

Формат выгрузки данных (минимальный перечень) 2021-2025

ПК Mediameter

В данном документе описан формат выгрузки телеметрических данных с устройств респондентов, полученных с помощью системы Mobile peoplemeter

1. Формат поставки результатов совпадений сигнатур по телесмотрению и радиослушанию

1.1. Данные по телесмотрению и радиослушанию находятся в поддиректории `\tv\yyyy\mm\`, где yyyy – год, mm – месяц (номер).

1.2. Данные представлены в формате CSV (текстовый файл, каждая строка соответствует одному событию, значения разделены точкой с запятой)

1.3. Формат строки события:

DeviceID; Login; Start DateTime; Finish DateTime; Channel; Reliability; Mean offset; Offset deviation
где

###	Tun	Описание
DeviceID	string	Идентификатор устройства
Login	string	Индивидуальный логин панелиста (последний зарегистрированный на устройстве)
Start datetime	string	Дата начала события (в формате DD.MM.YYYY HH:MM:SS) – местное время с учетом серверной коррекции часов на устройстве
Finish datetime	string	Дата окончания события (в формате DD.MM.YYYY HH:MM:SS) – местное время с учетом серверной коррекции часов на устройстве
Channel	int	Идентификатор канала (при этом, 0 – означает специальное событие: интервал в течение которого устройство работало непрерывно)
Reliability	int	Параметр, который означает надежность определения события (0...100)
Mean offset	int	Среднее время сдвига распознанного события по сравнению с эталонным эфирным временем события на сервере записи (сек.)
Offset deviation	int	Стандартное отклонение времени сдвига распознанного события по сравнению с эталонным эфирным временем события на сервере записи (сек.)

1.4. Пример поставки результатов совпадений сигнатур по телесмотрению и радиослушанию:

```
8604786671;user12345;30.07.2019 09:00:27;31.07.2019 08:55:58;0;100;0;0
8604786671;user12345;30.07.2019 10:13:34;30.07.2019 12:17:59;1111;75;0;0
8604786671;user12345;30.07.2019 12:17:59;30.07.2019 12:43:35;16;55;-5;1
8604786671;user12345;30.07.2019 12:43:35;30.07.2019 12:58:09;20;4;-27;2
```

1.5. Файлы данных

Файлы архива "сырых" данных имеют стандартное название вида:

20151018_104031_mpm_XXXXXX_**20151014_20151018**.csv

где

20151018_104031 - дата и время формирования отчета

20151014 - начальная дата отчета

20151018 - конечная дата отчета

Отчеты формируются ежедневно. Каждый отчет включает в себя данные за последние 4 дня. Таким образом, в архиве, как правило, на каждую дату есть несколько файлов, которые содержат в себе информацию за одну и ту же дату. При этом, наиболее актуальные данные (включающие допоставки за последние 3 дня) будут содержаться в файле с более поздним сроком формирования отчета.

1.6. Список ID каналов и радиостанций дан в Приложении №1

1.7. Параметры Reliability, Mean_offset, Offset_deviation могут быть использованы в системе пост-обработки данных, для фильтрации ложных срабатываний и более точного определения версии ТВ канала.

2. Формат выгрузки данных по контактам с мобильными приложениями

2.1. Данные по контактам респондентов с мобильными приложениями находятся в поддиректории `\package\yyyy\mm\`, где yyyy – год, mm – месяц (номер).

2.2. Данные представлены в формате CSV (текстовый файл, каждая строка соответствует одному событию, значения разделены точкой с запятой)

2.3. Формат строки события имеет вид:

Date;Time;DeviceID;Login;Length;Package

где

###	Tun	Описание
Date	string	Дата события (год-месяц-день в формате YYYY-MM-DD)
Time	string	Время события (час:минуты:секунды в формате HH:MM:SS) - местное время с учетом серверной коррекции часов на устройстве
DeviceID	string	Идентификатор устройства
Login	string	Индивидуальный логин панелиста (последний зарегистрированный на устройстве)
Length	int	Длительность контакта в секундах
Package	string	Название пакета приложения (см. app_db)

2.4. Пример строки данных:

2018-09-04; 21:16:09; 2687471523; user123456; 15; com.google.android.apps.photos;

2018-09-04; 21:16:24; 2687471523; user123456; 100; com.whatsapp;

2018-09-04; 21:27:54; 2687471523; user123456; 207; com.vkontakte.android;

2.5. Файлы архива контактов с мобильными приложениями имеют стандартное название вида:

20151018_104031_mpm_XXXXXX_**20151014_20151018**.csv

где

20151018_104031 - дата и время формирования отчета

20151014 - начальная дата отчета

20151018 - конечная дата отчета

Отчеты формируются ежедневно. Каждый отчет включает в себя данные за последние 4 дня. Таким образом, в архиве, как правило, на каждую дату есть несколько файлов, которые содержат в себе информацию за одну и ту же дату. При этом, наиболее актуальные данные (включающие допоставки за последние 3 дня) будут содержаться в файле с более поздним сроком формирования отчета.

