Amazon Mechanical Turk
Hands-on session

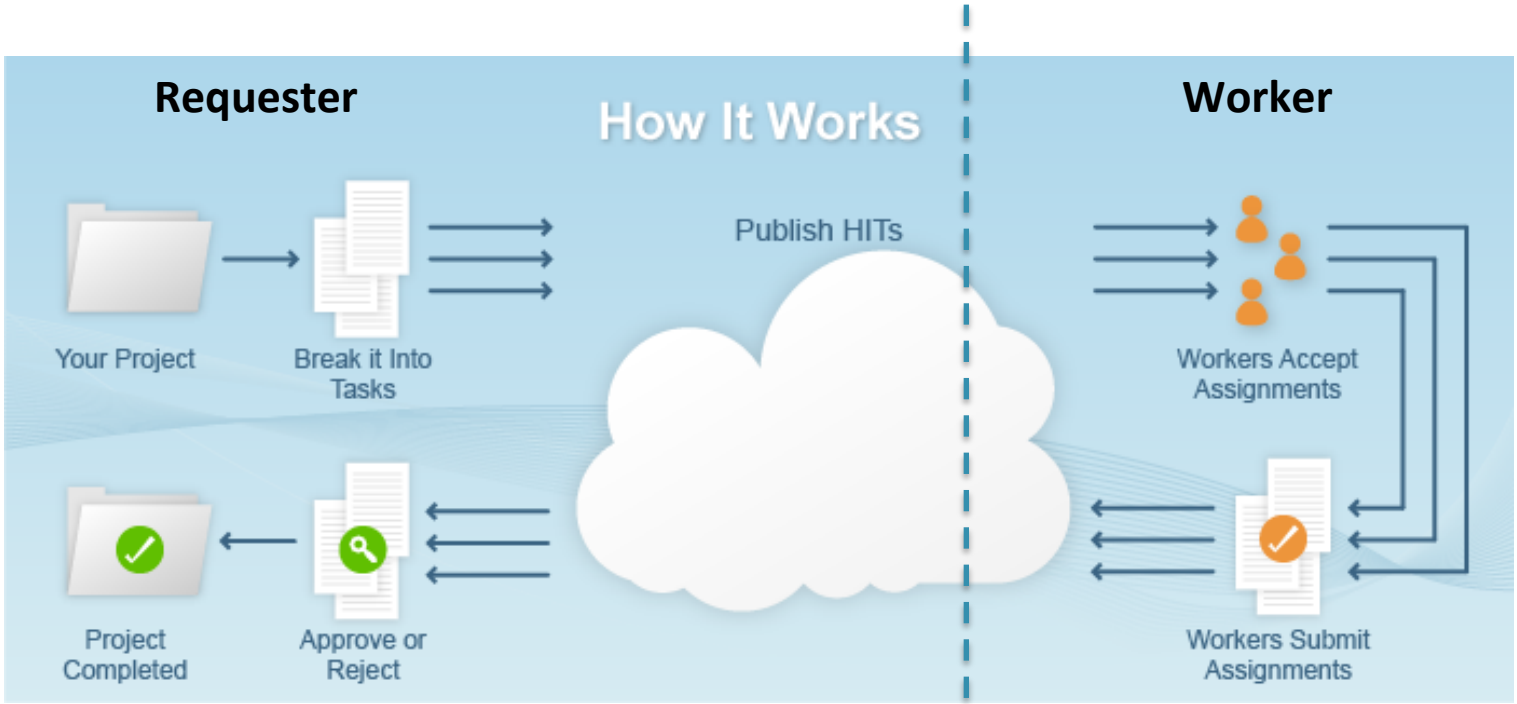
Maribel Acosta

aifb
KIT
Karlsruher Institut für Technologie

amazon mechanical turk™

Artificial Artificial Intelligence

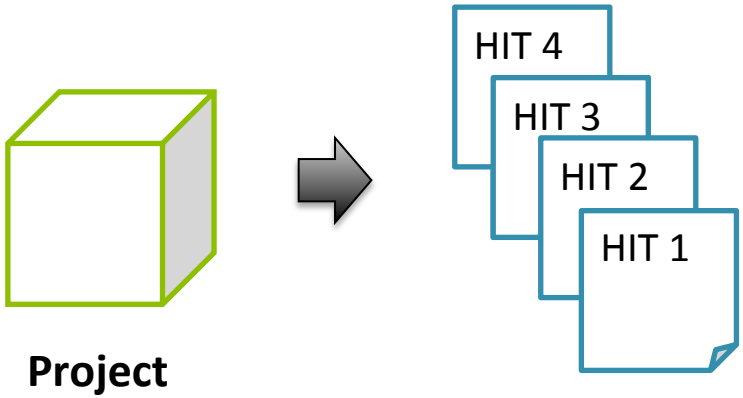
MTurk Basic Concepts (1)



Source: https://requester.mturk.com/tour/how_it_works

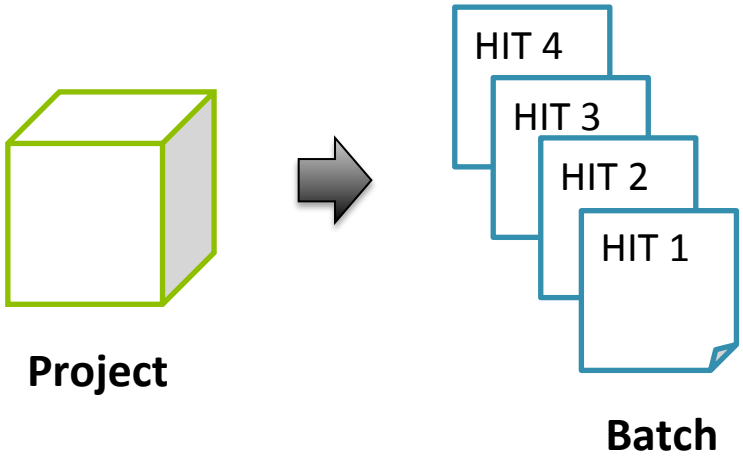
- **Requester:** creates and submits tasks to the platform.
- **Worker:** person who solves the tasks.
- **Human Intelligence Task (HIT):** work unit.

MTurk Basic Concepts (2)



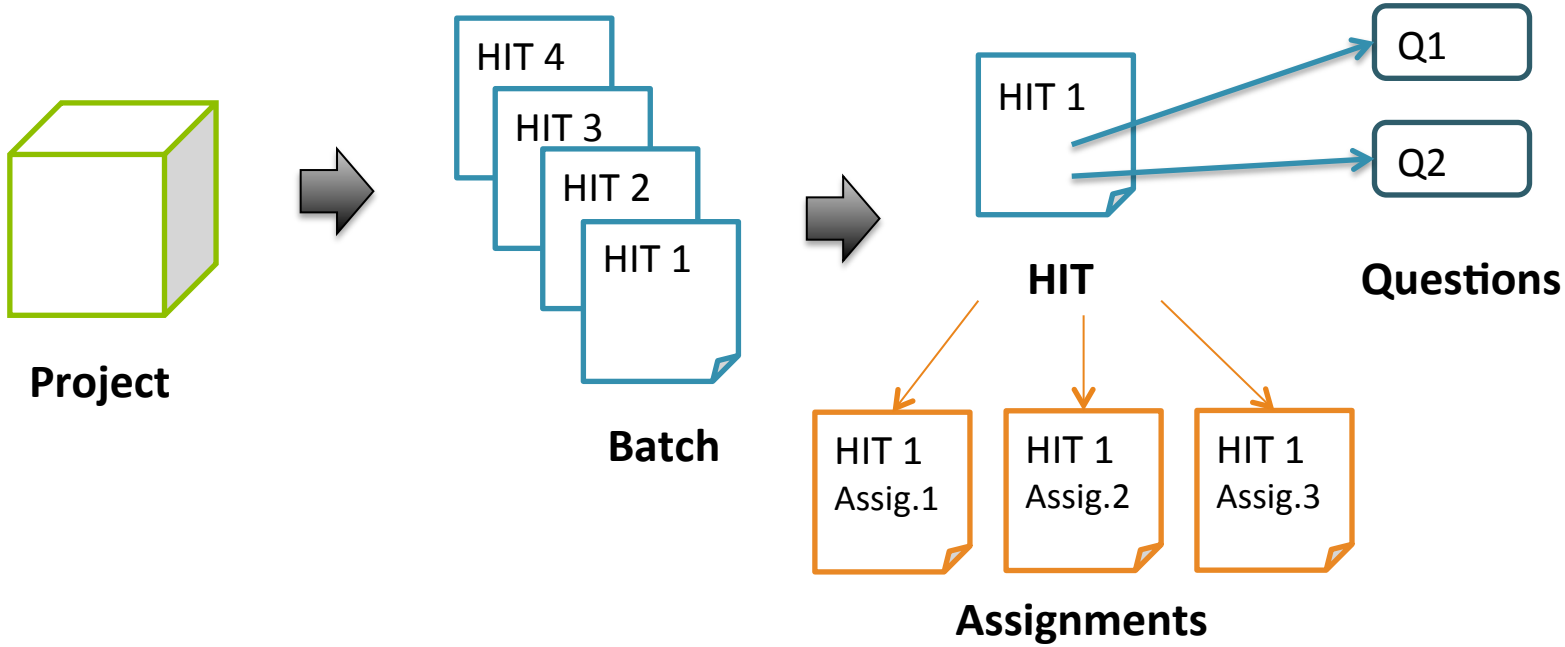
- **Project**: HIT HTML + HIT metadata
  - The elements that stay the <u>same</u> in every HIT are denominated **template**
  - The data that will <u>vary</u> from HIT to HIT are specified via **variables**
- NOTE: If no variables are specified in the project, we will create a single HIT
- **Variables:** allow creating several HITs in the project

