

**Эффективный и масштабируемый
сбор обучающих данных с
помощью краудсорсинга: практика**

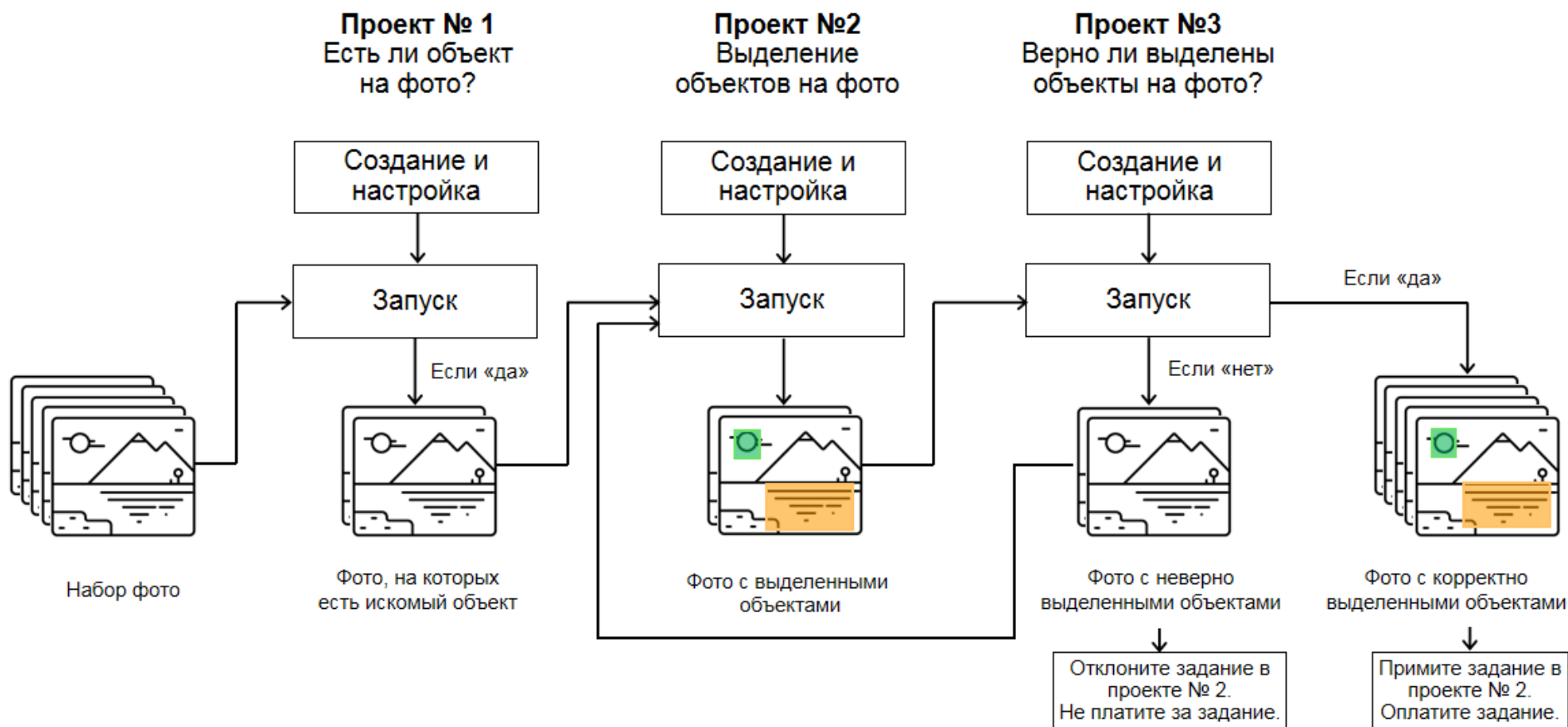
Создание и настройка проектов по сбору данных

Инструкция

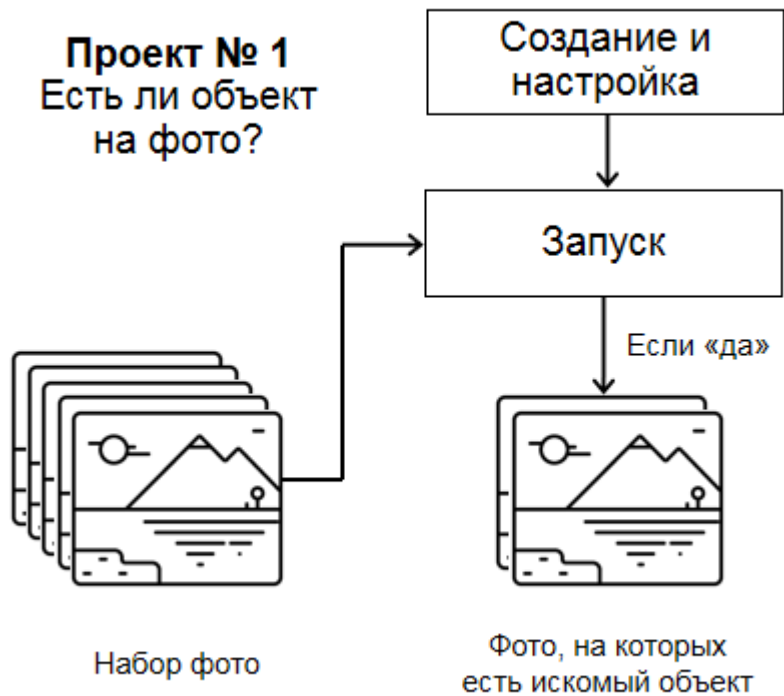
Содержание

Схема построения процесса	3
Проект № 1. Есть ли объект на фото?	4
Создание проекта	5
Создание пула	8
Подготовка и загрузка файла с заданиями	10
Получение результатов	13
Проект № 2. Выделение объектов на фото	14
Создание проекта	15
Создание пула	17
Подготовка и загрузка файла с заданиями	20
Проект № 3. Верно ли выделены объекты на фото?	21
Создание проекта	22
Создание пула	24
Подготовка и загрузка файла с заданиями	26
Получение результатов	27
Загрузка результатов проверки	28
Проверка ответов онлайн и анализ результатов	29

Схема построения процесса

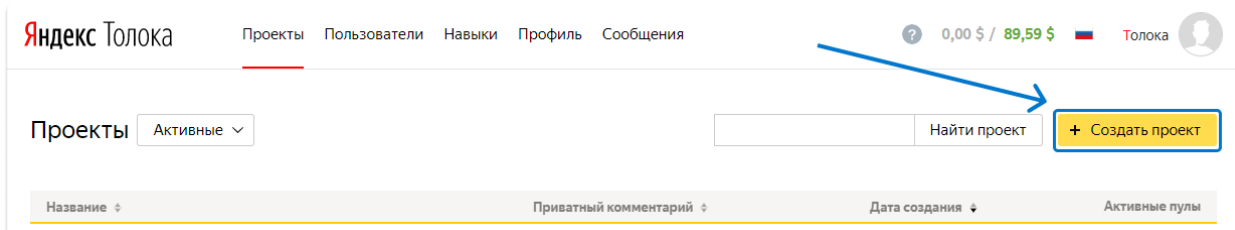


Проект № 1. Есть ли объект на фото?

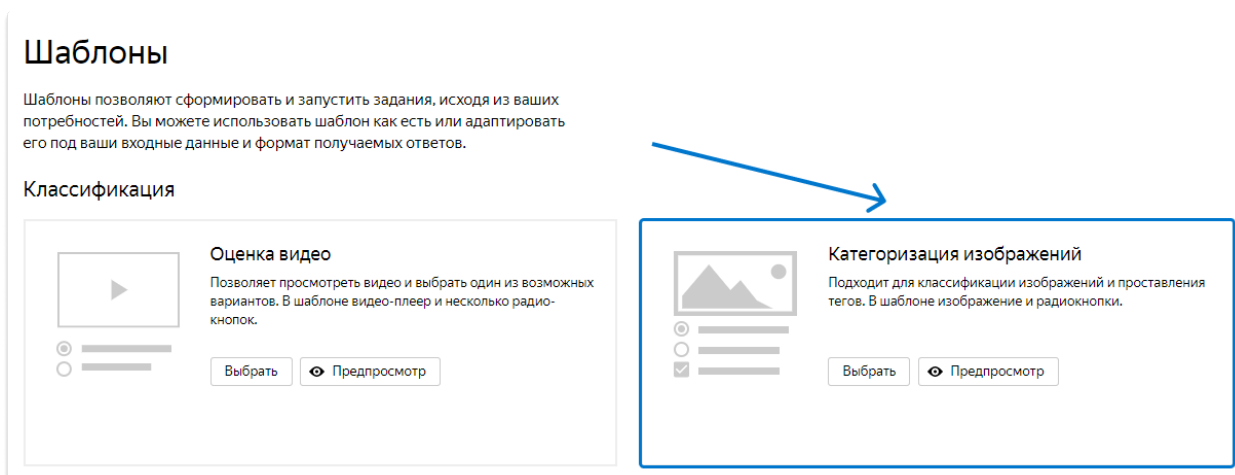


Создание проекта

1. Нажмите кнопку **+ Создать проект**

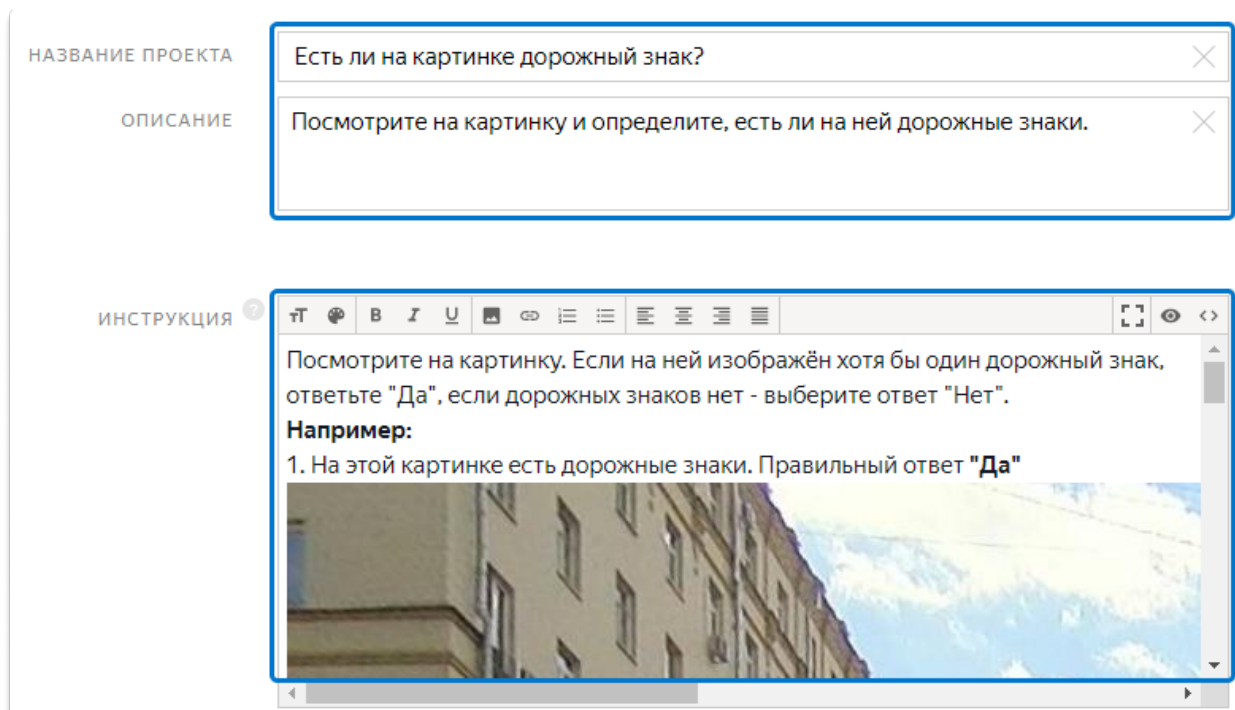


2. Выберите шаблон **Категоризация изображений**.



