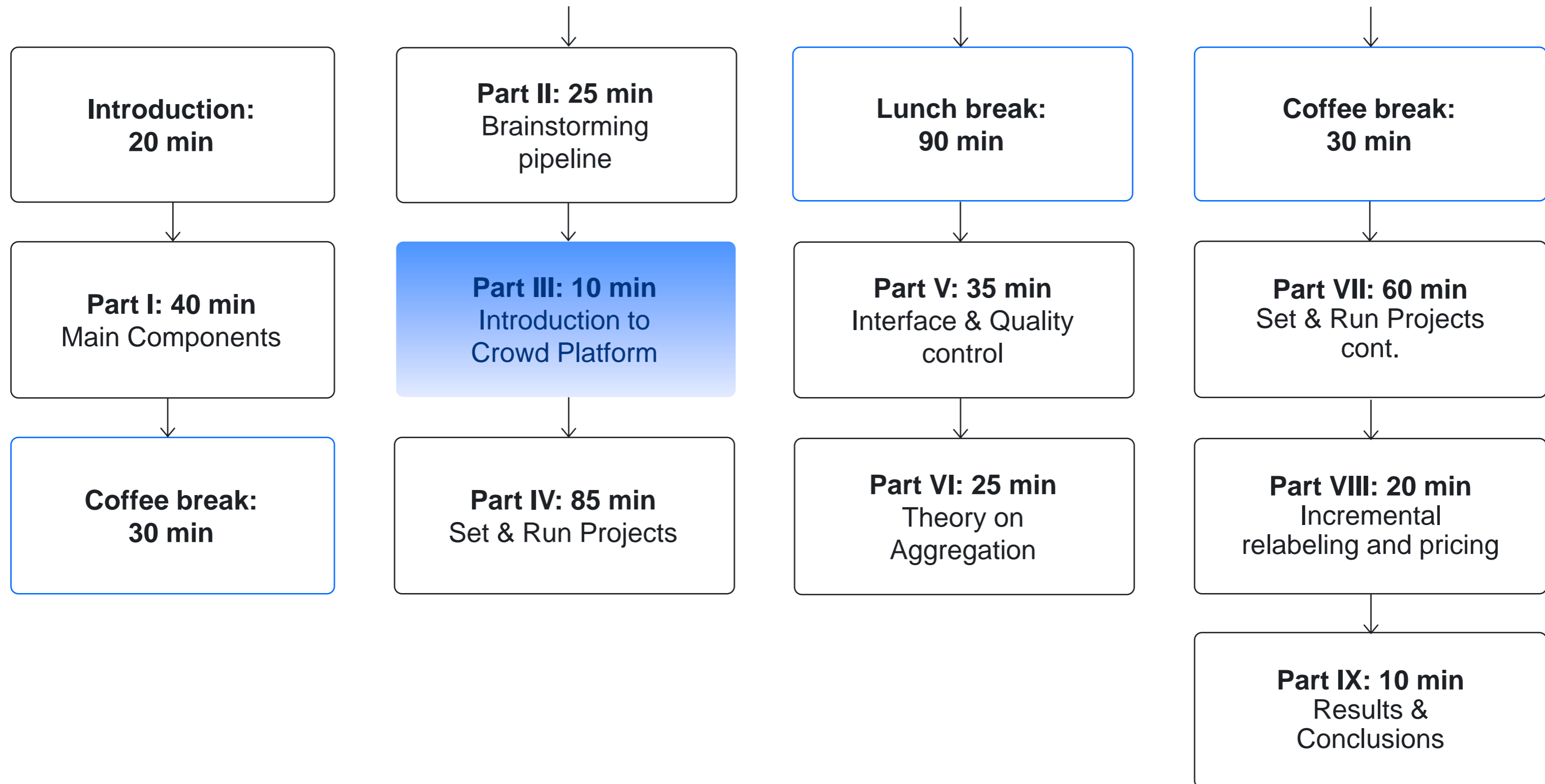Part III

Introduction to Toloka for requesters

Evfrosiniya Zerminova,
Head of Data Analysis and Research Group, Toloka

Tutorial outline

Introduction:
20 min

Part I: 40 min
Main Components

Coffee break:
30 min

Part II: 25 min
Brainstorming pipeline

Part III: 10 min
Introduction to Crowd Platform

Part IV: 85 min
Set & Run Projects

Lunch break:
90 min

Part V: 35 min
Interface & Quality control

Part VI: 25 min
Theory on Aggregation

Coffee break:
30 min

Part VII: 60 min
Set & Run Projects cont.