MTurk Basic Concepts (3)



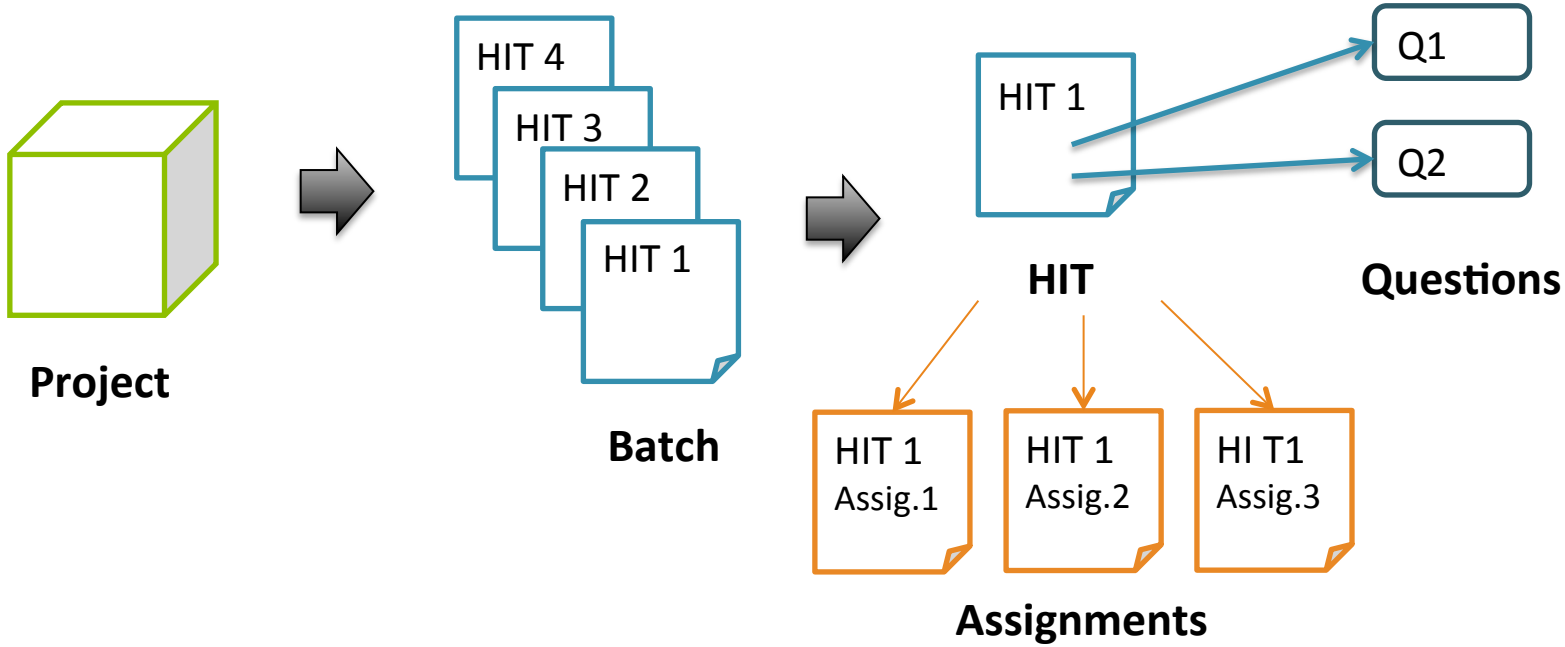
- **Batch:** Group of HITs created by instantiating the variable(s) of a project
- The values of the variables are specified in (CSV, TSV) files:
  - Each **column** corresponds to a variable
  - Each **row** is an instance -> HIT
  - Each file corresponds to a batch
- We can create **several batches** for the same project

MTurk Basic Concepts (4)



- **HIT:** Work unit. The same HIT can be solved by 1 or more workers (assignments)
- **Assignment:** How many workers should solve one exact same HIT
- **Questions:** A single HIT may contain one or several questions

MTurk Basic Concepts (5)

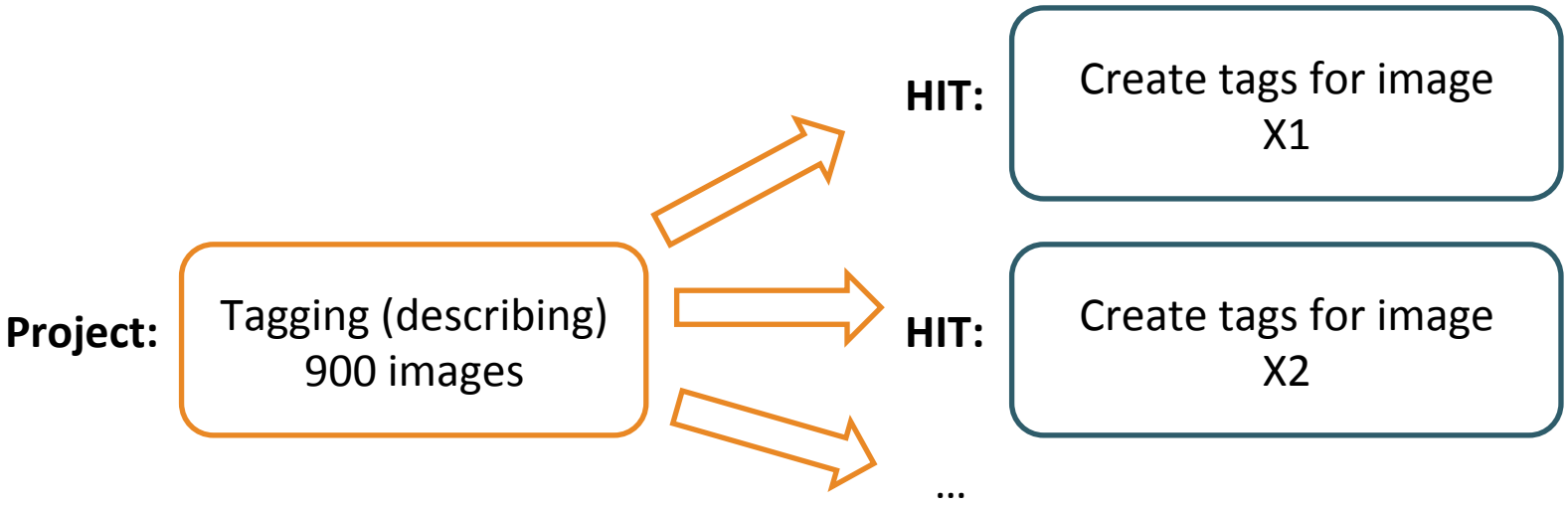


Total cost of the project = No. of HITs x No. of Assignments x (Reward per HIT + Fee)

MTurk Basic Concepts (6)

Example of Human Intelligence Tasks (HITs)

- Projects can be broken into smaller tasks called HITs
- A HIT represents a single work unit