3. Дайте проекту понятное название и краткое описание.

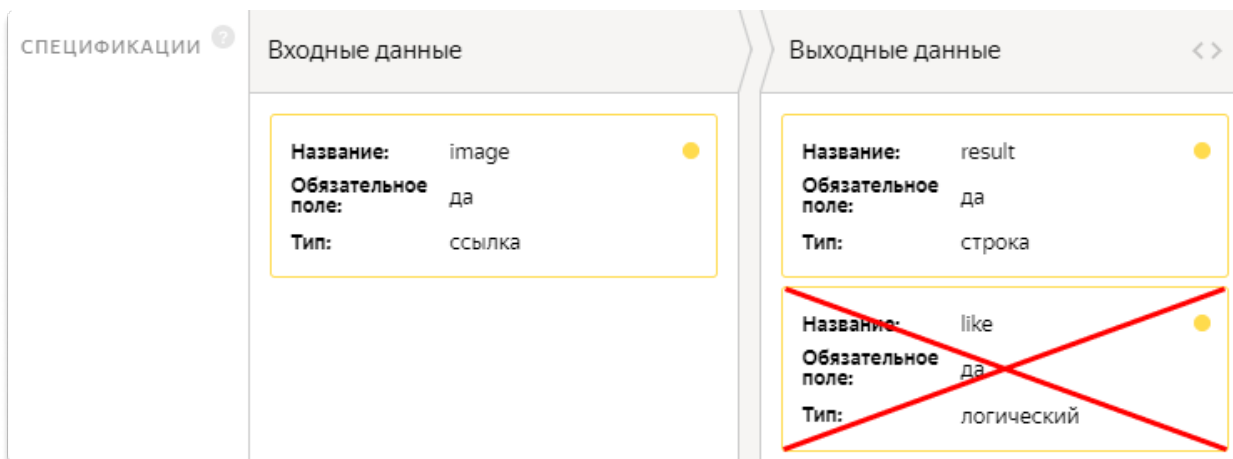
Важно: Их увидят реальные исполнители Толоки в списке доступных заданий.



4. Напишите простую и понятную инструкцию.

5. Определите поля для ВХОДНЫХ И ВЫХОДНЫХ ДАННЫХ в блоке **Спецификации**.

- Во входное поле **image** мы передадим ссылки на картинки, чтобы показать их исполнителям.
TSV-файл со ссылками мы впоследствии [загрузим в пул](#).
- В выходное поле **result** с типом строка запишем вариант ответа, выбранный исполнителем.
- В поле **like** в шаблоне передаётся ответ на вопрос «*Нравится фото?*». В нашем проекте этой галочки не будет, поэтому и выходное поле для неё нам не нужно — удалим его.



6. Создайте интерфейс задания в блоке HTML.

- Удалите строку с компонентом чекбокс:

```
{{field type="checkbox" name="like" label="Do you like the photo?" hotkey="q"}}
```

- Добавьте вопрос: есть ли на картинке искомый объект. Например:
<div>Есть ли на картинке дорожный знак?</div>
- Измените label на свои варианты ответов. Например:

```
{{field type="radio" name="result" value="OK" label="Да" hotkey="1"}}  
{{field type="radio" name="result" value="BAD" label="Нет" hotkey="2"}}  
{{field type="radio" name="result" value="404" label="Ошибка загрузки"  
hotkey="3"}}
```

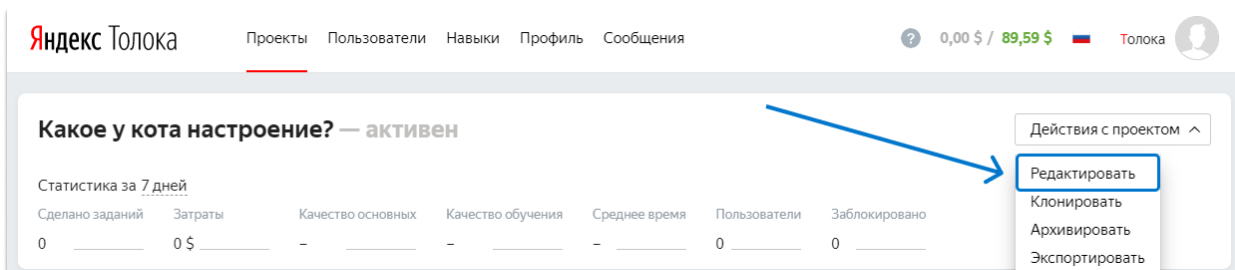


7. Блоки CSS и JavaScript оставьте без изменений.
8. Click the **Preview** button to see the performer's view of the task.
9. Нажмите кнопку **Предпросмотр**, чтобы увидеть получившееся задание.
*Вы увидите на странице четыре задания со стандартными картинками.
Количество заданий на странице вы сможете задать в пуле.*



10. Отметьте радиобаттоны в предпросмотре и убедитесь, что задание можно выполнить.
11. Чтобы сохранить изменения в проекте, нажмите **Сохранить**.

Совет. Если нужно что-то исправить, отредактируйте проект и сохраните его.



Создание пула

1. Нажмите **Добавить пул**.
2. Дайте пулу любое удобное название и описание. Они доступны только вам.
3. Укажите параметры пула:
 - Цена за страницу заданий (например, \$0.01).
 - Перекрытие. Это количество пользователей, которые выполнят одну и ту же страницу с заданиями. Например, 3.
 - Время на выполнение страницы заданий. Например, 600 секунд.

Цена за страницу заданий

ЦЕНА 0.01 НАДБАВКА 0.005

+ Динамическое ценообразование

Перекрытие

ПЕРЕКРЫТИЕ 3

ДИНАМИЧЕСКОЕ ПЕРЕКРЫТИЕ Выкл.

Настройки заданий

ВРЕМЯ НА ЗАДАНИЕ 600 СРОК ДЕЙСТВИЯ 2020-09-02

ЧАСТОТА ПОКАЗА КАПЧИ Нет ИНТЕРВАЛ ЗАКРЫТИЯ 0

СОХРАНЯТЬ ПОРЯДОК ЗАДАНИЙ Нет ПРИОРИТЕТ 0

КОНТЕНТ ДЛЯ ВЗРОСЛЫХ Есть ОТЛОЖЕННАЯ ПРИЕМКА Нет

ДЕДЛАЙН

4. Добавьте **Фильтры** для отбора исполнителей, которым будет доступно задание.
 - Чтобы ваше задание было доступно только исполнителям, владеющим русским языком, выберите готовый набор **Русскоязычные исполнители**.

Перекрытие

ПЕРЕКРЫТИЕ 3

ДИНАМИЧЕСКОЕ ПЕРЕКРЫТИЕ Выкл.

Настройки заданий

ВРЕМЯ НА ЗАДАНИЕ 600 СРОК ДЕЙСТВИЯ 2020-09-02

ЧАСТОТА ПОКАЗА КАПЧИ Нет ИНТЕРВАЛ ЗАКРЫТИЯ 0

СОХРАНЯТЬ ПОРЯДОК ЗАДАНИЙ Нет ПРИОРИТЕТ 0

КОНТЕНТ ДЛЯ ВЗРОСЛЫХ Есть ОТЛОЖЕННАЯ ПРИЕМКА Нет

Требуемые навыки

Вычисляемые

Багфикс версия операционной системы

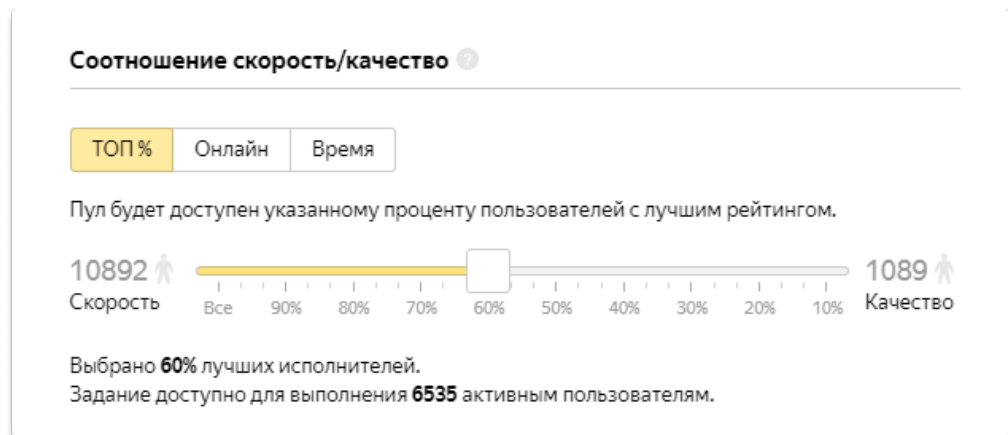
Багфикс версия ядра агента

Добавить фильтр

Скопировать настройки из...

Создать навык




- По [желанию](#), укажите процент лучших по рейтингу исполнителей в блоке **Соотношение скорость/качество**.



Соотношение скорость/качество ?

ТОП % Онлайн Время

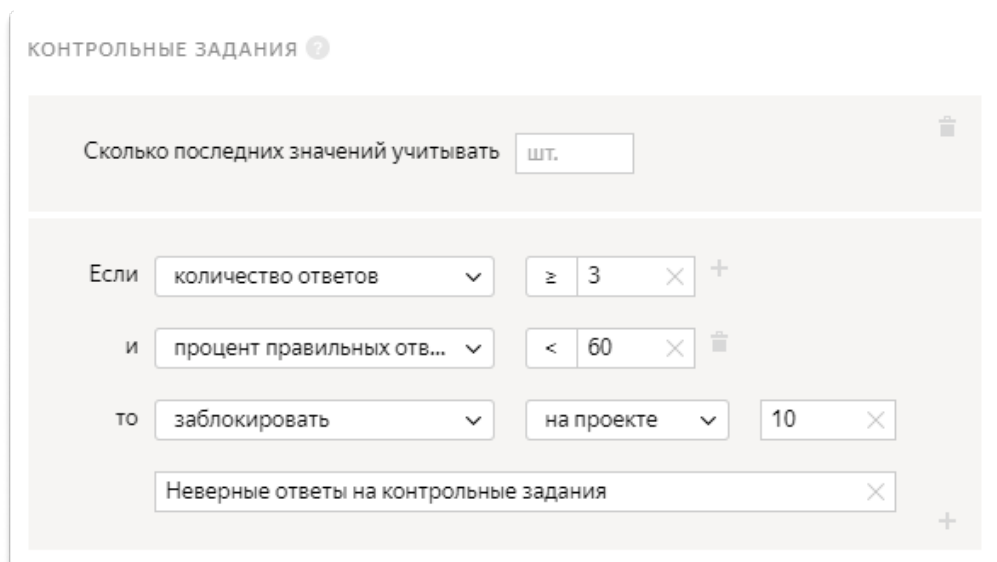
Пул будет доступен указанному проценту пользователей с лучшим рейтингом.

