п.).
API	Набор программных методов для интеграции с внешними системами
REST, REST API RESTful API	(англ. Representational State Transfer Application Programming Interface) – стандартизированный метод взаимодействия приложений в рамках архитектуры REST

## 2. Введение

Модуль Customer Data Platform (CDP) является частью ПО «ВИ КРУТИЛКА».

ПО «ВИ КРУТИЛКА» (публичное фирменное наименование – AmberData, DMP AmberData, CDP AmberData, Amber Tag Manager) – многофункциональная технологическая платформа по сбору, обработке, хранению и управлению данными об интернет-пользователях. Предметом сбора и анализа в DMP AmberData становятся профили пользователей.

## 3. Основные функции модуля

Существуют 3 основных типа данных о профилях пользователей:

- First-Party Data - собственные данные, собранные и используемые на своих площадках.
- Second-Party Data - собственные данные, которые вы используете, но их собрали и подготовили для использования сторонние системы.
- Third-Party Data - данные третьих лиц, которые можно использовать при нехватке своих - для увеличения охвата, или для уточнения собственных данных.

Технически для сбора данных на страницах сайтов и/или в рекламных сообщениях размещают специальные коды тэгов, которые обеспечивают идентификацию профилей пользователей с помощью уникального идентификатора. В дальнейшем происходит синхронизация этих уникальных идентификаторов профилей с уникальными идентификаторами наших партнеров и получение имеющихся у них данных о профилях этих пользователей – таким образом, формируются Third party Data.

Модуль Customer Data Platform (CDP) AmberData предназначен для обработки данных партнеров о профилях пользователей, приведению их к правильному виду и дальнейшему использованию в системах потребителей данных.

## 4. Компоненты и структура модуля CDP

- Модуль CDP
  - UI – пользовательский интерфейс, включающий следующие возможности:
    - Просмотр/фильтрация результирующей агрегированной таблицы профилей пользователей;
    - Визуализация/сравнение отфильтрованных профилей, сгруппированных по сконфигурированным ключам;
    - Создание сегментов на основе фильтров, применяемых к отфильтрованной таблице.
  - API – провайдер данных о сущностях, используемых в процессе работы модуля CDP, предоставляющий программный интерфейс для:
    - CDP-UI;
    - CDP-Syncer;
    - CDP-Mango;
    - CDP-Flow-constructor;
    - CDP-EASY;
    - Внешних систем.
  - Syncer – Компонент, реализующий загрузку данных и создание сущностей в clickhouse, в соответствии с декларативным конфигурированием наборов данных в CDP-API.
  - Mango – API для запросов к данным clickhouse, с учетом прав владения объектами в CDP. Реализует динамический Swagger, согласно спецификации CDP-API.
  - Flow-constructor – Компонент с интерфейсом для пользователей, позволяющий посмотреть/отредактировать пайплайн преобразования данных в clickhouse.
  - EASY - External Amberdata Syncers - Компоненты для синхронизации сформированных сегментов профилей пользователей со сторонними системами.

## 5. Описание компонентов модуля CDP

### 5.1. CDP-UI

Пользовательский интерфейс, включающий следующие возможности:

- Просмотр/фильтрация результирующей агрегированной таблицы профилей пользователей;
- Визуализация/сравнение отфильтрованных профилей, сгруппированных по сконфигурированным ключам;
- Создание сегментов на основе фильтров, применяемых к отфильтрованной таблице.

К компоненту прилагается инструкция пользователя и документация, с требованиями для запуска и конфигурации приложений.

Поставляется в виде Docker image с обфусцированным javascript/typescript кодом приложения.

## 5.2. CDP-API

Провайдер данных о сущностях, используемых в процессе работы модуля CDP, предоставляющий программный интерфейс для:

- CDP-UI;
- CDP-Syncer;
- CDP-Mango;
- CDP-Flow-constructor;
- CDP-EASY;
- Внешних систем.

Для работы API необходима база PostgreSQL.

К компоненту прилагается документация, с требованиями для запуска и конфигурации приложений и спецификация API.

Поставляется в виде Docker image с скомпилированным/обфусцированным кодом приложения.