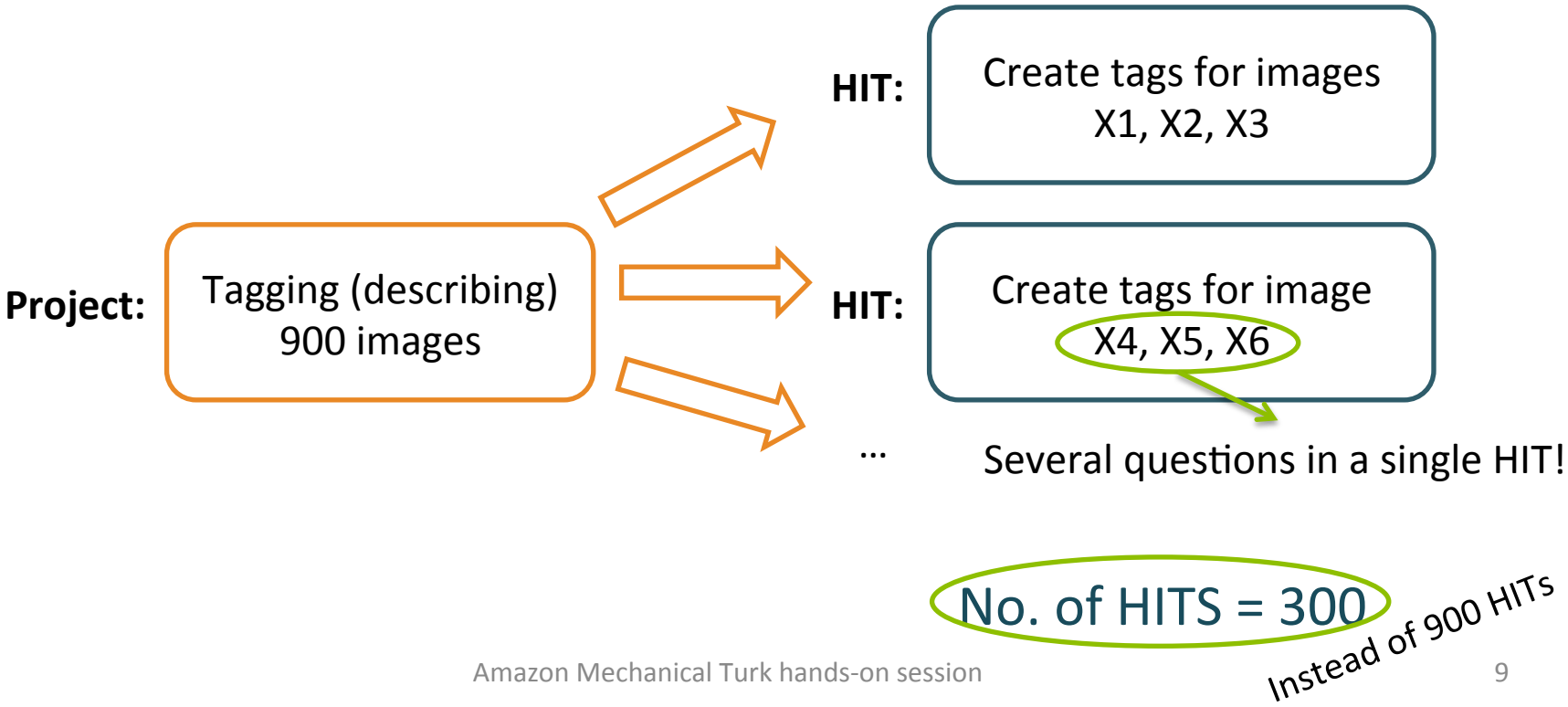


No. of HITS = 900

MTurk Basic Concepts (7)

Example of Human Intelligence Tasks (HITs)

- Projects can be broken into smaller tasks called HITs
- A HIT represents a single work unit

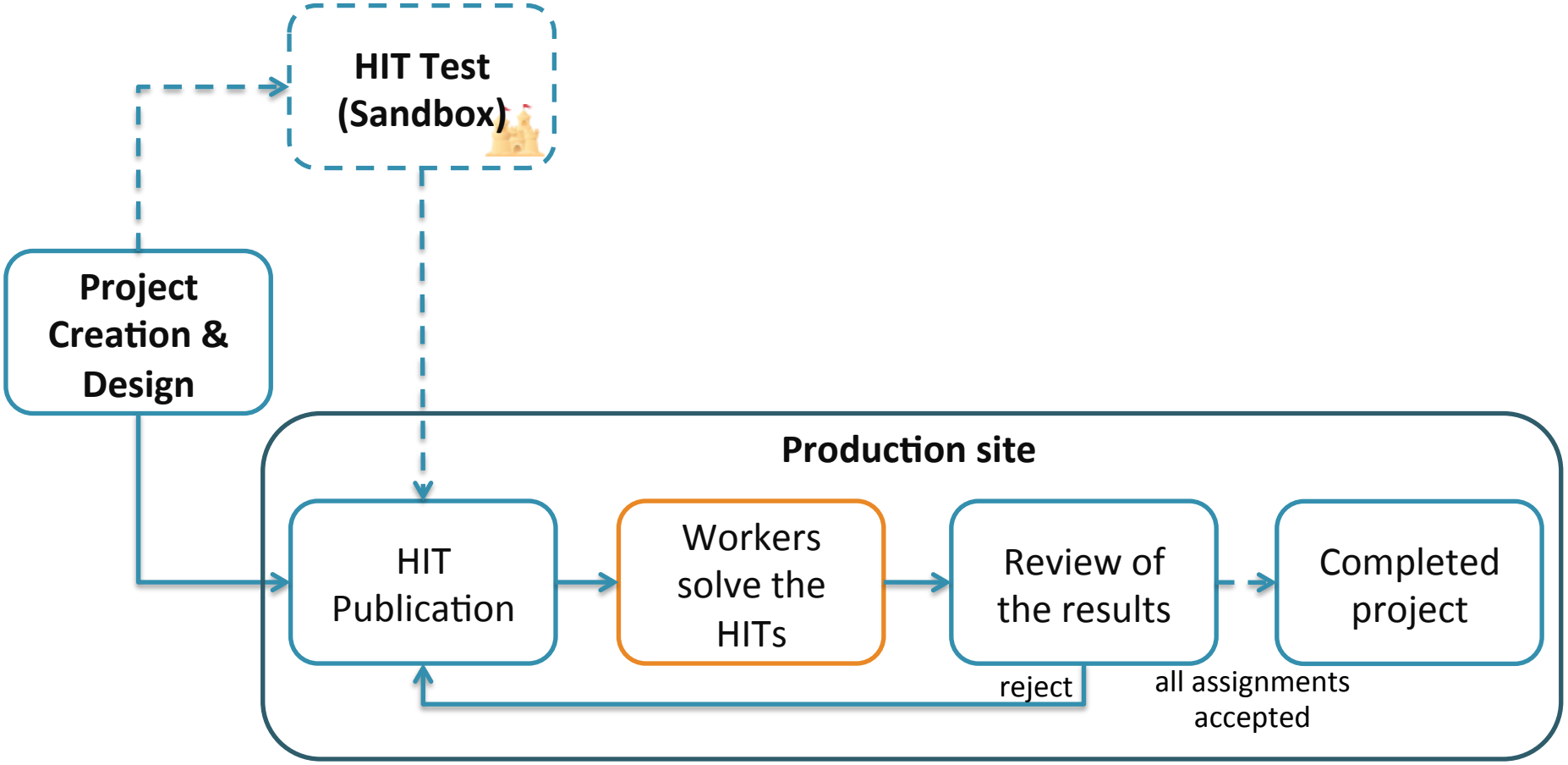


MTurk Basic Concepts (8)

When creating a project or individual HITs, the **HIT properties** must be specified:

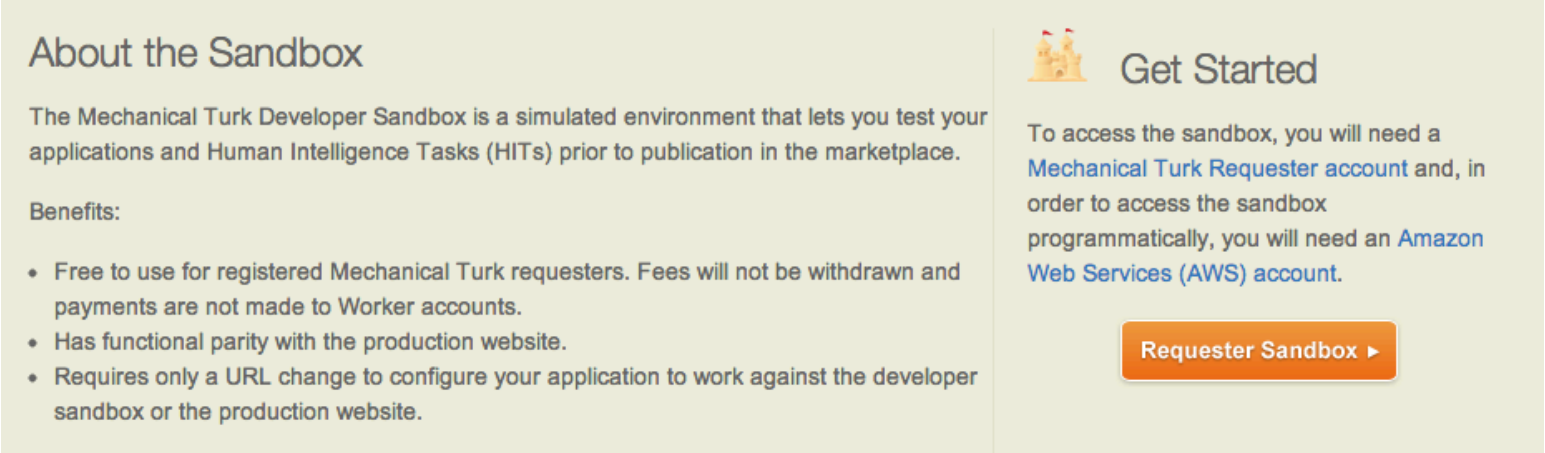
- General information: includes the title and description of the HIT, as well as keywords which are used by worker for searching HITs
- HIT duration time: time allotted to solve the HIT (before it is given to another worker)
- HIT life time: how long will the HIT be available on the platform
- # Assignments: number of different persons that will perform the exact same HIT
- Reward: payment for correctly solving each assignment

MTurk Workflow for Requesters



MTurk Sandbox

The Sandbox is a simulated MTurk environment to test HITs.