10892   1089 

Скорость Все 90% 80% 70% 60% 50% 40% 30% 20% 10% Качество

Выбрано **60%** лучших исполнителей.
Задание доступно для выполнения **6535** активным пользователям.

5. [Настройте контроль качества](#). Выберите блок **Контрольные задания**, чтобы заблокировать исполнителей, которые дают неверные ответы. Например:



КОНТРОЛЬНЫЕ ЗАДАНИЯ ?

Сколько последних значений учитывать

Если

и

то

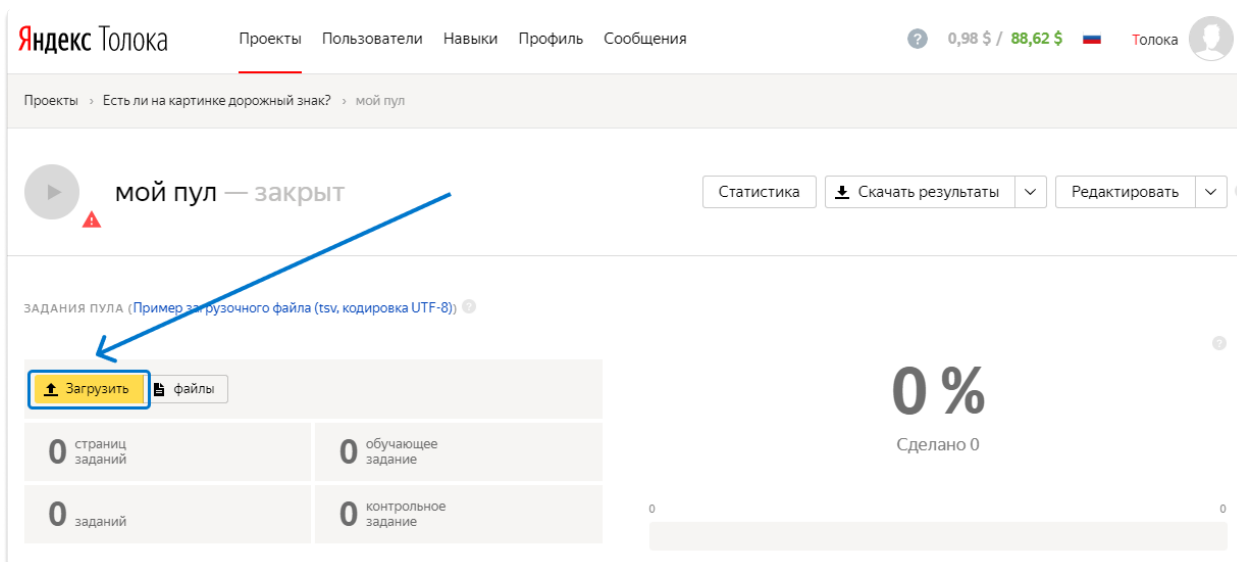
Правило сработает, когда исполнитель выполнит 3 контрольных задания в пуле. Если процент правильных ответов окажется меньше 60, т.е. из 3 заданий он ответит верно всего на 1, то он будет заблокирован на проекте на 10 дней. Правило будет пересчитано после каждого следующего контрольного задания.

По желанию добавьте другие [правила контроля качества](#).

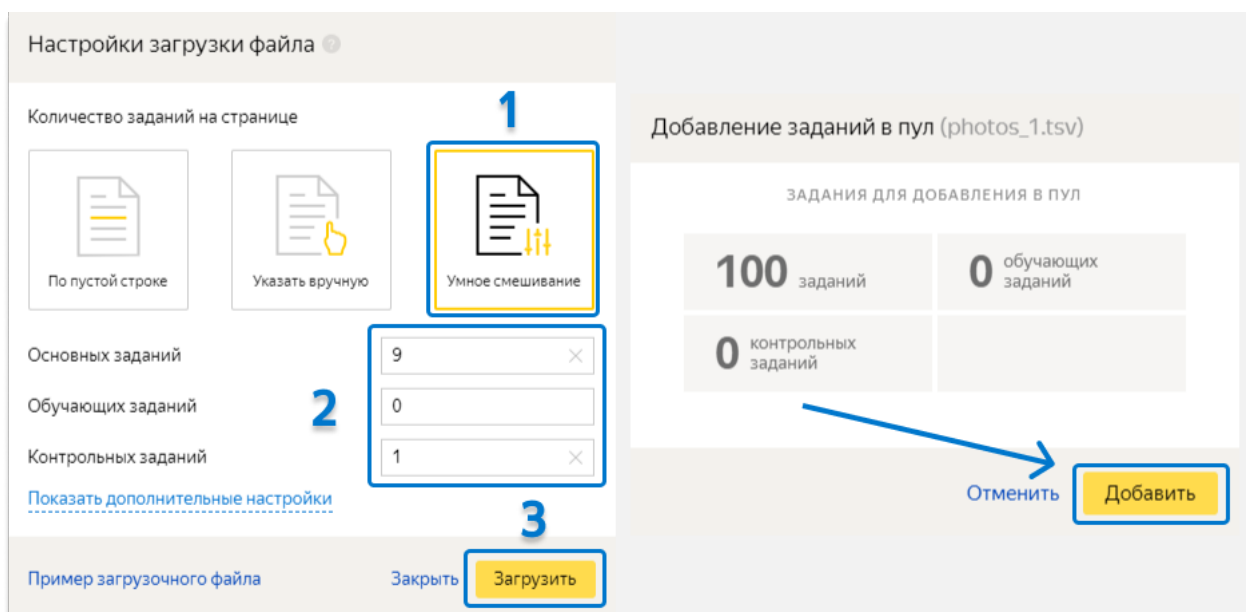
6. Сохраните пул.

Подготовка и загрузка файла с заданиями

1. Скачайте TSV-файл со ссылками на картинки. Перейдите по ссылке или QR-коду на карточке с логином и паролем, которые вы получили в начале практической части.
2. Загрузите скачанный TSV-файл в пул.
Важно! Если вы поменяли название входного поля в проекте, измените название в файле INPUT:<ваше имя поля>.



Выберите **Умное смешивание** и укажите количество заданий на странице. Например: 9 основных и 1 контрольное.



3. Создайте контрольные задания.
Справка. Контрольные задания — это задания с заранее известным ответом. Они нужны, чтобы следить за качеством ответов исполнителя. Заданный вами ответ сравнивается с ответом исполнителя, если они не совпадают, значит исполнитель ответил неверно.

- Нажмите **Разметить** → **Создать контрольные**.

Яндекс Толока | Проекты | Пользователи | Навыки | Профиль | Сообщения | 0,00 \$ / 89,59 \$ | Толока

Проекты > Есть ли на картинке дорожный знак? > мой пул

мой пул — закрыт

Статистика | Скачать результаты | Редактировать

задания пула (Пример загрузочного файла (txt, кодировка UTF-8))

Загрузить | файлы | Удалить | **Разметить**

0 страниц заданий | 0 обучающее задание | 100 заданий | 0 контрольное задание

0 %
Сделано 0

Разметить задания

Сделайте из основных заданий контрольные или обучающие.

Контрольные — задания для проверки исполнителей. Содержат правильные ответы, которые сравниваются с ответами исполнителей.

Обучающие — задания для тренировки исполнителей. Содержат правильные ответы и подсказки.

[Подробнее](#)

Основные 100 | Контрольные 0 | Обучающие 0

Создать контрольные

Создать обучающие

Скачать

ID	Перекрывтие	Ответы исполнителей	Обновлено
...e395e3a2	3	0	03.09.2019 15:23:22
...e395e3cd	3	0	03.09.2019 15:23:22

- Отметьте галочкой выходное поле **result**, которое будет учитываться при проверке совпадения ответа пользователя с эталоном, выберите вариант ответа и нажмите кнопку **Сохранить и перейти к следующему**.

Проекты > Есть ли на картинке дорож... > мой пул > Загруженные задания > Разметка заданий

Создайте ещё 1

Создать контрольное задание

1. Укажите правильные ответы
2. Отметьте, какие поля учитывать

Поле	Значение
<input checked="" type="checkbox"/> result	OK

1

3

Сохранить и перейти к следующему

Распределение правильных ответов на контрольные задания

Создайте контрольные задания, чтобы увидеть графики распределения ответов.

Есть ли на картинке дорожный знак?

1 Да 2 Нет 3 Ошибка загрузки

2

- Укажите правильные ответы у 10 заданий. В небольших пулах контрольными должны быть менее 10% заданий.

Совет. На панели справа показано распределение правильных ответов на контрольные задания. Старайтесь, чтобы разные варианты ответов встречались примерно в равной степени

Проекты > Есть ли на картинке дорож... > мой пул > Загруженные задания > Разметка заданий

Контрольных: 10

Создать контрольное задание

1. Укажите правильные ответы
2. Отметьте, какие поля учитывать

<input type="checkbox"/>	Поле	Значение
<input type="checkbox"/>	result	

Распределение правильных ответов на контрольные задания

result

40%	BAD
60%	OK

- Закончив разметку, проверьте, что в параметрах пула появилось число размеченных вами заданий.

4. Запустите пул.

Важно. Помните, что задание будут выполнять реальные пользователи Толоки. Тщательно проверьте все настройки перед тем, как нажать кнопку **Пуск**.

Яндекс Толока

Проекты Пользователи Навыки Профиль Сообщения

0,00 \$ / 89,59 \$ Толока

Проекты > Есть ли на картинке дорожный знак? > мой пул

мой пул — закрыт

Статистика Скачать результаты Редактировать

ЗАДАНИЯ ПУЛА (Пример загрузочного файла (tsv, кодировка UTF-8))

Загрузить файлы Удалить Разметить Предпросмотр

~30 страниц заданий	0 обучающее задание
90 заданий	10 контрольное задание

0 %
Сделано 0

Получение результатов

1. Дождитесь окончания разметки пула. Обновляйте страницу, чтобы следить за прогрессом.
2. Нажмите стрелочку рядом с кнопкой **Скачать результаты** и запустите [агрегацию по методу Дэвида-Скина](#).

