

# КАК ЗАГРУЗИТЬ И ОПУБЛИКОВАТЬ ДАТАСЕТ

Чтобы попасть на экран «**Добавление датасета**», необходимо:

- Войти в свой личный кабинет, нажав на иконку в заголовке веб-сервиса;
- Нажать «**Добавить**» и в выпадающем списке выбрать «*Датасет*».

Для отправки датасета на модерацию необходимо:

- Заполнить «*Описание датасета*» и «*Технические характеристики*»;
- Выбрать значения из списков в разделе «*Классификация датасета*»;
- Загрузить «*Условия использования*»;
- Добавить «*Ссылку на датасет*»;
- Нажать «**Отправить на модерацию**».

## Описание датасета

Оптимальная длина описания — до 2000 символов с пробелами.

В описании датасета необходимо раскрыть следующую информацию:

- Общее описание, что из себя представляет датасет;
- Источник данных для датасета;
- Особенности данных в датасете;
- Подробно раскрытые отличия платной и бесплатной версии;
- Другую информацию.

## Пример описания датасета

- В датасете Microsoft Malware находится информация о различных устройствах и факте поражения их вредоносным ПО.
- Источником является открытая информация от компании Microsoft.
- В датасете есть информация об устройствах различного размера — информация о компьютерах, планшетах, телефонах и т.п.
- Расширенная версия содержит в 4.6 раза больше данных, чем демонстрационная.

## Технические характеристики

Описание технических характеристик датасета.

В этом блоке необходимо раскрыть следующие пункты:

- Размер датасета;
- Форматы файлов в датасете;
- Полнота данных, наличие «кривых» данных;
- Распределение данных;
- Другие характеристики.

## Классификация датасета

Необходимо выбрать подходящие значения из справочников «*Отрасль*», «*Тип данных*», «*Применение*».

**Отрасль** — для какой отрасли применим датасет (например, «*Медицина*»).

**Тип данных** — данные какого типа представлены в датасете (например, .csv).

**Применение** — для каких целей можно применять датасет (например, «*Бинарная классификация*»).

В случае, если Вы не нашли подходящего значения, выберите наиболее подходящий вариант из существующих и напишите обращение в Службу поддержки с просьбой добавить значение в справочник.

## Условия использования

Файл, описывающий условия использования датасета. Может содержать текст лицензии.

Допустимые форматы файла: .pdf, .doc, .docx, .md.

Допустимый размер: не более 128 Мб.

## Ссылка на датасет

Разработчик должен загрузить ссылку на файлообменник, где хранятся файлы с датасетом.

**Загружается ограниченная демо версия датасета.** Продажа полной версии датасета осуществляется при содействии сотрудников ML-Place.

## Модерация датасета

После того, как датасет был загружен, он отправляется на модерацию. Посмотреть список датасетов на модерации можно в личном кабинете в разделе «*Мои датасеты*».

Модерация занимает до 72 часов и происходит в рабочие дни.

С модерации датасет может вернуться на доработку или же может быть опубликованным в каталоге датасетов. В случае, если датасет вернулся на доработку, его карточку можно увидеть в личном кабинете в разделе «**Мои датасеты**» — «**На доработке**». По нажатию на кнопку «*Редактировать*» можно перейти на экран редактирования карточки датасета.

**Комментарии о том, что нужно доработать, приходят разработчику на электронную почту.**