- Log in as **requester**: preview and test the interface of your HITs
  - https://requestersandbox.mturk.com
- Log in as **worker**: solve your own HITs to test their functionalities and result output
  - https://workersandbox.mturk.com
- **Best practice**: Always test your HITs (**as requester and worker**) before publishing them in the production site

Managing HITs in MTurk

There are three different mechanism to manage your HITs in MTurk:

MTURK WEB INTERFACE

Hands On!

- **Project:** Crowdsourcing DBpedia triples to verify the links to external web pages

```
prefix dbpedia-ont:<http://dbpedia.org/ontology/>
prefix foaf:<http://xmlns.com/foaf/0.1/>
SELECT *
WHERE {
 ?s dbpedia-ont:wikiPageExternalLink ?o;
    foaf:name ?s_name;
    foaf:isPrimaryTopicOf ?s_wikipage .
} LIMIT 200
```

Triple to crowdsource

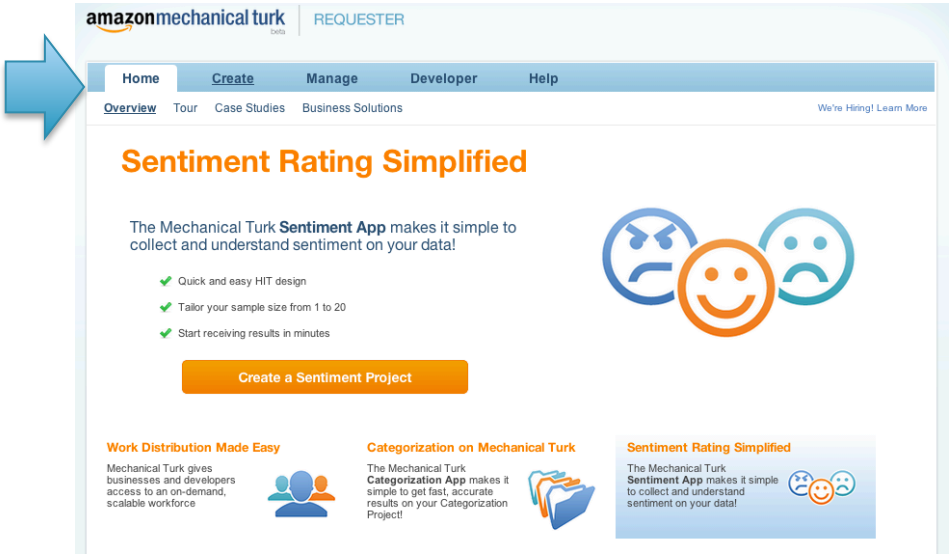
Triples to build the UI

MTurkDemo/data/sparql.csv

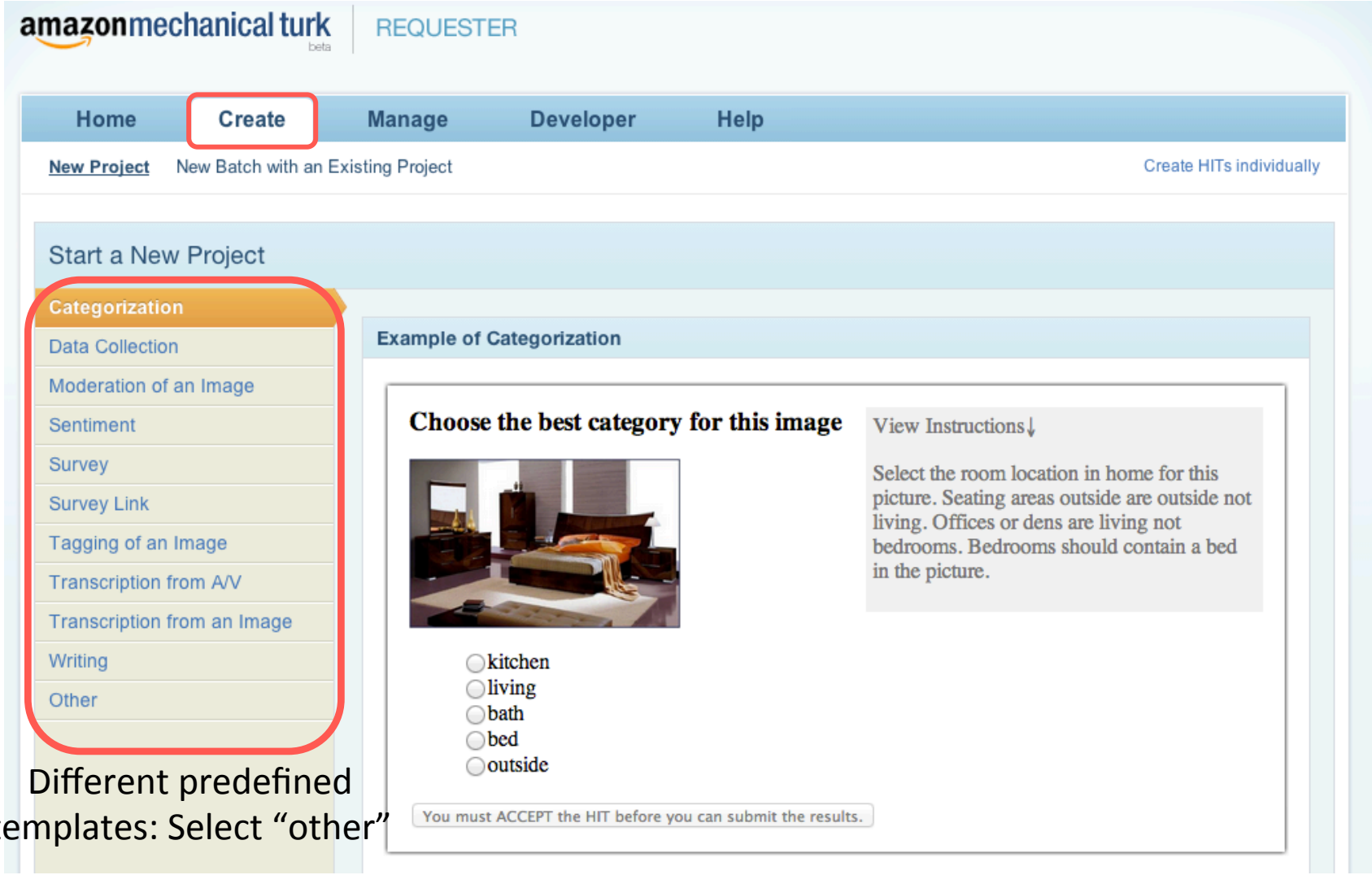
http://dbpedia.org/sparql

Hands On!