Яндекс Толока

Проекты Пользователи Навыки Профиль Сообщения

0,00 \$ / 8,29 \$ Toloka requester

Проекты > Есть ли на картинке дорожный знак? > мой пул

мой пул — закрыт

Статистика Скачать результаты Редактировать

К списку операций
Агрегация результатов по методу Дэвида-Скина
Агрегация результатов по навыку исполнителя

ЗАДАНИЯ ПУЛА (Пример загрузочного файла (tsv, кодировка UTF-8))

Загрузить файлы Разметить Предпросмотр

30 страниц заданий 0 обучающее задание

90 заданий 10 контрольное задание

100 %

Сделано 30, принято 30

Просмотреть задания

3. Перейдите к списку операций и дождитесь завершения агрегации.
Совет. Процесс может занять от 5 до 20 минут. В это время вы можете начать работу над следующим проектом. Обновляйте страницу, чтобы следить за прогрессом.

Яндекс Толока

Проекты

Агрегация результатов успешно запущена. Перейти к списку операций.

4. Когда агрегация завершится, скачайте результаты.

Яндекс Толока

Проекты Пользователи Навыки Профиль Сообщения

0,00 \$ / 8,29 \$ Toloka requester

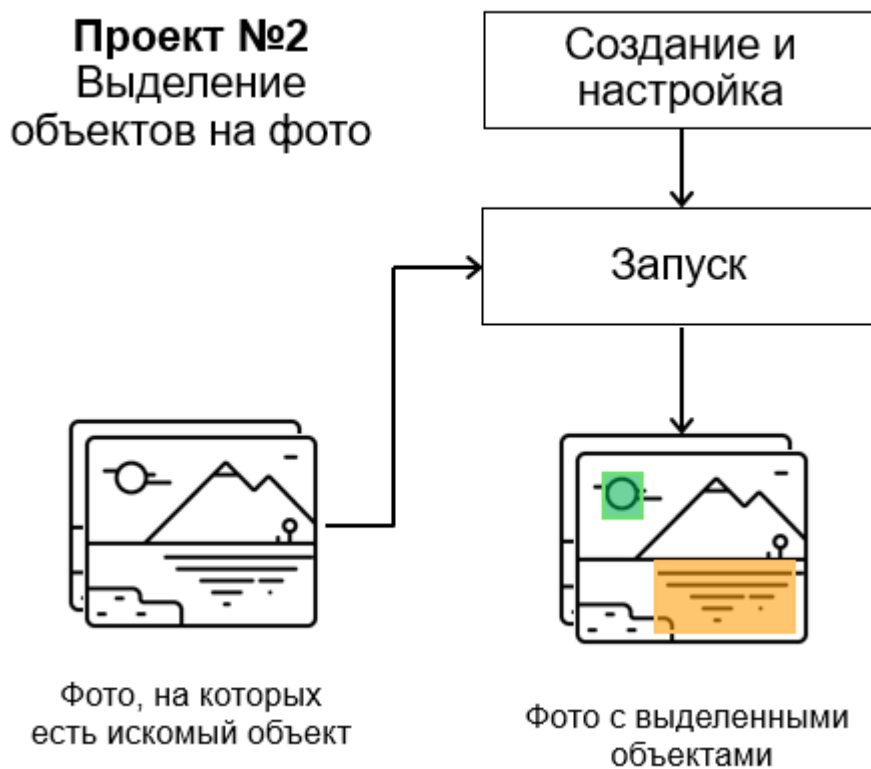
Операции

ПРОЕКТ: Есть ли на картинке дорожный зн... ПУЛ: мой пул

Id	Тип операции	Время запуска	Время завершения	Прогресс	Статус	Файлы
4d34ес...	Агрегация результатов по методу Дэвида-Скина	05.08.2019 21:00:36	05.08.2019 21:03:11	100%	Завершена	Скачать

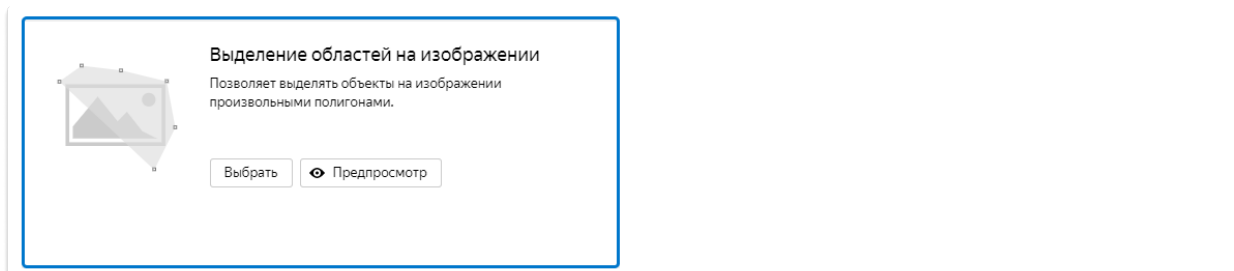
5. Используйте полученный TSV-файл [в проекте № 2](#).

Проект № 2. Выделение объектов на фото



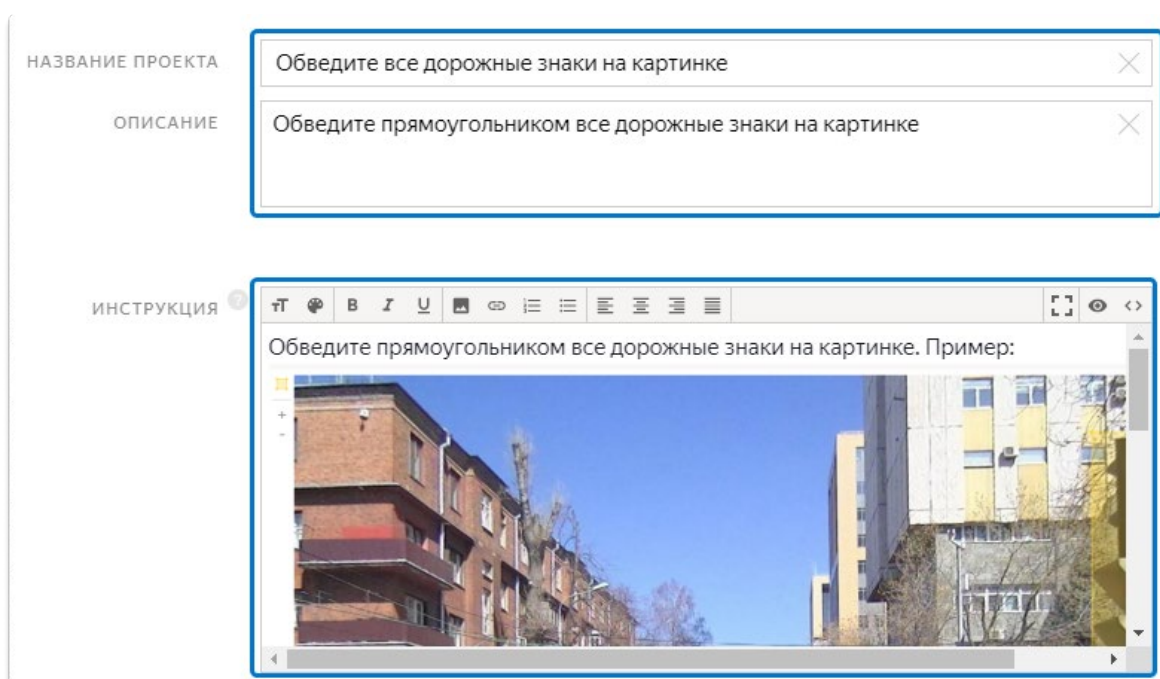
Создание проекта

1. Нажмите кнопку **+** **Создать проект** и выберите шаблон **Выделение областей на изображении**.

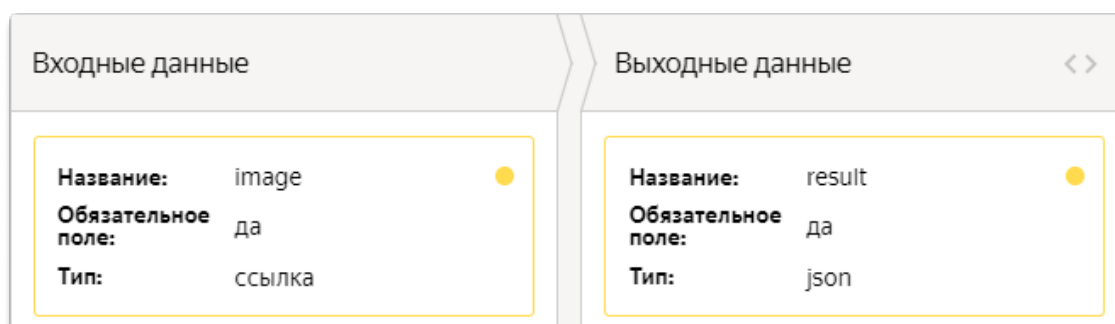


2. Дайте проекту понятное название и краткое описание.

Важно: Их увидят реальные исполнители Толоки в списке доступных заданий.

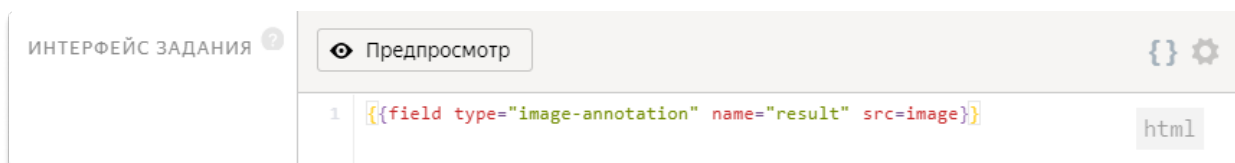


3. Напишите простую и короткую инструкцию.
4. Определите поля для **входных и выходных данных** в блоке **Спецификации**:
 - Во входное поле **image** мы передадим ссылки на картинки, чтобы показать их исполнителям.
TSV-файл со ссылками мы впоследствии [загрузим в пул](#).
 - В выходное поле **result** запишем координаты выделенного объекта в формате JSON.



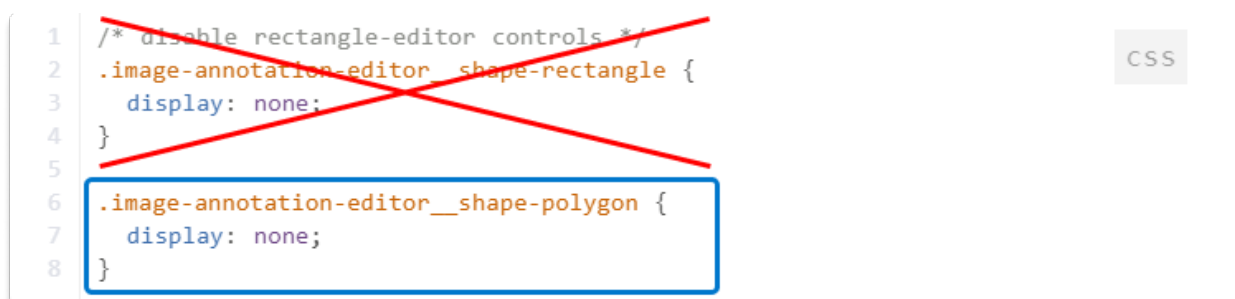
5. Создайте интерфейс задания в блоке HTML.