Part VIII: 20 min
Incremental relabeling and pricing

Part IX: 10 min
Results & Conclusions

Key types of instances in Toloka

Project

► Defines the structure of tasks

► Defines how to perform them

Configure in a project

► Input and output data types

► Task interface

► Task instruction

Pool

► Is a batch of tasks

► Defines access of performers

Configure in a pool

► Performer filters
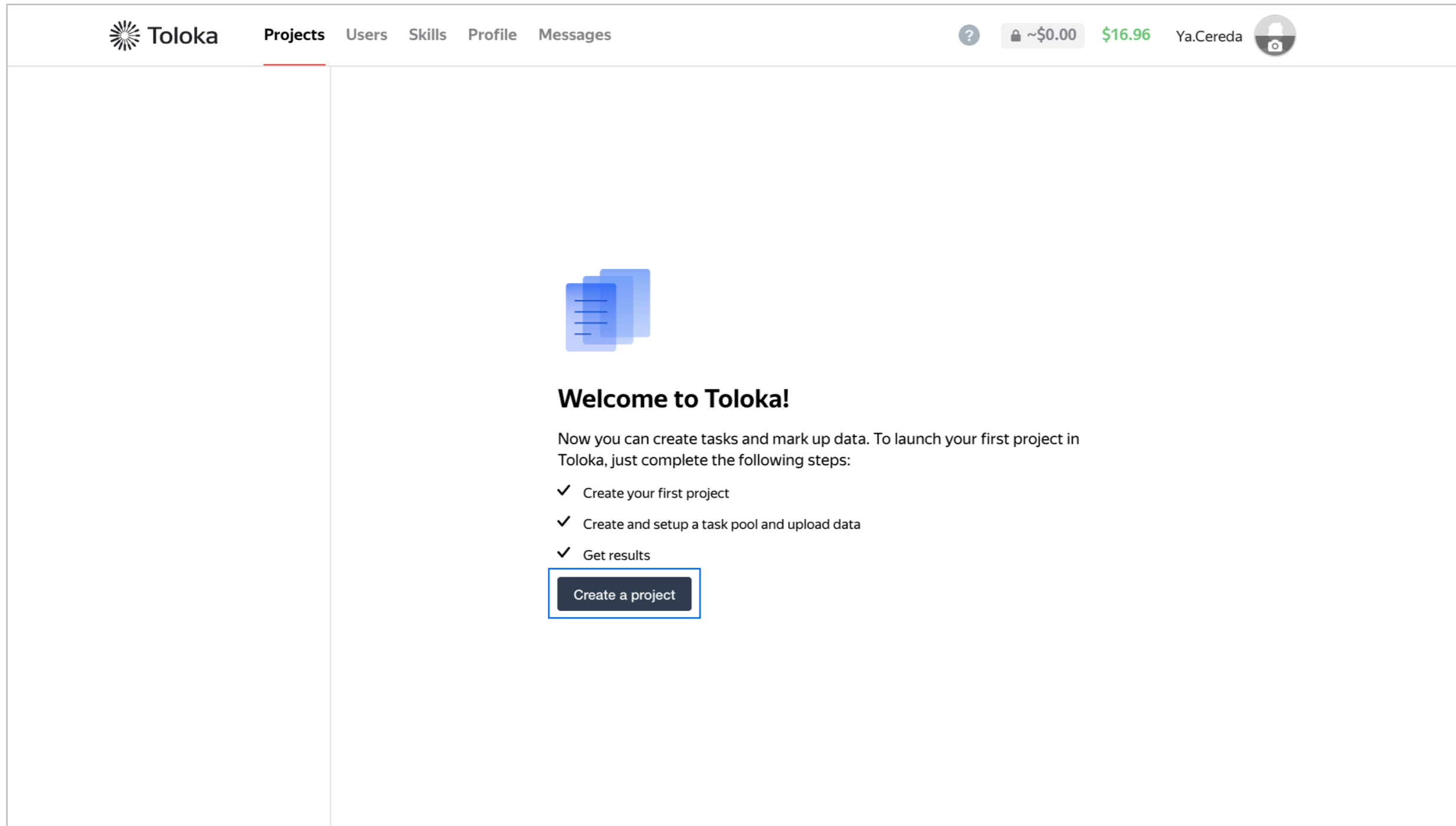
► Quality control mechanisms