- Go to Mturk Sandbox as a **requester**:
  - https://requestersandbox.mturk.com/
- Click on **Sign In**
  - Email address: own_tp@gmx.li
  - Password: sourcrowd
- Now we are at "home"

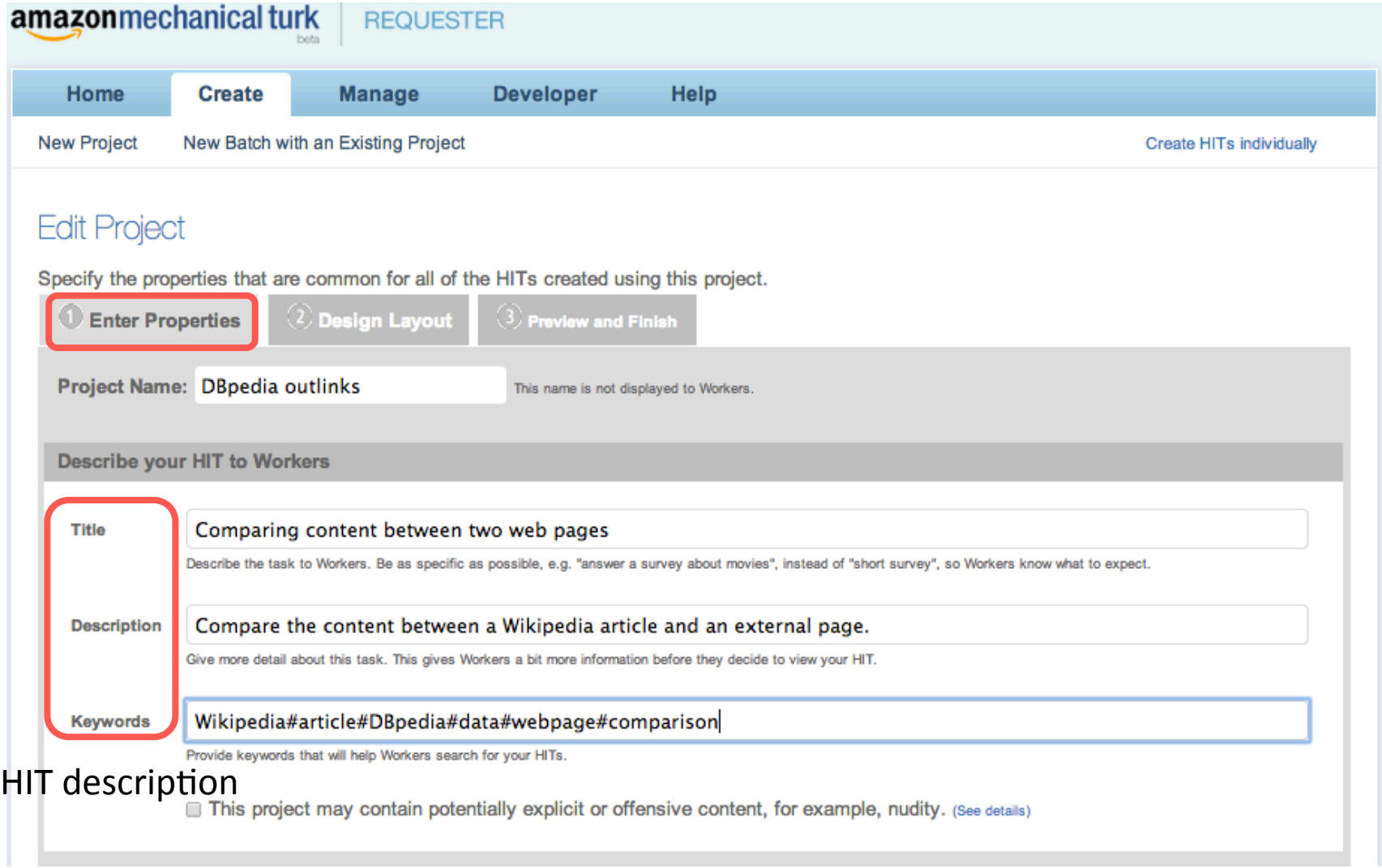


1. Creating a Project



Different predefined templates: Select "other"

2. Setting up the HIT Properties (1)



HIT description

2. Setting up the HIT Properties (2)

Setting up your HIT

Reward per assignment	$ 0.04 Tip: Consider how long it will take a Worker to complete each task. A 30 second task that pays $0.05 is a $6.00 hourly wage.
Number of assignments per HIT	3 How many unique Workers do you want to work on each HIT?
Time allotted per assignment	10 Minutes Maximum time a Worker has to work on a single task. Be generous so that Workers are not rushed.
HIT expires in	15 Days Maximum time your HIT will be available to Workers on Mechanical Turk.
Results are automatically approved in	15 Days After this time, all unreviewed work is approved and Workers are paid.

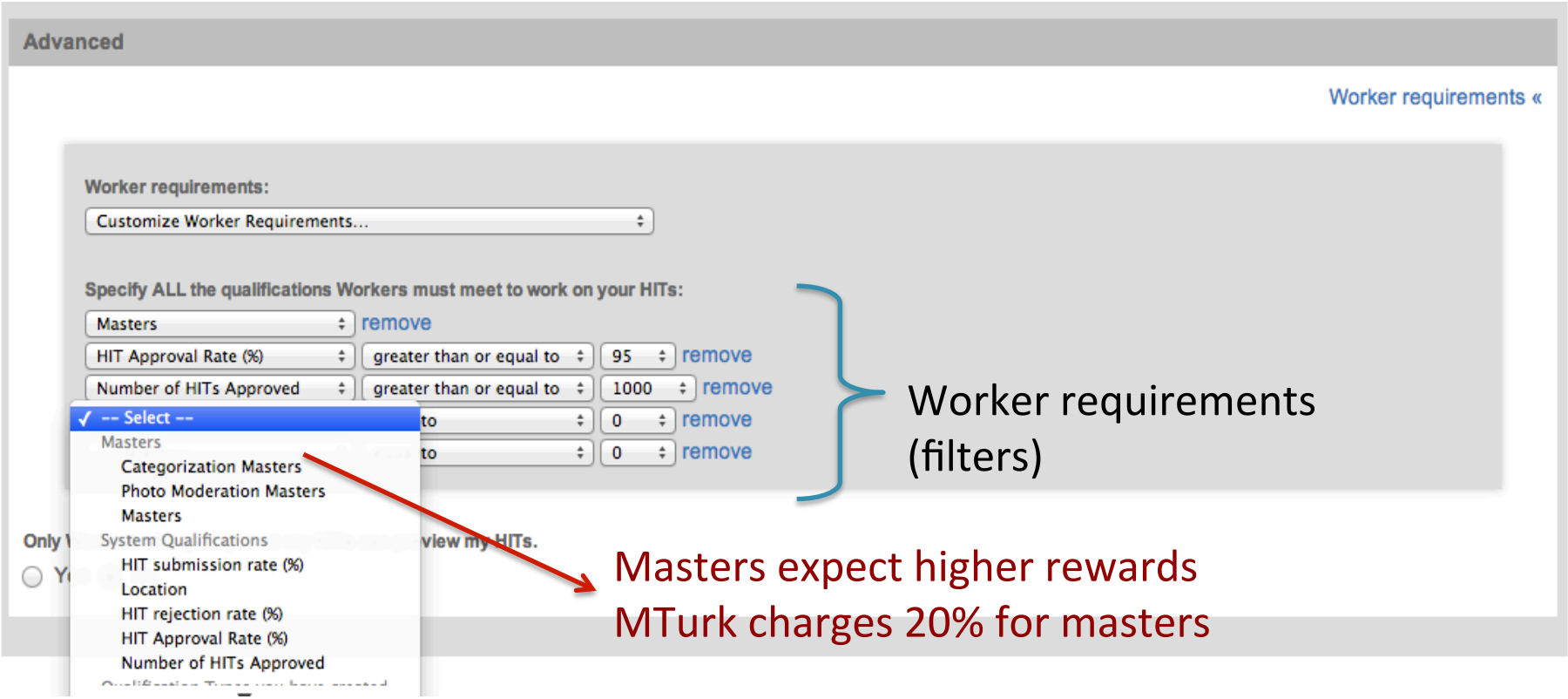
HIT properties

Very **IMPORTANT**:
Set up quality mechanisms
Masters are selected by default

Advanced »

Save | Design Layout

3. Selecting Qualifications



4. Defining the Task

Edit Project

Use the HTML editor below to design the layout of your HIT. This layout is common for all of the HITs created with this project. You can define variables for data that will vary from HIT to HIT (Learn more).

1 Enter Properties | 2 Design Layout | 3 Preview and Finish

Project Name: DBpedia outlinks — This name is not displayed to Workers.

Frame Height 400 — Height in pixels of the frame your HIT will be displayed in to Workers. Adjust the height appropriately to minimize scrolling for Workers.

Format | Font | U I B A | Source

WYSIWYG HTML editor

Instructions

Jump to questions

In this task, you will help us verify whether the links to external pages contained in Wikipedia articles are correct or not. The content of these external pages should be related to the content of the Wikipedia article. In this task, you will verify whether this is the case or not. In order to solve this task, we will provide the Wikipedia article and an external website that the article links to.

Your job: Compare whether the Wikipedia article and the external website are related.
Try to **refresh the page** if the content is not displayed properly.

Example of incorrect data

In the following example, we are checking whether the external web pages are related to the Wikipedia article "John Two-Hawks".

body h2

Save | Preview

4. Defining the Task (with Variables)

Edit Project

Use the HTML editor below to design the layout of your HIT. This layout is common for all of the HITs created with this project. You can define variables for data that will vary from HIT to HIT (Learn more).

1 Enter Properties | 2 Design Layout | 3 Preview and Finish

Project Name: DBpedia outlinks — This name is not displayed to Workers.