В шаблоне автоматически добавлен редактор для выделения области на картинке. Это означает, что исполнитель увидит картинку и сможет выделить область инструментами редактора. Поэтому в интерфейсе можно ничего не менять.



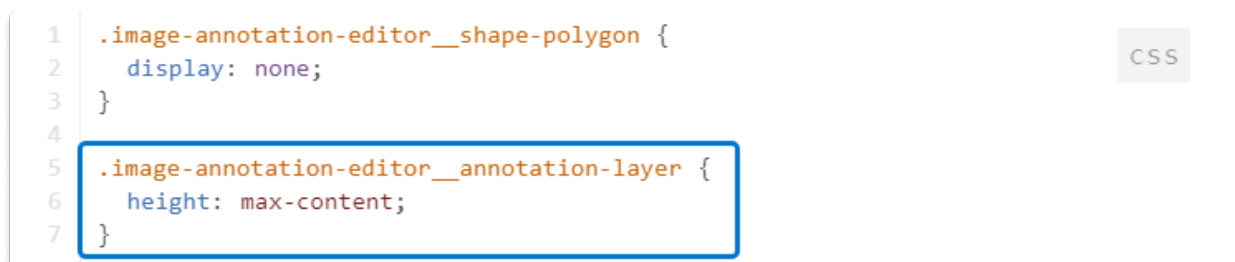
6. По умолчанию редактор даёт возможность выделить область на картинке произвольным многоугольником. Чтобы изменить этот инструмент на прямоугольник, удалите имеющийся код из поля CSS и добавьте следующий код:

```
.image-annotation-editor__shape-polygon { display: none;}
```



7. Добавьте код для настройки высоты интерфейса разметки задания по размеру изображения:

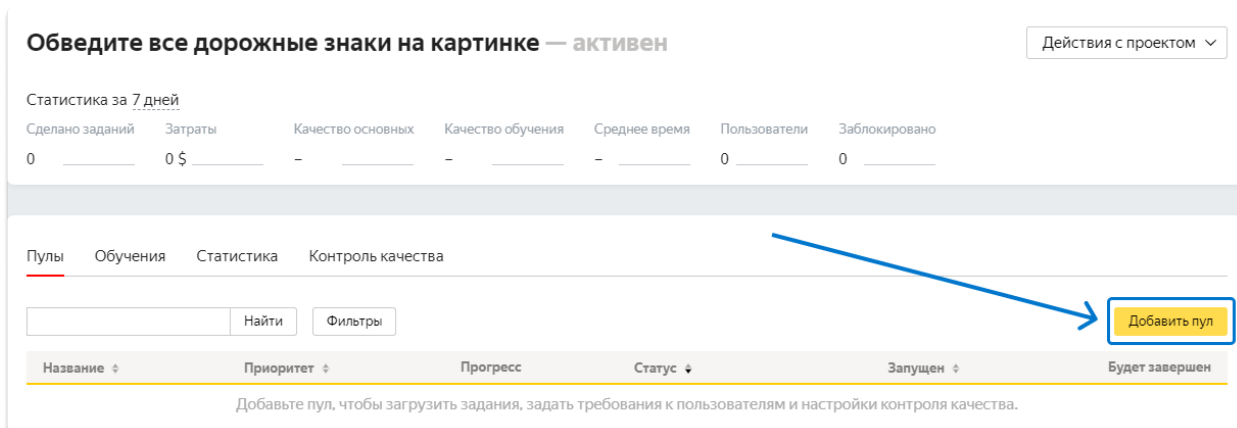
```
.image-annotation-editor__annotation-layer { height: max-content;}
```



8. Нажмите кнопку **Предпросмотр**, чтобы увидеть получившееся задание. Убедитесь, что его можно выполнить.
9. Чтобы сохранить изменения в проекте, нажмите **Сохранить**.

Создание пула

1. Нажмите **Добавить пул**.



Обведите все дорожные знаки на картинке — активен Действия с проектом ▾

Статистика за 7 дней

Сделано заданий	Затраты	Качество основных	Качество обучения	Среднее время	Пользователи	Заблокировано
0	0 \$	-	-	-	0	0

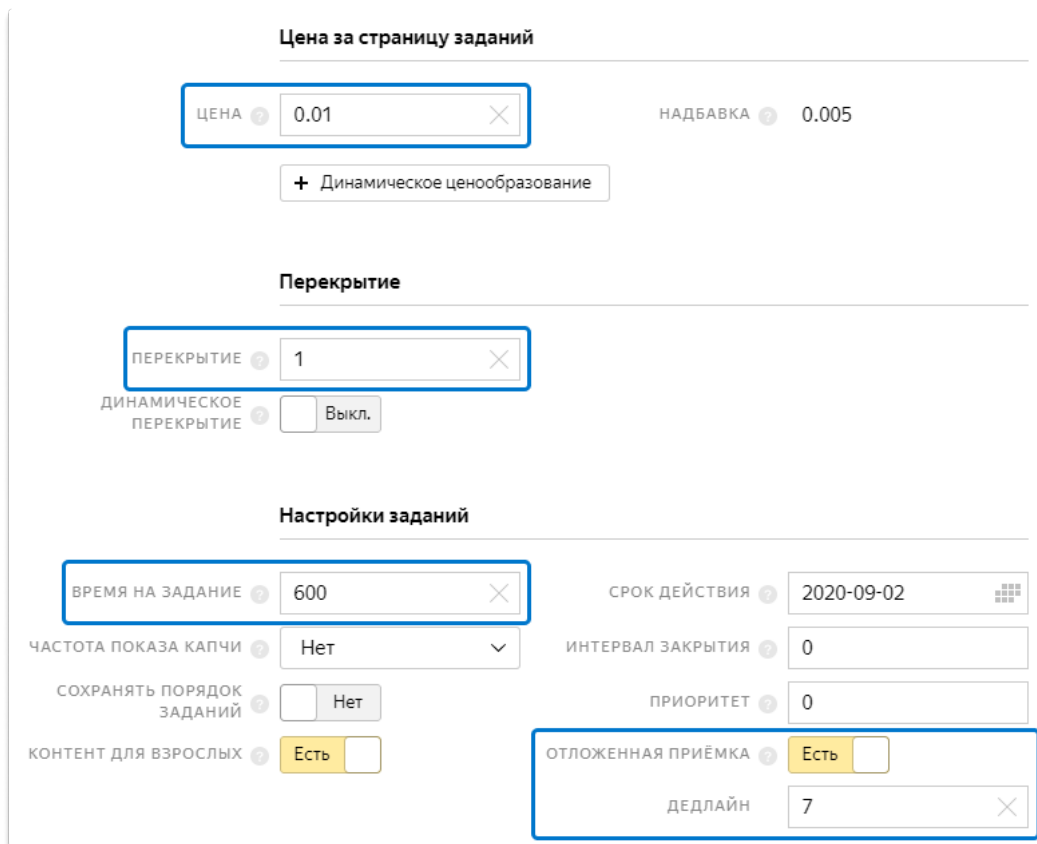
Пулы | Обучения | Статистика | Контроль качества

Название ▾	Приоритет ▾	Прогресс	Статус ▾	Запущен ▾	Бюджет завершен
------------	-------------	----------	----------	-----------	-----------------

Добавьте пул, чтобы загрузить задания, задать требования к пользователям и настройки контроля качества.

2. Укажите [параметры пула](#):

- Цена за страницу заданий (например, \$0.01)
- Перекрытие. Это количество пользователей, которые выполняют одну и ту же страницу с заданиями. Для заданий с проверкой ответов обычно устанавливают перекрытие 1.
- Время на выполнение страницы заданий. Например, 600 секунд
- Включите опцию **Отложенная приёмка** и укажите количество дней на проверку. Например, 7.



Цена за страницу заданий

ЦЕНА НАДБАВКА

Перекрытие

ПЕРЕКРЫТИЕ

ДИНАМИЧЕСКОЕ ПЕРЕКРЫТИЕ Выкл.

Настройки заданий

ВРЕМЯ НА ЗАДАНИЕ СРОК ДЕЙСТВИЯ

ЧАСТОТА ПОКАЗА КАПЧИ ИНТЕРВАЛ ЗАКРЫТИЯ

СОХРАНЯТЬ ПОРЯДОК ЗАДАНИЙ Нет ПРИОРИТЕТ

КОНТЕНТ ДЛЯ ВЗРОСЛЫХ Есть ОТЛОЖЕННАЯ ПРИЁМКА Дедлайн

3. Добавьте **Фильтры** для отбора исполнителей, которым будет доступно задание.
- Чтобы ваше задание было доступно только исполнителям, владеющим русским языком, выберите готовый набор **Русскоязычные исполнители**.

ВЫЧИСЛИМЫЕ

Регион по номеру телеф... = Россия

или = Украина

или = Казахстан

или = Беларусь

и

ПРОФИЛЬ

Языки = Русский

- Создайте навык **«выделение областей»**, который будет присваиваться пользователям после выполнения заданий в пуле. С помощью этого навыка мы запомним этих пользователей и не допустим их к проверке заданий в следующем проекте. Нажмите кнопку **Создать навык**.

Требования к пользователям ? [Скопировать настройки из...](#)

Добавить фильтр

Создать навык

- Введите название навыка и, если хотите, добавьте описание. Оно доступно только вам. Оставьте навык приватным, как это сделано по умолчанию, и нажмите **Добавить**.

Добавление навыка

НАЗВАНИЕ

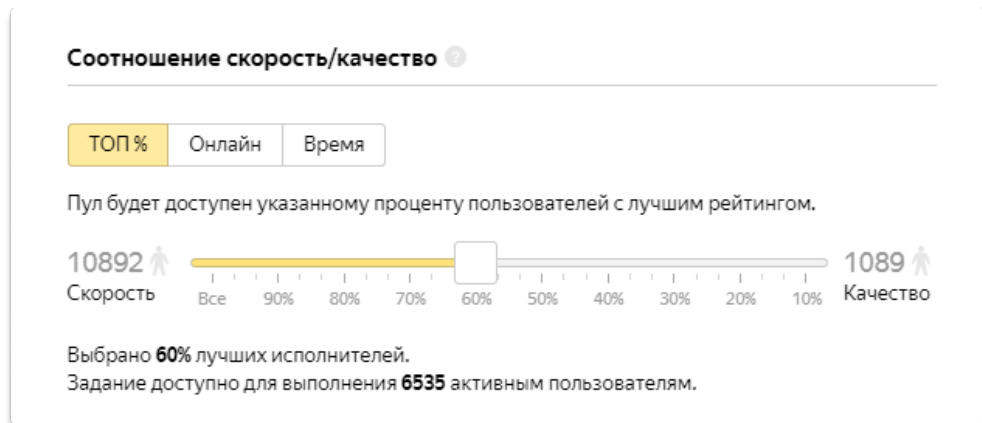
выделение областей

ОПИСАНИЕ

Публичный? Нет ?