► Overlap settings

Task

► A particular input data

► Results for it from performers

Project: creation & configuration

Project: creation

Project: configure name and instruction

Project: configure in/out data types

Project: configure interface

Project: use the task preview to see how it works

Project: saving

Edit project   Back to the old interface   Cancel   Finish

✓ General information

✓ Task interface

✓ Instructions for performers

✓ Translations

Pool:
creation & configuration

Pool: creation

Are there shoes in the picture? — active

Project actions ⌄

Statistics for 7 days

Submitted tasks	Spent	Quality: control tasks	Quality: training tasks	Average submit time	Users	Banned users
0	0 $	–	–	–	0	0

Pools   Training   Statistics   Quality control

Active and closed   Archived   Filters   [Search]   Add a pool

Pools can be archived manually or automatically (automatic archiving applies to pools with no activity for 30 days).

Title ⇕	Priority ⇕	Progress	Status ⇕	Started ⇕	To be completed

To launch a project, you first need to add a pool, set user filters and quality control rules, and upload tasks.

50 ⌄

Pool: configure name and description

POOL NAME (VISIBLE ONLY TO YOU) ❓   Are there shoes in the picture? ✕

☑ Use project description

PUBLIC DESCRIPTION ❓   Look at the picture and decide if there are shoes in it.

☐ Add a private description

Pool: configure pricing and overlap

Price per task suite

Each task suite can have one or multiple tasks on the same page. Enter the total price for all tasks in the suite.

PRICE IN US DOLLARS ? | 0.01   ✕ | FEE ?   0.005

[+ Dynamic pricing]

Overlap

Specify how many performers you want to complete each task in the pool.

OVERLAP ? | 3   ✕ |

DYNAMIC OVERLAP ? | Off |

Pool: configure timing and post verification

Parameters

TIME PER TASK SUITE IN SECONDS ❓ 300 ✕

KEEP TASK ORDER ❓ No

POOL CLOSING DATE ❓ 2022-08-24 ▦

WAITING TIME FOR THE POOL TO CLOSE IN SECONDS ❓ 0

POOL PRIORITY WITHIN THE PROJECT ❓ 0

Pool: filter performers by their profiles

Performers

Filter performers who can access the task.
Toloka has users from different countries,
so don't forget to filter by language and region. **Learn more**

ADULT CONTENT ? | Yes |

| Add filter ▾ | | Create a skill |

PERFORMER PROFILE

| Languages ▾ | = | English ✕ | 🗑 +

Pool: filter performers by computed properties

COMPUTED

| Browser | ⌄ | = | YANDEX_BROWSER | ⌄ | 🗑 | + |

AND

COMPUTED

| Client | ⌄ | = | Toloka web version | ⌄ | 🗑 | + |

Pool: filter performers by custom skills

| Add filter ▾ | Create skill |

SKILLS

| Segmentation ✕ | > | 70 ✕ | 🗑 + |

Pool: control quality via a rule on golden set

GOLDEN SET ?

History size | items

If | Number of responses | ∨ | ≥ | 3 | ✕ | +

and | % correct responses | ∨ | < | 60 | ✕

then | Ban | ∨ | on project | ∨ | 10 | ✕

Golden set | ✕

Pool: control quality via a rule on majority vote

MAJORITY VOTE ❓

Accept as majority 2 ✕

History size items

If Number of responses ⌄ ≥ 10 ✕ +

and % correct responses ⌄ < 50 ✕ 🗑

then Ban ⌄ on project ⌄ 10 ✕

Majority vote ✕

Pool: control quality via a rule on post verification

OFFLINE ACCEPTANCE ?