Frame Height: 400 — Height in pixels of the frame your HIT will be displayed in to Workers. Adjust the height appropriately to minimize scrolling for Workers.

Normal | Font | Source

Question

The content in "External page" corresponds to the topic covered in the Wikipedia article?

About:

${s_name}

Wikipedia article about: ${s_name}

Template: elements that stay the <u>same</u> in every HIT

Variables: data that will <u>vary</u> from HIT to HIT. Are denoted as follows: ${var_name}

5. Previewing the Template

Edit Project

This is how your HIT will look to Mechanical Turk Workers. Before you publish these HITs, any variables in the HIT will be replaced with the input data that you provide when you publish the HIT. You can download a sample of the input file for this project or learn more about acceptable file formats

Download sample

1 Enter Properties | 2 Design Layout | 3 Preview and Finish

This is what the workers will see

Project Name: DBpedia outlinks This name is not displayed to Workers.

Comparing content between two web pages

Requester: mac21 Reward: $0.04 per HIT HITs available: 0 Duration: 10 Minutes

Qualifications Required: HIT Approval Rate (%) greater than or equal to 95 , Number of HITs Approved greater than or equal to 1000

HIT Preview

Question

The content in "External page" corresponds to the topic covered in the Wikipedia article?

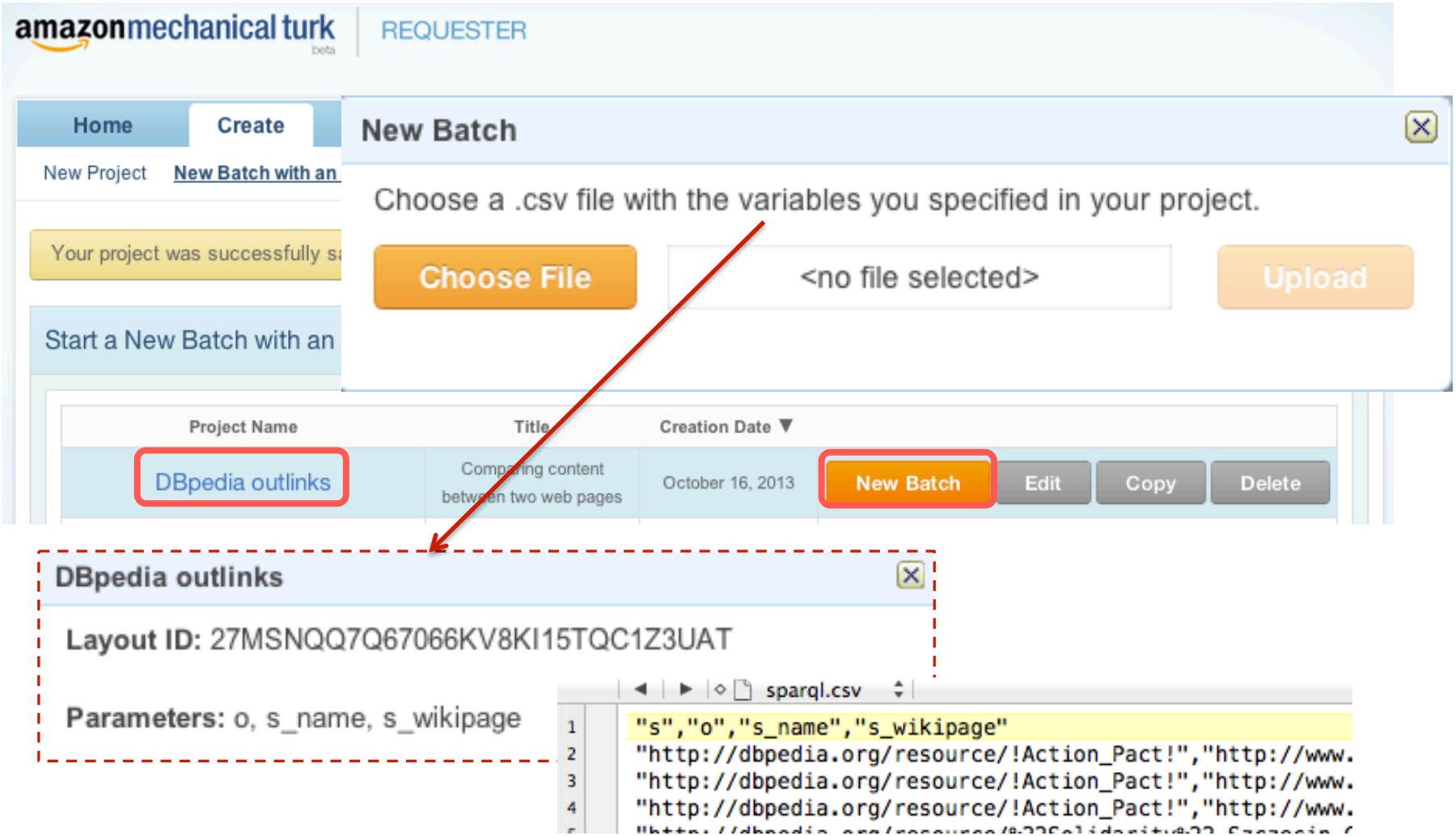
About:

${s_name}

The variables will be replaced by the input data

Wikipedia article about: ${s_name}

6. Creating Batches



7. Previewing the HITs

DBpedia outlinks

Comparing content between two web pages

Requester: mac21 **Reward:** $0.04 per HIT **HITs available:** 200 **Duration:** 10 Minutes

Qualifications Required: HIT Approval Rate (%) greater than or equal to 95 ,

Number of HITs Approved greater than or equal to 1000

HIT Preview

Question

The content in "External page" corresponds to the topic covered in the Wikipedia article?

About:

!Action Pact!

Wikipedia article about: !Action Pact!

Create account Log in

Article Talk Read Edit Search

!Action Pact!

Variables are replaced with the data from the input file

Showing HIT 1 of 200 Next HIT

8. Publishing the HITs

DBpedia outlinks

Batch Summary

HITs		
Number of HITs in this batch:		200
Number of assignments per HIT:	x	3
Total number of assignments in this batch:		600
Cost		
Reward per Assignment:		$0.040
	x	600
Estimated Total Reward:		$24.000
Estimated Fees to Mechanical Turk:	+	$3.000
Estimated Total Cost:		$27.000
Your Available Balance:		$10,000.000
Your Projected Balance:		$9,973.000

Estimated Total Cost: $27.000 (this is the amount that will be deducted from your Available Balance when you click "Publish HITs")

Your Available Balance: $10,000.000 (before clicking "Publish HITs")

Your Projected Balance: $9,973.000 (after clicking "Publish HITs")

Summary of the project:

- # of HITs
- Rewards
- Total payment
- Account balance

9. Retrieving the Results

Review Results

Select the check boxes on the left to approve or reject results. You only pay for approved results. To evaluate results offline, select Download CSV.

For additional batch information, view batch details.

DBpedia outlinks 10

Customize View | Filter Results | Upload CSV | Download CSV

1 of 1 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Approve | Reject

	HIT ID ▲	Worker ID	Lifetime Approval Rate	Input Data	Q1
☐	2N6Q1SNQQ7Q685ZCMH3NU12T1XJS8I	A3QBKUNGNMLHTF	98% (131/133)	BatchId:58757;	correct
☐	HIT ID ▲	Worker ID	Lifetime Approval Rate	Input Data	Q1

Approve | Reject

MTURK SUMMARY

Project/HIT Creation & Design (1)

- The requester is able to create **projects** or **individual HITs**
- Build **user-friendly interfaces** (using web technologies)
- Then, the **HIT properties** must be specified:
  - General information: includes the title and description of the HIT, as well as keywords which are used by worker for searching HITs.
  - HIT duration time: time allotted to solve the HIT (before it is given to another worker).
  - HIT life time: how long will the HIT be available on the platform.
  - # Assignments: number of different persons that will perform the same HIT.
  - Reward: payment for correctly solving each assignment.