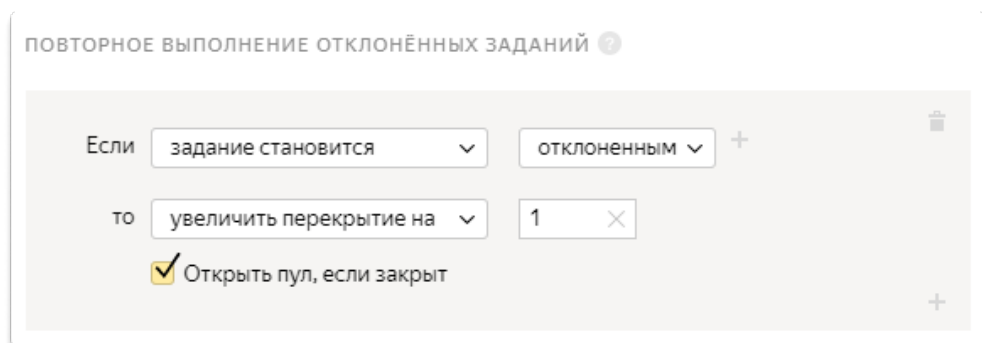
Отменить **Добавить**

- По желанию, укажите процент лучших по рейтингу исполнителей в блоке [Соотношение скорость/качество](#).

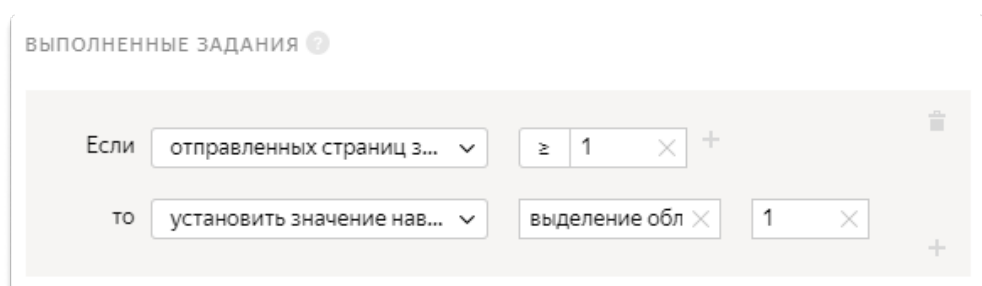


4. [Настройте контроль качества](#).

- Верните отклонённые задания в пул, чтобы их выполнили другие исполнители. Добавьте правило [Повторное выполнение отклонённых заданий](#).



- [Выполненные задания](#). Добавьте правило, чтобы отметить исполнителя, выполнившего хотя бы одно задание в пуле.



Совет. Если в выпадающем списке не появляется название созданного вами навыка, сохраните пул, а затем снова откройте режим редактирования.

- По желанию добавьте другие [правила контроля качества](#).
Совет. Контрольные задания и мнение большинства в такого типа проектах не используются, т.к. для этого разметка пользователей должна в точности совпасть с эталонной, что практически невозможно.

5. Сохраните пул.

Подготовка и загрузка файла с заданиями

1. Откройте [файл с агрегированными результатами](#) из проекта № 1.
2. Отберите только те картинки, на которых есть пригодные для разметки изображения (ответы ОК). Используйте текстовый редактор или редактор электронных таблиц.
3. Скопируйте столбец с отобранными ссылками на новую страницу или в новый документ и дайте название столбцу INPUT:image.
Важно! Если вы поменяли название входного поля в проекте, измените название в файле INPUT:<ваше имя поля>.
4. Сохраните файл в формате TSV.

Совет. Если у вас компьютер на базе UNIX OS (Linus, MacOS), для отбора изображений вы можете воспользоваться командами `awk`:

```
awk 'BEGIN {OFS = FS = "\t";} $2=/OK/ {print $1}' <aggregated_res>.tsv >  
<filtered_res>.tsv
```

5. [Загрузите файл](#) в пул, выбрав **Указать вручную**. Поставьте 1 задание на странице.

Настройки загрузки файла

Количество заданий на странице

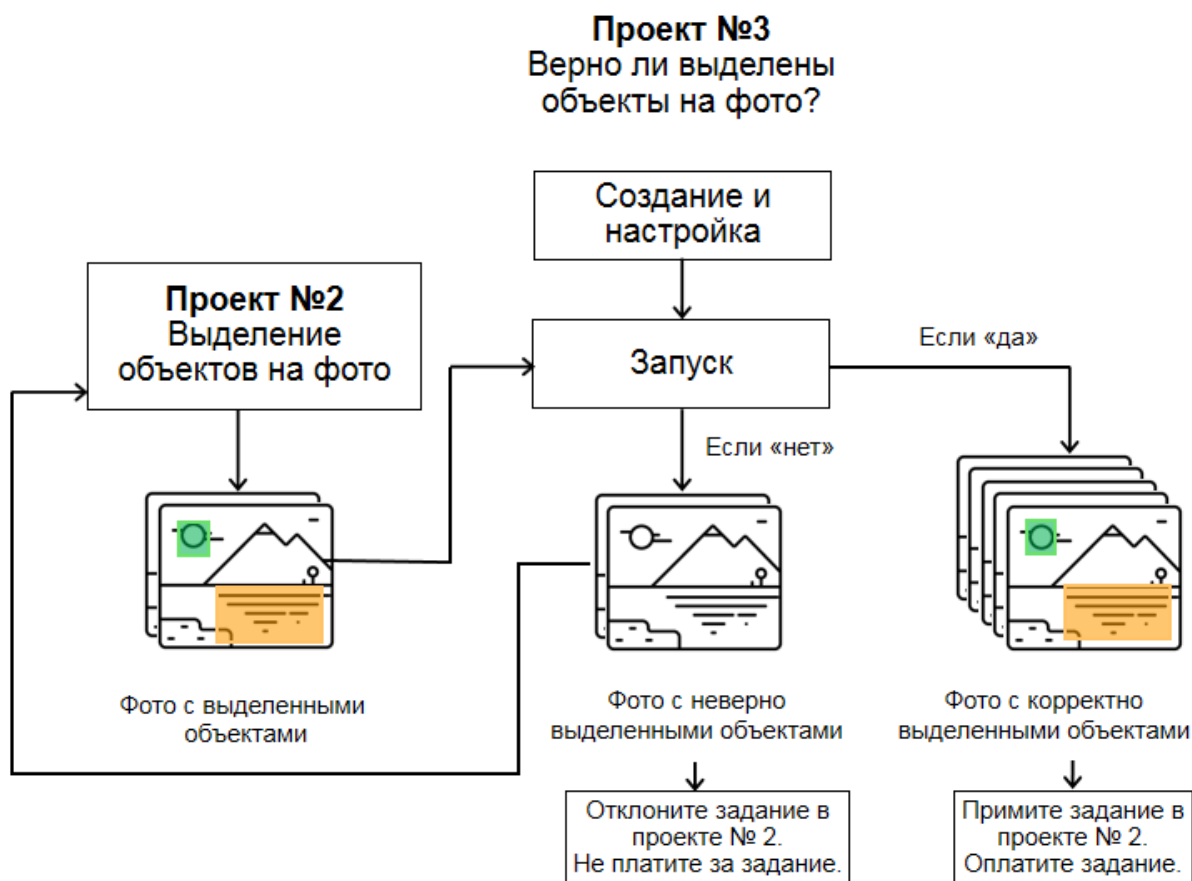
По пустой строке Указать вручную Умное смешивание

Заданий на странице: 1

Пример загрузочного файла Заккрыть Загрузить

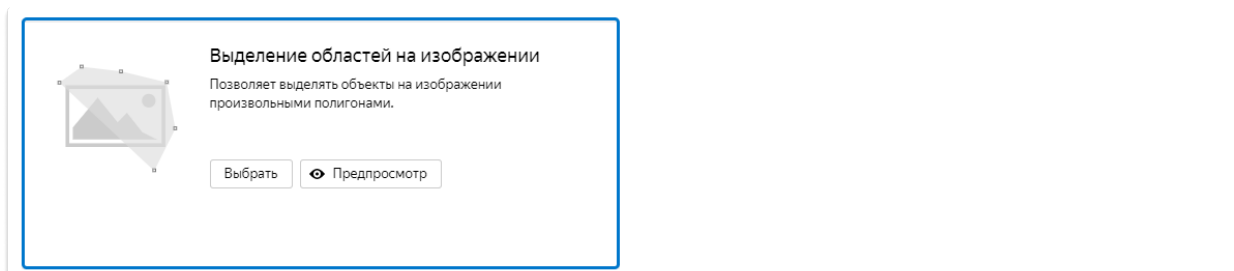
6. Запустите пул.

Проект № 3. Верно ли выделены объекты на фото?



Создание проекта

1. Нажмите кнопку **+ Создать проект** и выберите шаблон **Выделение областей на изображении**. Этот шаблон содержит редактор для выделения областей на изображении, с помощью которого мы покажем исполнителям разметку, полученную в предыдущем проекте.



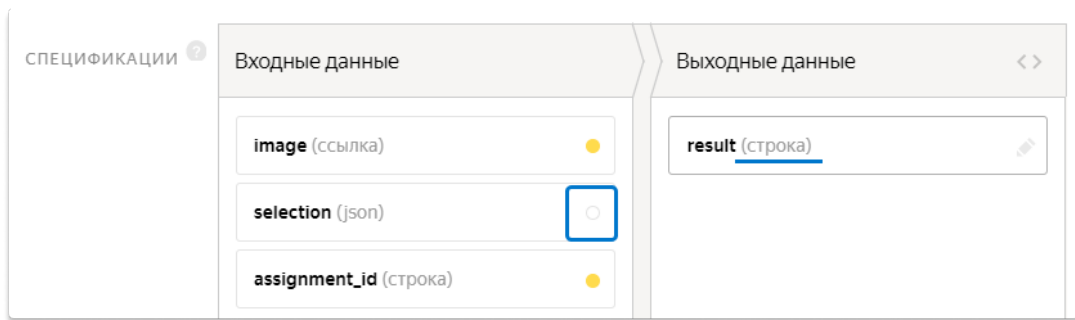
2. Дайте проекту понятное название и краткое описание.
Важно: Их увидят реальные исполнители Толоки в списке доступных заданий.

НАЗВАНИЕ ПРОЕКТА: Верно ли выделены дорожные знаки

ОПИСАНИЕ: Посмотрите на картинку и определите, правильно ли выделены дорожные знаки

ИНСТРУКЦИЯ: Посмотрите на картинку и ответьте на вопрос: **Все ли дорожные знаки выделены верно?**
Например, на этой картинке все дорожные знаки обведены правильно:

3. Напишите простую и понятную инструкцию.
4. Определите поля для входных и выходных данных в блоке **Спецификации**:
 - Во входное поле **image** мы передадим ссылки на картинки, чтобы показать их исполнителям.
TSV-файл со ссылками мы впоследствии [загрузим в пул](#).
 - В поле **assignment_id** с типом «строка» передадим номер выполненного задания из предыдущего проекта.
 - В поле **selection** с типом «json» передадим координаты выделенных в предыдущем задании объектов. Сделайте это поле необязательным.
 - В выходное поле **result** с типом «строка» запишем ответ исполнителя: верно или неверно.



5. Создайте интерфейс задания.

