Amazon Mechanical Turk
Hands-on session

Maribel Acosta
MTurk Basic Concepts (1)

• **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 same in every HIT are denominated **template**
  - The data that will vary 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)

Project → Batch → HIT → Questions

Assignments

HIT 4
HIT 3
HIT 2
HIT 1

HIT 1

Q1
Q2

HIT 1
Assig.1

HIT 1
Assig.2

HI T1
Assig.3

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

![Diagram of project and HITs]

- Project: Tagging (describing) 900 images
- HIT 1: Create tags for image X1
- HIT 2: Create tags for image X2
- ... 

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

Project: Tagging (describing) 900 images

HIT: Create tags for images X1, X2, X3

HIT: Create tags for image X4, X5, X6

⋯ Several questions in a single HIT!

No. of HITS = 300

Amazon Mechanical Turk hands-on session
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

Project Creation & Design

HIT Test (Sandbox)

Production site

HIT Publication
Workers solve the HITs
Review of the results
Completed project

reject
all assignments accepted
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](https://requestersandbox.mturk.com)
- Log in as **worker**: solve your own HITs to test their functionalities and result output
  - [https://workersandbox.mturk.com](https://workersandbox.mturk.com)
- **Best practice**: Always test your HITs (as requestor and worker) before publishing them in the production site
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

```sparql
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
```

MTurkDemo/data/sparql.csv

http://dbpedia.org/sparql
Hands On!

- Go to Mturk Sandbox as a **requester**:
  - [https://requestersandbox.mturk.com/](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**

- **Title**: Comparing content between two web pages
  - Describe the task to Workers. Be as specific as possible, e.g. "answer a survey about movies", instead of "short survey", so Workers know what to expect.

- **Description**: Compare the content between a Wikipedia article and an external page.
  - "This gives Workers a bit more information before they decide to view your HIT."

- **Keywords**: Wikipedia#article#DBpedia#data#webpage#comparison
2. Setting up the HIT Properties

<table>
<thead>
<tr>
<th>HIT properties</th>
<th>Very IMPORTANT:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reward per assignment</strong></td>
<td>Set up quality mechanisms</td>
</tr>
<tr>
<td><strong>Number of assignments per HIT</strong></td>
<td>Masters are selected by default</td>
</tr>
<tr>
<td><strong>Time allotted per assignment</strong></td>
<td></td>
</tr>
<tr>
<td><strong>HIT expires in</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Results are automatically approved in</strong></td>
<td></td>
</tr>
</tbody>
</table>
3. Selecting Qualifications

Worker requirements (filters)

Masters expect higher rewards
MTurk charges 20% for masters
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).


Project Name: DBpedia outlinks

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.

Instructions

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".
4. Defining the Task (with Variables)

**Template:** elements that stay the same in every HIT

**Variables:** data that will vary from HIT to HIT. Are denoted as follows: `${\text{var\_name}}$"
5. Previewing the Template

This is what the workers will see

The variables will be replaced by the input data.
6. Creating Batches

Amazon Mechanical Turk hands-on session
7. Previewing the HITs

Variables are replaced with the data from the input file.
8. Publishing the HITs

<table>
<thead>
<tr>
<th>DBpedia outlinks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch Summary</td>
</tr>
<tr>
<td>HITs</td>
</tr>
<tr>
<td>Number of HITs in this batch:</td>
</tr>
<tr>
<td>Number of assignments per HIT:</td>
</tr>
<tr>
<td>Total number of assignments in this batch:</td>
</tr>
<tr>
<td>Cost</td>
</tr>
<tr>
<td>Reward per Assignment:</td>
</tr>
<tr>
<td>Estimated Total Reward:</td>
</tr>
<tr>
<td>Estimated Fees to Mechanical Turk:</td>
</tr>
<tr>
<td>Estimated Total Cost:</td>
</tr>
</tbody>
</table>

**Summary of the project:**
- # of HITs
- Rewards
- Total payment
- Account balance

Amazon Mechanical Turk hands-on session 26
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**

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

<table>
<thead>
<tr>
<th>HIT ID ▲</th>
<th>Worker ID</th>
<th>Lifetime Approval Rate</th>
<th>Input Data</th>
<th>Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2N6Q1SNQQ7Q685ZCMH3NU12T1XJS8l</td>
<td>A3QBKUNGNMLHTF</td>
<td>98% (131/133)</td>
<td>BatchId:56757; correct</td>
<td></td>
</tr>
</tbody>
</table>
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:
  
  - **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

---

**Worker requirements**

Amazon Mechanical Turk hands-on session
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](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

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

Source: [https://requestersandbox.mturk.com/tour/choose_the_right_tool](https://requestersandbox.mturk.com/tour/choose_the_right_tool)

Amazon Mechanical Turk hands-on session
IS THERE MORE THAN MTURK?
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

Comparing content between two pages

Instructions

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". The following link shows a website not related to "John Two-Hawks".
Why CrowdFlower? (2)

Quality control mechanisms

Jobs 261779 Comparing Content between two pages

- Overview
- Data
- Edit
- Gold
- Analytics
- Skills
- Reports

Unit #333574812

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

About: Action Pact!

Wikipedia article about: $ja_name$

Action Pact! were a punk rock band, formed in 1981 as the Bad Samaritans by guitarist Wild Planet, bassist Dr. Phibes, and drummer...

Allowes 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

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

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

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

• CrowdFlower Platform
  http://crowdflower.com