Project/HIT Creation & Design (2)

- Selection of **MTurk quality control** mechanisms:

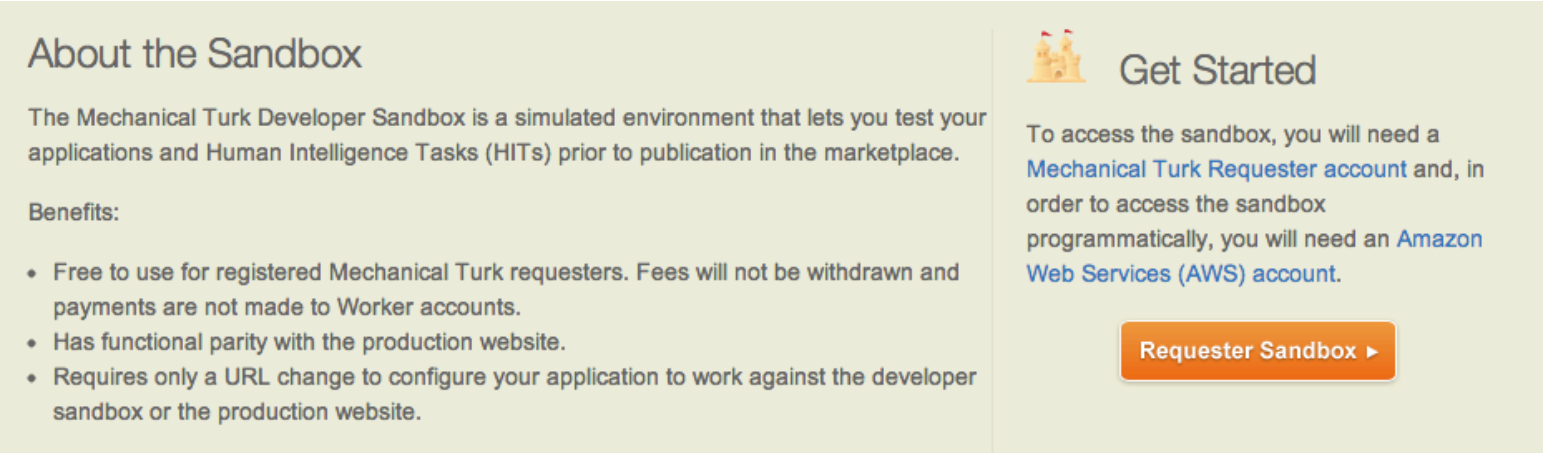
Worker requirements

- **High quality workers**
  - Masters
  - Photo moderation masters
  - Categorization masters
    Masters expect higher rewards
    MTurk charges 20% for masters
- **System qualifications**
  - Location by country
  - HIT submission rate (%)
  - HIT approval/rejection rate (%)
  - (Absolute) Number of HITs approved
- **Qualification types**
  - Simply granted or attributed via customized tests

- These filters are automatically performed by the platform

HIT Test

- **Best practice:** Always test your HITs before publishing them
  1. Perform **technical tests** (both as requester and worker) in the MTurk Sandbox environment.



Source: https://requester.mturk.com/developer/sandbox

  2. Publish a small subset of tasks in the production site to test **usability** and **responsiveness**.

Run live HITs

- **HIT publication:**
  - Make the HITs available to the workers

- **Review the results:**
  - Monitor the submitted assignments constantly
  - Download the results
  - Accept/reject assignments, provide feedback when rejecting
  - Block spammers (optional)

- **Update HIT/Project:**
  - Extend/expire HITs or modify other HIT properties
  - Add additional assignments

Lessons Learned

- Introduce yourself on Worker forums (regular requester)

- Be responsive to workers
  - Reply to emails with questions about tasks

- Use monitoring tools:
  - Forums
  - Turkopticon (Source:http://turkopticon.differenceengines.com/)

Choosing the Right Tool

Tool Comparison Table

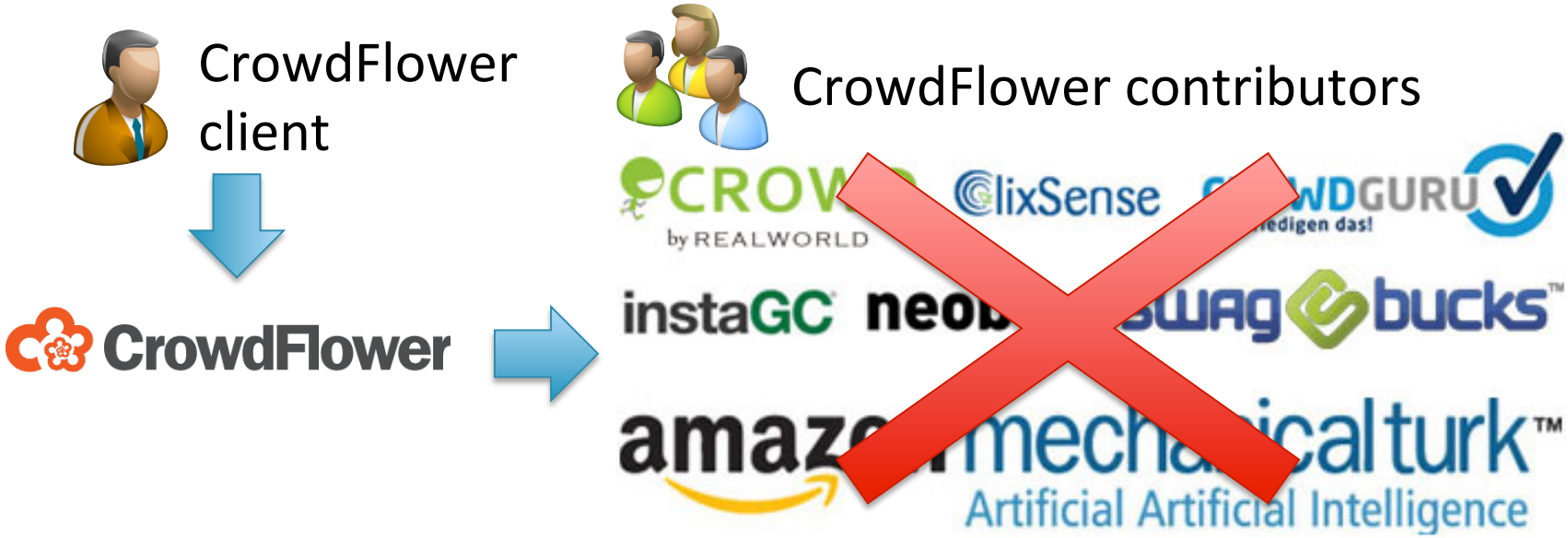
	Web Interface	Command Line Tools	API
Creating and managing your work			
Start with our sample HTML templates	✔		
Create HITs visually with an HTML editor	✔		
Create and manage your HITs in batches	✔	✔	
Support tab-delimited input files	✔	✔	
Manage HITs created via the CLT or API		✔	✔
Define HITs in XML		✔	✔
Host HITs on your own server		✔	✔
Can be integrated into back-end systems			✔
Create notifications indicating when HITs are updated			✔

Source: https://requestersandbox.mturk.com/tour/choose_the_right_tool

IS THERE MORE THAN MTURK?

CrowdFlower

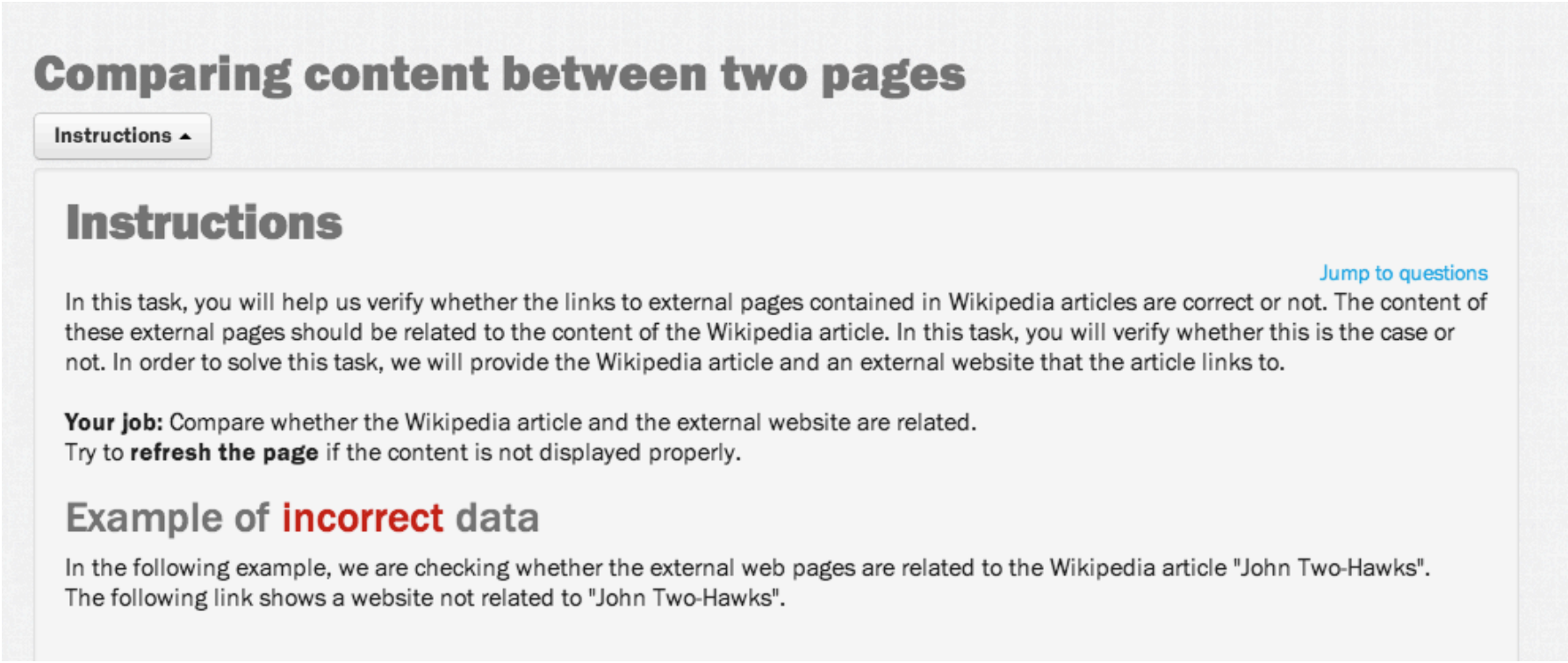
CrowdFlower Platform



- **Client:** Creates and submits jobs (MTurk = requester)
- **Contributor:** person who solves the jobs (MTurk = worker)
- **Job:** unit work (MTurk = task)