- Добавьте в редактор для выделения областей параметр **"annotations=selection"** чтобы показать выделенный объект исполнителю
- Измените имя редактора для выделения областей. Например, **"object"**.
- Добавьте две радиокнопки:

```

{{field type="radio" name="result" value="OK" label="Correct" hotkey="1"}}
{{field type="radio" name="result" value="BAD" label="Incorrect" hotkey="2"}}

```

```

{{field type="image-annotation" name="object" annotations=selection src=image}}
{{field type="radio" name="result" value="OK" label="Верно" hotkey="1"}}
{{field type="radio" name="result" value="BAD" label="Неверно" hotkey="2"}}

```

- Удалите из интерфейса инструмент для выделения области: в этом задании не нужно исправлять или добавлять разметку. Добавьте в блок CSS код:

```
.image-annotation-editor__shape-polygon {display: none;}
```

- Добавьте код для настройки высоты интерфейса разметки задания по размеру изображения:

```
.image-annotation-editor__annotation-layer {height: max-content;}
```

```

1  /* disable rectangle-editor controls */
2  .image-annotation-editor__shape-rectangle {
3    display: none;
4  }
5
6  /* disable polygon-editor controls */
7  .image-annotation-editor__shape-polygon {
8    display: none;
9  }
10
11 .image-annotation-editor__annotation-layer {
12   height: max-content;
13 }

```

6. Нажмите кнопку **Предпросмотр**, чтобы увидеть получившееся задание.
7. Отмечайте радиокнопки в задании, чтобы убедиться, что его можно выполнить.
8. Сохраните изменения, нажав кнопку **Сохранить**.

Создание пула

1. Нажмите **Добавить пул**.
2. Дайте пулу любое удобное название и описание. Они доступны только вам.
3. Укажите [параметры пула](#):
 - Цена за страницу заданий (например, \$0.01).
 - Перекрытие. Например, 3.
 - Время на выполнение страницы заданий. Например, 600 секунд.
 - Сохранять порядок заданий. Активируйте переключатель чтобы быстрее получить нужное число ответов на задание для определения [мнения большинства](#).

Цена за страницу заданий

ЦЕНА 0.01 НАДБАВКА 0.005

+ Динамическое ценообразование

Перекрытие

ПЕРЕКРЫТИЕ 3

ДИНАМИЧЕСКОЕ ПЕРЕКРЫТИЕ Выкл.

Настройки заданий

ВРЕМЯ НА ЗАДАНИЕ 600 СРОК ДЕЙСТВИЯ 2020-09-04

ЧАСТОТА ПОКАЗА КАПЧИ Нет ИНТЕРВАЛ ЗАКРЫТИЯ 0

СОХРАНЯТЬ ПОРЯДОК ЗАДАНИЙ Да ПРИОРИТЕТ 0

КОНТЕНТ ДЛЯ ВЗРОСЛЫХ Есть ОТЛОЖЕННАЯ ПРИЁМКА Нет

ДЕДЛАЙН

4. Добавьте [Фильтры](#) для отбора исполнителей, которым будет доступно задание.
 - Чтобы ваше задание было доступно только исполнителям, владеющим русским языком, выберите готовый набор **Русскоязычные исполнители**.
 - Ограничьте доступ к проверочному заданию исполнители, которые участвовали в выделении объектов в предыдущем проекте. Установите фильтр **«выделение областей = отсутствует»** (пустое поле).

НАВЫКИ

выделение областей = Отсутствует

- По [желанию](#), укажите процент лучших по рейтингу исполнителей в блоке **Соотношение скорость/качество**.

5. [Настройте контроль качества.](#)

- [Мнение большинства.](#) Добавьте правило контроля мнением большинства. Укажите сколько ответов считать большинством. Для перекрытия 3 укажите 2 ответа.

МНЕНИЕ БОЛЬШИНСТВА

Считать большинством

Сколько последних значений учитывать

Если

и

то

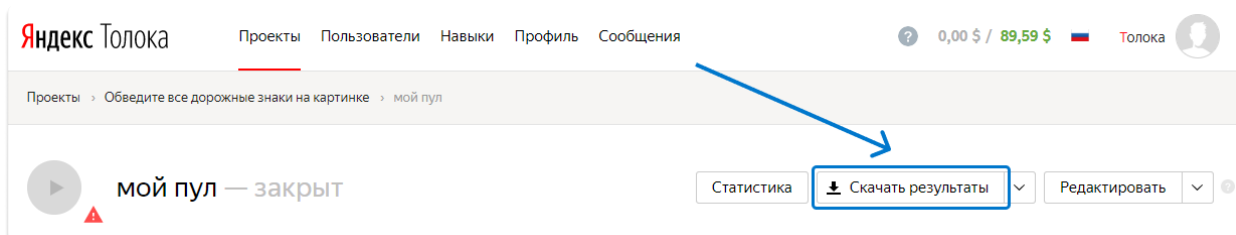
Важно! Правило начинает действовать, когда количество ответов на задание равно перекрытию. Чтобы быстрее получить нужное число ответов, включите опцию **Сохранять порядок заданий** в параметрах пула.

По желанию добавьте другие [правила контроля качества.](#)

6. •Сохраните пул.

Подготовка и загрузка файла с заданиями

1. Дождитесь окончания разметки пула в проекте № 2.
2. Откройте страницу пула в проекте № 2 и нажмите **Скачать результаты**.



- Замените галочку **Принятые** на **Не проверенные**.
- Снимите галочки **ссылка**, **id пользователя**, **статус**, **время старта** и **Разделять ответы пустой строкой**. Вы получите список выполненных заданий.

The screenshot shows a dialog box titled 'Скачать результаты'. It has two main sections: 'Статус' and 'Поля'. In the 'Статус' section, the 'Не проверенные' checkbox is checked, while 'Активные', 'Принятые', 'Отклоненные', 'Пропущенные', and 'Просроченные' are unchecked. In the 'Поля' section, the 'id ответа' checkbox is checked, while 'ссылка', 'id задания', 'id пользователя', 'статус', 'время старта', 'время отправки', 'время принятия', 'время отклонения', 'время пропуска', 'время истечения', and 'цена' are unchecked. At the bottom, there are three checkboxes: 'Скачать данные за период' (unchecked), 'Разделять ответы пустой строкой' (unchecked), and 'Не учитывать ответы заблокированных пользователей' (unchecked). At the very bottom, there are two buttons: 'Закрыть' and 'Скачать результаты'.

3. Измените название столбца **"OUTPUT:result"** на **"INPUT:selection"**, а **"ASSIGNMENT:assignment_id"** на **INPUT:assignment_id**. Столбец **"INPUT:image"** оставьте без изменений. Сохраните файл в формате TSV.

Важно: Вы можете использовать для подготовки файла обычный блокнот. При переносе данных в Microsoft Excel, воспользуйтесь **Мастером импорта текстов**. В шаге 2 укажите (нет) в поле **Ограничитель строк**. Таким образом вы сохраните все кавычки в JSON-объектах.

	A	B	C
1	INPUT:image	INPUT:selection	INPUT:assignment_id
2	http://storage.mds.yandex.net	"[{"data":{"p1":{"x":0.09141,"y":0.0000031c4c--5d713938310d7100e3a0fc84	
3	http://storage.mds.yandex.net	"[{"data":{"p1":{"x":0.04141,"y":0.0000031c4c--5d71393d7e170d00d8cd80b6	
4	http://storage.mds.yandex.net	"[{"data":{"p1":{"x":0.06094,"y":0.0000031c4c--5d713940310d7100e3a0fc91	

- Откройте [страницу пула в проекте № 3](#).
- [Загрузите файл в пул](#), выбрав **Указать вручную**. Задайте количество заданий на странице. Например: 10.
- Откройте Предпросмотр пула и проверьте, что в задании отображаются картинки с выделенными объектами. Если выделенных объектов нет, проверьте кавычки в INPUT:selection и убедитесь, что верно задали параметры компонента [в шаге 5 создания проекта](#).
- Если всё в порядке, запустите пул.

Получение результатов

- Дождитесь завершения разметки пула.
- Нажмите стрелочку рядом с кнопкой **Скачать результаты** и запустите [агрегацию по методу Дэвида-Скина](#).

Яндекс Толока | Проекты | Пользователи | Навыки | Профиль | Сообщения | 0,00 \$ / 8,29 \$ | Toloka requester

Проекты > Верно ли выделены дорожные знаки? > pool

pool — архивный

Статистика | Скачать результаты | Редактировать

К списку операций
Агрегация результатов по методу Дэвида-Скина
 Агрегация результатов по навыку исполнителя

ЗАДАНИЯ ПУЛА (Пример загрузочного файла (tsv, кодировка UTF-8))

Загрузить | файлы | Предпросмотр

7 страниц заданий | 0 обучающее задание
 65 заданий | 0 контрольное задание

100%
 Сделано 21, принято 21
 Просмотреть задания

- Перейдите к списку операций и дождитесь завершения агрегации. Это может занять от 5 до 20 минут..

Яндекс Толока | Проекты

Агрегация результатов успешно запущена. [Перейти к списку операций.](#)

- Скачайте результаты.

Яндекс Толока | Проекты | Пользователи | Навыки | Профиль | Сообщения | 0,00 \$ / 8,29 \$ | Toloka requester

Операции

ПРОЕКТ: Верно ли выделены дорожные зна... | ПУЛ: pool

Id	Тип операции	Время запуска	Время завершения	Прогресс	Статус	Файлы
08091a...	Агрегация результатов по методу Дэвида-Скина	05.08.2019 22:51:41	05.08.2019 22:53:33	100%	Завершена	Скачать

Загрузка результатов проверки

Мы запустили второе задание с отложенной приёмкой. Это значит, что в течение срока, указанного в поле **Дедлайн** в настройках пула, задания нужно проверить.

1. Откройте [файл с агрегированными результатами](#) из проекта № 3 в редакторе электронных таблиц.

Совет. Если у вас не установлен редактор, воспользуйтесь онлайн редактором электронных таблиц на Яндекс.Диске, подключённом к вашему аккаунту заказчика. Перейдите на <https://disk.yandex.ru/client/disk> и нажмите **Создать** → **Таблицу**. Скопируйте в неё агрегированные данные из TSV-файла.