3. Формат выгрузки данных по работе экрана

3.1. Данные по работе экрана находятся в поддиректории \screen\yyyy\mm\, где yyyy – год, mm – месяц (номер).

3.2. Данные представлены в формате CSV (текстовый файл, каждая строка соответствует одному событию, значения разделены точкой с запятой)

3.3. Формат строки события имеет вид:

Date;Time;DeviceID;Login;Duration

где

###	Tun	Описание
Date	string	Дата <i>начала</i> события включения экрана (год-месяц-день в формате YYYY-MM-DD)
Time	string	Время <i>начала</i> события включения экрана (час:минуты:секунды в формате HH:MM:SS)
DeviceID	string	Идентификатор устройства
Login	string	Индивидуальный логин панелиста (последний зарегистрированный на устройстве)
Duration	int	Длительность работы включенного экрана (сек)

3.4. Пример строки данных:

2016-07-13; 12:23:57; 2687471523; user123456; 65

3.5. Файлы архива данных гео-локации имеют стандартное название вида:

20151018_104031_mpm_XXXXXX_**20151014_20151018**.csv

где

20151018_104031 - дата и время формирования отчета

20151014 - начальная дата отчета

20151018 - конечная дата отчета

Отчеты формируются ежедневно. Каждый отчет включает в себя данные за последние 4 дня. Таким образом, в архиве, как правило, на каждую дату есть несколько файлов, которые содержат в себе информацию за одну и ту же дату. При этом, наиболее актуальные данные (включающие допоставки за последние 3 дня) будут содержаться в файле с более поздним сроком формирования отчета.

4. Формат выгрузки данных по состоянию батареи

4.1. Данные по состоянию батареи находятся в поддиректории **\battery\yyyy\mm**, где yyyy – год, mm – месяц (номер).

4.2. Данные представлены в формате CSV (текстовый файл, каждая строка соответствует одному событию, значения разделены точкой с запятой)

4.3. Формат строки события имеет вид:

Date;Time;DeviceID;Login;State;Charge

где

###	Tun	Описание
Date	string	Дата <i>начала</i> события включения экрана (год-месяц-день в формате YYYY-MM-DD)
Time	string	Время <i>начала</i> события включения экрана (час:минуты:секунды в формате HH:MM:SS) - местное время с учетом серверной коррекции часов на устройстве
DeviceID	string	Идентификатор устройства
Login	string	Индивидуальный логин панелиста (последний зарегистрированный на устройстве)
State	int	Уровень заряда батареи от 0 до 1000 (1000 = 100%)
Charge	int	Подключение и тип зарядного устройства (0 – зарядка не подключена, 1 – АС зарядка, 2 – зарядка от USB)

4.4. Пример строки данных:

2016-07-13; 12:23:10; 2687471523; user123456; 510; 1

2016-07-13; 12:24:32; 2687471523; user123456; 530; 1

4.5. Файлы архива данных состояния батареи имеют стандартное название вида:

20151018_104031_mpm_XXXXXX_**20151014_20151018**.csv

где

20151018_104031 - дата и время формирования отчета

20151014 - начальная дата отчета

20151018 - конечная дата отчета

Отчеты формируются ежедневно. Каждый отчет включает в себя данные за последние 4 дня. Таким образом, в архиве, как правило, на каждую дату есть несколько файлов, которые содержат в себе информацию за одну и ту же дату. При этом, наиболее актуальные данные (включающие допоставки за последние 3 дня) будут содержаться в файле с более поздним сроком формирования отчета.

5. Формат выгрузки данных по распознаванию аудио-вотермарок

5.1. Данные по распознаванию аудио-вотермарок находятся в поддиректории **\watermark\yyyy\mm**, где yyyy – год, mm – месяц (номер).

5.2. Данные представлены в формате CSV (текстовый файл, каждая строка соответствует одному событию, значения разделены точкой с запятой)

5.3. Формат строки события имеет вид:

Date;Time;DeviceID;Login;WatermarkValue;Package

где

###	Тип	Описание
Date	string	Дата события (год-месяц-день в формате YYYY-MM-DD)
Time	string	Время события включения экрана (час:минуты:секунды в формате HH:MM:SS) - местное время с учетом серверной коррекции часов на устройстве
DeviceID	string	Идентификатор устройства
Login	string	Индивидуальный логин панелиста (последний зарегистрированный на устройстве)
WatermarkValue	int	Числовое значение вотермарки (0 ... 2147483648)
Package	string	Имя пакета, в том случае, если распознавание произошло в момент работы зарегистрированного приложения