## 5.3. CDP-Syncer

Компонент, реализующий загрузку данных и создание сущностей в clickhouse, в соответствии с декоративным конфигурированием наборов данных в CDP-API.

К компоненту прилагается документация, с требованиями для запуска и конфигурации приложений.

Поставляется в виде Docker image со скомпилированным кодом приложения.

## 5.4. CDP-Mango

API для запросов к данным clickhouse, с учетом прав владения объектами в CDP.

Реализует динамический Swagger, согласно спецификации CDP-API.

К компоненту прилагается документация, с требованиями для запуска и конфигурации приложений.

Поставляется в виде Docker image со скомпилированным кодом приложения.

## 5.5. CDP-Flow-constructor

Компонент с интерфейсом для пользователей, позволяющий посмотреть/отредактировать пайплайн преобразования данных в clickhouse.

К компоненту прилагается документация, с требованиями для запуска и конфигурации приложений.

Поставляется в виде Docker image со скомпилированным кодом приложения.

## 5.6. CDP-EASY (External Amberdata Syncers)

Компоненты для синхронизации сформированных сегментов профилей пользователей со сторонними системами.

К компоненту прилагается документация, с требованиями для запуска и конфигурации приложений.

Поставляется в виде Docker image со скомпилированным кодом приложения.

## **6. Технологическая основа**

Платформа AmberData построена на продуктах и технологиях с открытым исходным кодом, распространяемых под лицензией Apache 2.0.

При разработке модулей и компонентов решения используются следующие языки программирования:

- Java;
- Scala;
- JavaScript/TypeScript;
- Python;
- Clojure;
- SQL;
- PromQL;
- Bash.

Основные компоненты, отвечающие за обработку больших объемов данных, реализованы с использованием языка Java 8, Scala, Clojure.

Загрузка, обработка и хранение данных осуществляется на платформах Apache Hadoop, содержащая, как минимум, следующие компоненты:

- HDFS;
- YARN;
- Hive;
- Spark v > 3;
- Airflow;
- PrestoDB / PrestoSQL;
- Jupyter.

Оркестрация приложений в докере реализована на базе Kubernetes.

Обеспечена защита данных на следующих уровнях:

- Application (OAuth2),
- HTTPS для доступа из вне к DMP-UI.

## **7. Требования к техническим специалистам для разработки и эксплуатации платформы**

Требования к программистам и системным администраторам:

- 1) Знание технологий:

- Kubernetes;
- Apache Hadoop;
- Spark;
- Hive;
- PrestoDB;
- Airflow;
- Docker;
- Nginx;
- PostgreSQL;
- Clickhouse.

2) Знание языков программирования:

- Python;
- Bash;
- SQL.

3) Отличное знание Linux.

Требования к специалисту техподдержки сервиса:

- Знание Javascript/html на высоком уровне,
- Знание технических особенностей браузеров и механизмов работы с куками,
- Уверенное владение Python 3.6 - 3.8.

## 8. Информация, необходимая для установки Системы

Клиентская часть приложения не требует проведения установки. Работа осуществляется посредством web-браузера.

## 9. Вход в систему

Для осуществления входа в проект необходимо выполнить следующие действия:

1. Запустить браузер.
2. Указать следующий URL <https://my.amberdata.ru/> в адресной строке браузера.
3. В открывшемся окне авторизации (Рисунок 1) ввести логин и пароль, предоставленные администратором системы.
4. Нажать кнопку «**Войти**»

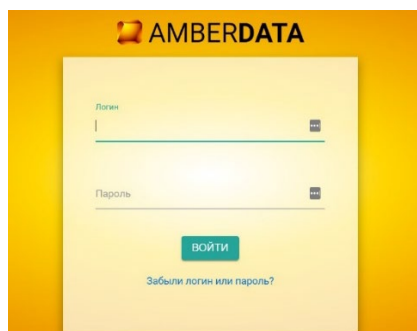


Рисунок 1 – *Окно авторизации*

В случае успешной аутентификации пользователю открывается экранная форма «**Последние добавленные коды**» и «**Последние добавленные категории**».