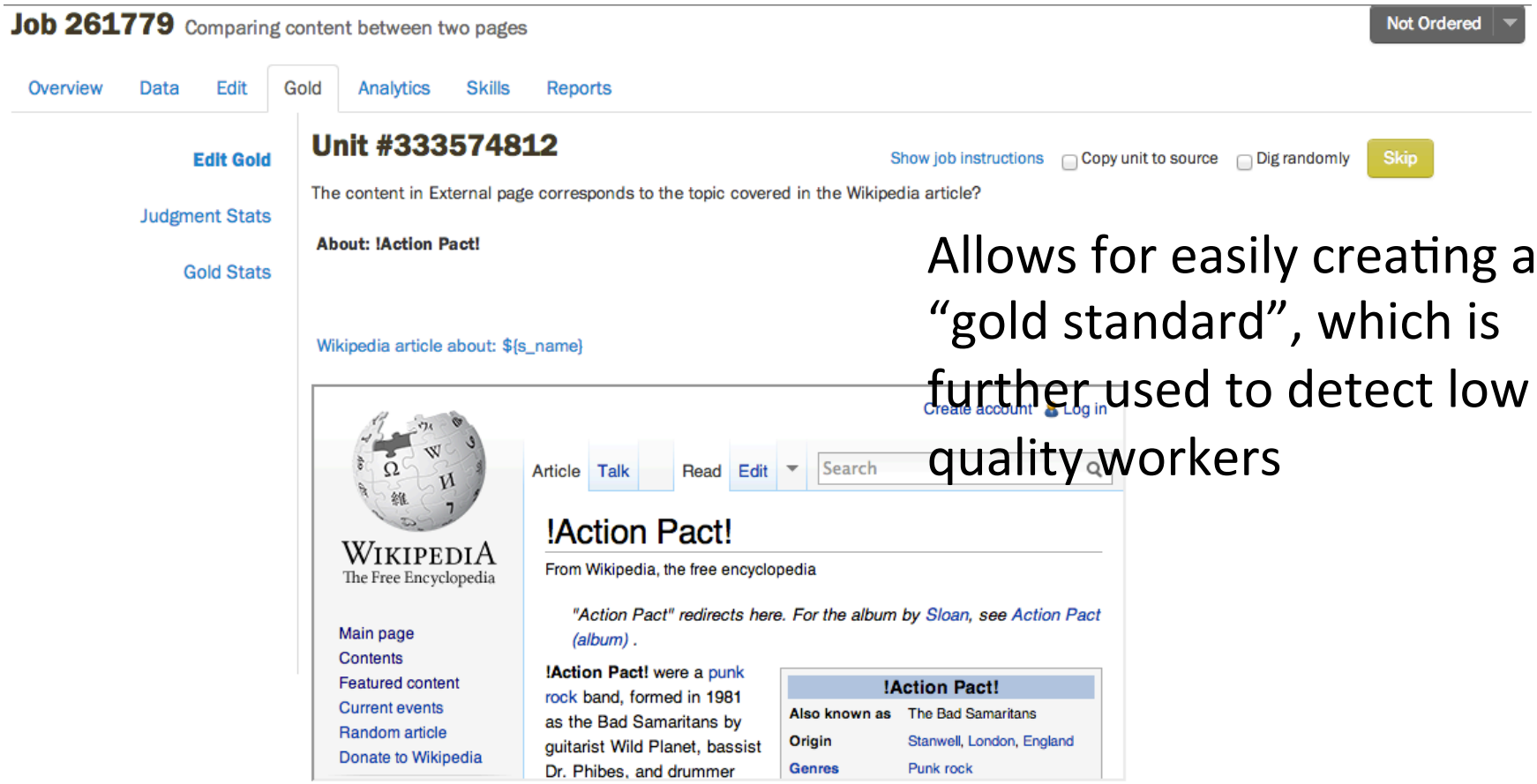
Why CrowdFlower? (1)

Neat UI



Why CrowdFlower? (2)

Quality control mechanisms



Allows for easily creating a "gold standard", which is further used to detect low quality workers

Why CrowdFlower? (3)

Report generation and analytics

Job 230234 Categorization of fashion images

Finished

Overview | Data | Edit | Gold | Contributors | Analytics | Skills | Reports

Summary

Contributors

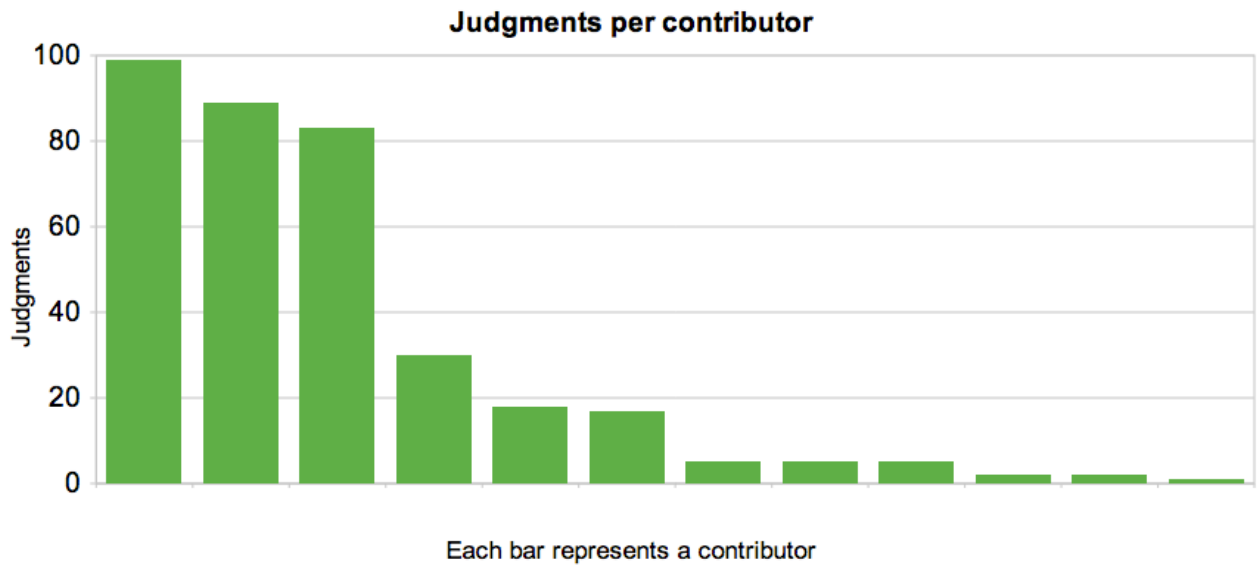
Quality

Gold

Distributions

Times

Each bar represents a contributor and the number of judgments they have submitted for this job. Contributors who have a low trust score and have submitted a significantly larger amount of judgments than other contributors are likely scammers. If your gold is working correctly, their work will be rejected. *(Only the top 100 contributors are displayed in this graph.)*



Why NOT CrowdFlower?

- At the beginning, clients must wait until their **projects are approved** by the CrowdFlower staff before publishing them
  - Wait time: From a couple of hours up to (5)* days

- Jobs must be specified in a **non-standard language**:
  - CML: CrowdFlower Markup Language

- There are certain **configurations** that cannot be executed in the platform

*Personal experience of the presenter

References

- AMT. Getting Started Guide. API Version 2012-03-25
  http://s3.amazonaws.com/awsdocs/MechTurk/latest/amt-gsg.pdf

- The Mechanical Turk Blog
  http://mechanicalturk.typepad.com/

- MTurk Java API
  http://people.csail.mit.edu/glittle/MTurkJavaAPI/

- CrowdFlower Platform
  http://crowdflower.com