5.4. Пример строки данных:

2018-09-04; 13:18:26; 703144029; user234567; 12345; com.google.android.youtube

5.5. Файлы архива данных по распознанным вотермаркам имеют стандартное название вида:

20151018_104031_mpm_XXXXXX_**20151014_20151018**.csv

где

20151018_104031 - дата и время формирования отчета

20151014 - начальная дата отчета

20151018 - конечная дата отчета

Отчеты формируются ежедневно. Каждый отчет включает в себя данные за последние 4 дня. Таким образом, в архиве, как правило, на каждую дату есть несколько файлов, которые содержат в себе информацию за одну и ту же дату. При этом, наиболее актуальные данные (включающие допоставки за последние 3 дня) будут содержаться в файле с более поздним сроком формирования отчета.

6. Формат выгрузки данных по установленным приложениям

6.1. Данные по распознаванию аудио-вотермарок находятся в поддиректории \inst_app\yyyy\mm\, где yyyy – год, mm – месяц (номер).

6.2. Данные представлены в формате CSV (текстовый файл, каждая строка соответствует одному событию, значения разделены точкой с запятой)

6.3. Формат строки имеет вид:

Date;DeviceID;Login;Package

где

###	Тип	Описание
Date	string	Дата события (год-месяц-день в формате YYYY-MM-DD)
DeviceID	string	Идентификатор устройства
Login	string	Индивидуальный логин панелиста (последний зарегистрированный на устройстве)
Package	string	Имя пакета приложения установленного на телефон на текущую дату (кроме системных приложений)

6.4. Пример строки данных:

2018-09-11;2011501020;ssp6841154;com.avito.android;

2018-09-11;2011501020;ssp6841154;com.viber.voip;

2018-09-11;2011501020;ssp6841154;ru.yandex.searchplugin;

2018-09-11;2011501020;ssp6841154;com.facebook.katana;

2018-09-11;2011501020;ssp6841154;ru.ok.android;

2018-09-11;2011501020;ssp6841154;com.google.android.youtube;

6.5. Файлы архива данных по установленным приложениям имеют стандартное название вида:

20151018_104031_mpm_XXXXXX_**20151014_20151018**.csv

где

20151018_104031 - дата и время формирования отчета

20151014 - начальная дата отчета

20151018 - конечная дата отчета

Отчеты формируются ежедневно. Каждый отчет включает в себя данные за последние 4 дня. Таким образом, в архиве, как правило, на каждую дату есть несколько файлов, которые содержат в себе информацию за одну и ту же дату. При этом, наиболее актуальные данные (включающие допоставки за последние 3 дня) будут содержаться в файле с более поздним сроком формирования отчета.

7. Формат выгрузки базы данных мобильных приложений

7.1. Для сопоставления названия пакета и мобильного приложения на ftp доступна выгрузка базы мобильных приложений (app_db). База ежедневно пополняется и включает в себя все приложения, которыми когда-либо пользовались респонденты.

7.2. Данные по всем мобильным приложениям (пакетам) обновляются раз в несколько дней и расположены в поддиректории \app_db, в файле *apps_db.csv*

7.3. Данные представлены в формате CSV (текстовый файл, каждая строка соответствует одному событию, значения разделены точкой с запятой).

7.4. Формат строки файла *apps_db.csv* имеет вид:

Имя пакета;Название;Категория;Разработчик;Ссылка;Дополнительно;Версия ОС;Возрастная категория;Загрузки;Оценок;Рейтинг;Выбор редакции;Лучший разработчик;Цена (RUB);Цена (USD)
где

Имя пакета	Имя пакета приложения (например, com.google.android.calculator)
Название	Название приложения в Google Play
Категория	Категория приложения
Разработчик	Разработчик приложения
Ссылка	Ссылка на сайт разработчика
Дополнительно	Дополнительная информация
Версия ОС	Минимальная поддерживаемая версия ОС
Возрастная категория	Возрастная категория (например, 12+)
Загрузки	Количество установок приложения
Оценок	Количество оценок приложения
Рейтинг	Рейтинг приложения (например: 4,5)
Выбор редакции	Текстовое поле (Да/Нет)
Лучший разработчик	Текстовое поле (Да/Нет)
Цена (RUB)	Цена в RUB (если есть)
Цена (USD)	Цена в USD (если есть)

7.5. Пример строки данных:

com.google.android.calculator;Калькулятор;Инструменты;Google Inc.;http://www.google.com/;;5.0+;3+;5000000;10110;4,4;Нет;Да;;

ПОСТАВКА ИНЫХ ДАННЫХ, КРОМЕ УКАЗАННЫХ ВЫШЕ ПРОИЗВОДИТСЯ ПО ДОГОВОРЕННОСТИ СТОРОН.