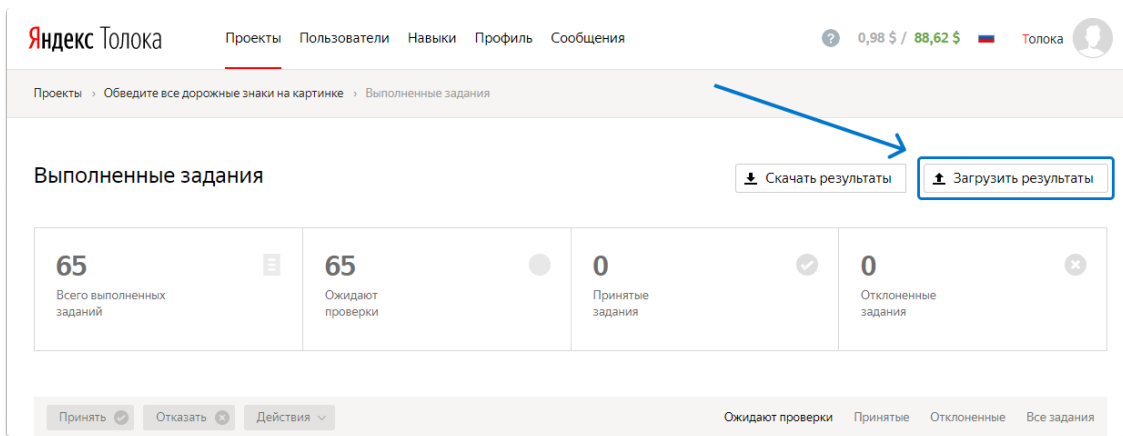
2. Добавьте столбцы:
 - **"ACCEPT:verdict"** — результат проверки задания.
 - **"ACCEPT:comment"** — комментарий для исполнителя с причиной отклонения задания (например, какой пункт инструкции нарушен).
3. Измените название столбца **"INPUT:assignment_id"** на **"ASSIGNMENT:assignment_id"**.
4. Заполните столбцы **"ACCEPT:verdict"** и **"ACCEPT:comment"**:
 - Если агрегированный результат по заданию ОК, поставьте «+» — задание будет принято.
 - Если результат оценки разметки BAD, поставьте «-» — задание будет отклонено. Напишите причину отклонения в поле **ACCEPT:comment**. Например: «Объект не выделен или выделен неверно».
5. Остальные столбцы теперь можно удалить. Сохраните отредактированный TSV-файл

	A	B	C
1	ASSIGNMENT:assignment_id	ACCEPT:verdict	ACCEPT:comment
2	000052efc4--5d0dab2fd3f5e2ecac89337b	+	
3	000052efc4--5d0daf0fb18cbb0152e8ec4f	+	
4	000052efc4--5d0db83e6a6ffd0155488e5e	-	Объект выделен неверно

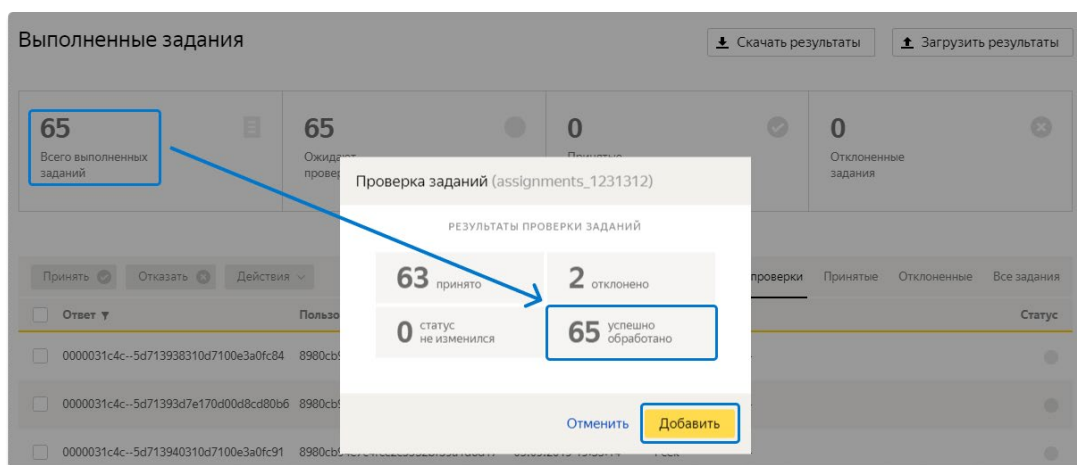
Совет. Если у вас компьютер на базе UNIX OS (Linus, MacOS), для отбора изображений вы можете воспользоваться командами `awk`:

```
awk 'BEGIN {FS=OFS="\t";} NR>1 {if($4~"OK"){ print $1, "+", "";}else{ print $1, "-", "Объект не выделен или выделен неверно.";}}' <post_accept_res>.tsv > <review_res>.tsv
```

5. Откройте [страницу пула в проекте № 2](#).
6. Нажмите кнопку [Проверить задания](#) на странице пула над прогресс-баром.
7. Нажмите **Загрузить результаты**.



8. Выберите получившийся файл и загрузите его в Толоку.
9. Проверьте, что все задания изменили статус на принятые и отклонённые.

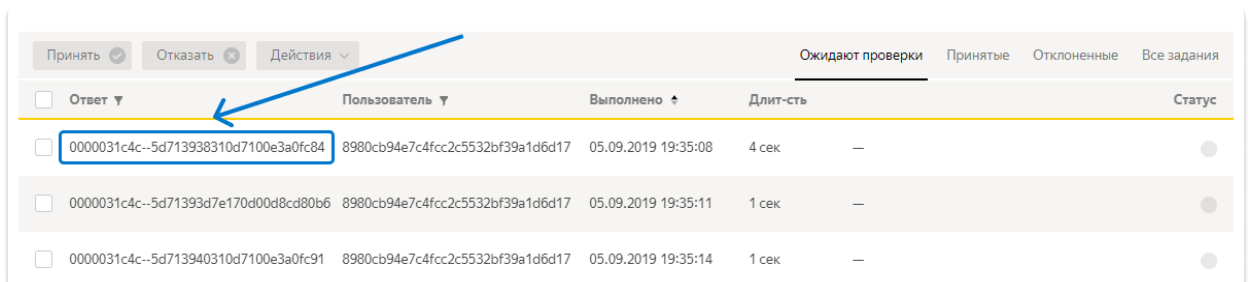


10. Если вы отклонили задания и настроили правило **Повторное выполнение отклонённых заданий**, то пул откроется заново, и эти задания будут вновь выданы исполнителям. После того, как пул закроется, выгрузите новые результаты и отправьте на проверку. Загрузите результаты. Повторяйте эти шаги до тех пор, пока все изображения из второго проекта не будут размечены верно.

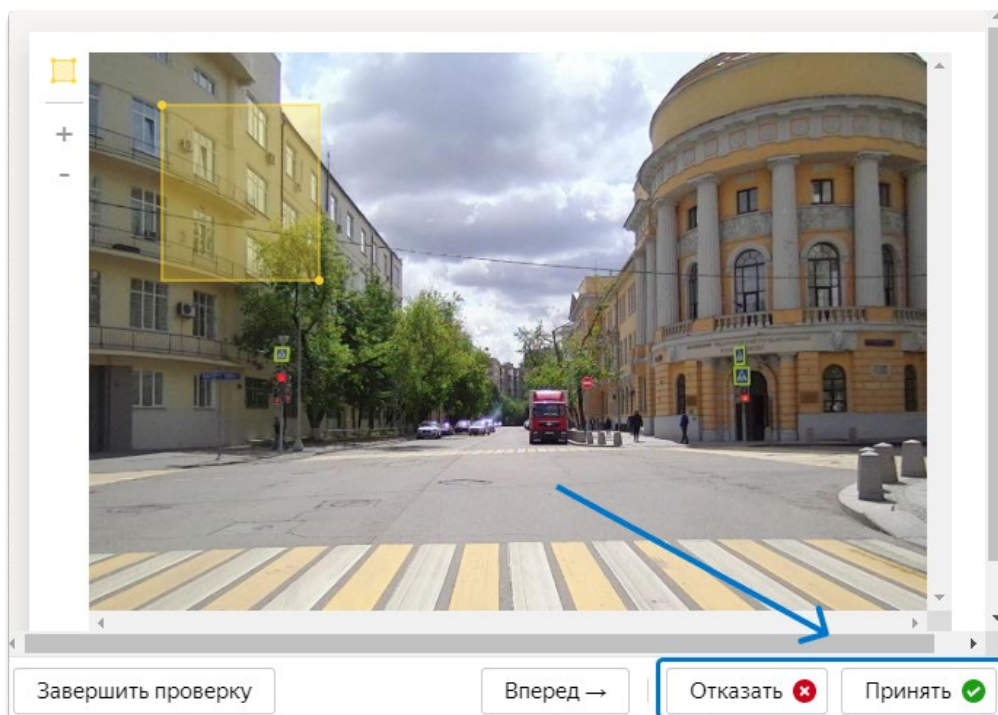
Проверка ответов онлайн и анализ результатов

Вы можете [проверить задания самостоятельно](#) в интерфейсе Толоки и получить итоговые результаты.

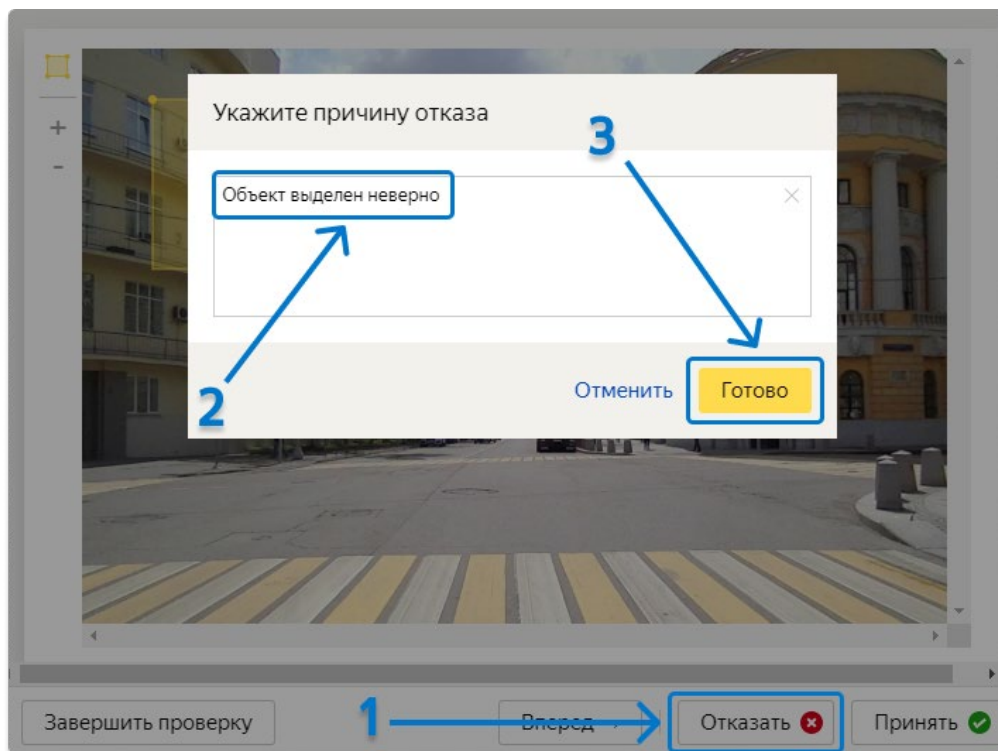
1. Откройте [страницу пула в проекте № 2](#).
2. Нажмите **Проверить задания** над индикатором выполнения.



- Откройте задание, кликнув по его ID. Проверьте правильность выделения объектов и нажмите **Принять** или **Отклонить**.



- Для отклоненных заданий введите комментарий (укажите причину).



- Когда все задания будут проверены, вернитесь на страницу пула и нажмите **Скачать результаты**. Выгрузите только принятые задания.
- Готово! Вы получили набор фотографий с выделенными объектами.

Поздравляем!