History size | items

If | Reviewed assignments ⌄ | ≥ | 3 ✕ +

and | % rejected assignments ⌄ | > | 35 ✕ 🗑

then | Ban ⌄ | on project ⌄ | 10 ✕

Rejected assignments ✕

Task uploading & golden set creation

Task: uploading

Does the image contain traffic lights? — closed

Statistics   ⬇ Download results ⌄   Edit ⌄

Download the sample file, add your task data, and upload the file to the pool.
The sample file uses TSV format, which is the same as CSV but using tab as the separator.
Make sure you choose UTF-8 encoding when saving the file. Learn more in the guide.

📄 Template for general tasks.tsv

📄 Template for control tasks.tsv

📄 Template for training tasks.tsv

⬆ Upload

0 task pages

0 training tasks

0 tasks

0 control tasks

0 %

Completed 0

0   0

Task: uploading

Task: edit for creation of control tasks

Does the image contain traffic lights? — closed

Statistics | ⬇ Download results ⌄ | Edit ⌄

Download the sample file, add your task data, and upload the file to the pool.
The sample file uses TSV format, which is the same as CSV but using tab as the separator.
Make sure you choose UTF-8 encoding when saving the file. Learn more in the guide.

📄 Template for general tasks.tsv

📄 Template for control tasks.tsv

📄 Template for training tasks.tsv

⬆ Upload | 📄 Files | 🗑 Delete | Edit

| 0 task pages | 0 training tasks |
| 100 tasks | 0 control tasks |

0 %

Completed 0

0   0

Task: control task creation

Edit tasks

Use main tasks as a starting point to create control tasks or training tasks.

Control tasks are for checking the quality of responses from performers. They contain correct responses to compare with actual responses.

Training tasks are for teaching performers how to complete tasks. They contain correct responses and hints.

Learn more

Main 100   Control tasks 0   Training tasks 0

| Create control tasks | | Create training tasks | | | Download |

ID ⇅	Overlap ⇅	Responses from performers ⇅	Last updated ⇅
...bdb34908	3	0	08/24/2021 4:23:52 PM

Task: create a control task by answer selection
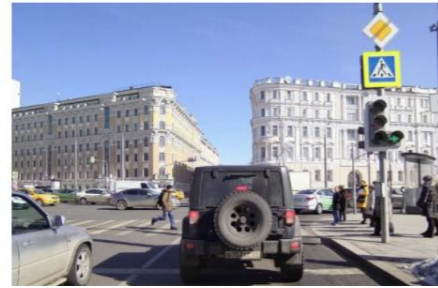
Run the pool & result downloading
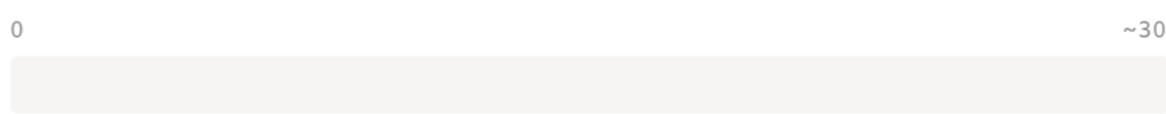
Pool running

Does the image contain traffic lights? — closed

Statistics | ⬇ Download results ⌄ | Edit ⌄

Download the sample file, add your task data, and upload the file to the pool.
The sample file uses TSV format, which is the same as CSV but using tab as the separator.
Make sure you choose UTF-8 encoding when saving the file. Learn more in the guide.

📄 Template for general tasks.tsv

📄 Template for control tasks.tsv

📄 Template for training tasks.tsv

⬆ Upload | 📄 Files | 🗑 Delete | Edit | 👁 Preview

~30 task pages

0 training tasks

90 tasks

10 control tasks

0 %
Completed 0

0   ~30

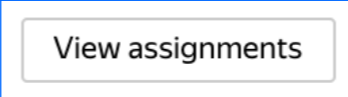# View and aggregate the results of your tasks

Monitor the statistics on how tasks are performed