

# Crowd-sourcing

- What is it good and not good for?

# Benoit et al.

- What is the alternative?
  - Expert manual annotation is expensive and rare
    - Valid and reliable
  - Non-expert crowdsourcing is cheap and abundant
    - Large N gives us central tendency and error distribution
    - Valid
- Go over design, and show results
- Why not machine?
- Keys to a good design?

# Categorization on Mechanical Turk

The Mechanical Turk **Categorization App** makes it simple to get fast, accurate results on your Categorization project!

- ✔ Quick and easy HIT design force up and down quickly
- ✔ Pre-qualified workers quality, cost-effective results
- ✔ Start receiving results in minutes minutes

Create a Categorization Project



*Crowdsourcing: Artificial artificial intelligence*

Categorization, Sentiment, Survey Experiments, Translation.....

Documentation: <http://docs.aws.amazon.com/AWSMechTurk/latest/RequesterUI/Welcome>

test — Choose the best category:

[How does it work?](#)

Categoryization Project

Choose Categories

Provide Instructions

Upload Data File

Preview

Checkout

Project

What would you like to call your project?

Categories

Provide the names of the top-level categories and subcategories that you want Workers to categorize items into.

Categories

[Undo Subcategories](#) ◀

[Add Category \(max=10\)](#)

Subcategories

[Add Subcategory \(max=10\)](#)

Note max of 10 categories and 10 subcategories

[Use Advanced Input dialog](#)

**Categorization Project**

Choose Categories

**Provide Instructions**

Upload Data File

Preview

Checkout

**Instructions**

Provide general instructions to workers doing your HIT.

**Selection Criteria**

Workers do best if you provide selection criteria on your categories. Specify what to include and/or exclude for those categories that might be ambiguous to Workers. ([see example](#))

**Categories**

b

a

**Selection Criteria for: b**

**Includes:**

**Excludes:**

Remove Criteria

Instructions and example criteria are likely to prove key to obtaining good results

a — Choose the best category:

[How does it work?](#)

**Categorization Project**

[Choose Categories](#)

[Provide Instructions](#)

**[Upload Data File](#)**

[Preview](#)

[Checkout](#)

**Upload Data File**

Please provide a .csv file of your data so that we can create your work items. [\(learn more\)](#)

**Choose File**

<no file selected>

**Upload**

After your file is uploaded, you can specify which columns to show Workers.

**Next »**

[Leave feedback for this page.](#)

Data to be labeled come in the form of a .csv file. Could be text, a link to outside content

**a** — Choose the best category:

[How does it work?](#)

### Categorization Project

[Choose Categories](#)[Provide Instructions](#)[Upload Data File](#)[Preview](#)[Checkout](#)

#### Upload Data File - Choose Fields

Each column from your .csv file is shown as a field below. Tell us which fields you want to show Workers and whether the fields contain text, a link to an image, or a link to a website.

Optionally, you can add a label to appear next to your data. For instance, if the field is a "description," you might want to add a label of "Description."

Field	Show Workers?	Type of Data	Label
Name	<input checked="" type="checkbox"/>	<Select type of data> <input type="button" value="v"/>	<input type="text"/>

If this doesn't look right, you can [upload a different file](#).

[Next »](#)[Leave feedback for this page.](#)

a — Choose the best category:

[How does it work?](#)**Categorization Project**[Choose Categories](#)[Provide Instructions](#)[Upload Data File](#)[Preview](#)[Checkout](#)**Preview of Work Items**

This is what Workers will see.

**[+] Instructions** [\(Open full instructions in a separate window\)](#)**name:** Tim Holden**Choose a category**

The interface Turkers see



[Home](#)[Create](#)[Manage](#)[Developer](#)[Help](#)[New Project](#) [New Batch with an Existing Project](#)[Create HITs individually](#)[a](#) — Choose the best category:[How does it work?](#)**Categorization Project**[Choose Categories](#)[Provide Instructions](#)[Upload Data File](#)[Preview](#)[Checkout](#)**Checkout**

Number of Items		12
Number of Workers per Item <small>(change)</small>	x	1
Number of Worker Submissions	=	12
Reward per Submission <small>(details)</small>	x	\$0.020 <input type="text"/>
Total Worker Rewards	=	\$0.240
Total Mechanical Turk fees <small>(details)</small>	+	\$0.108
<b>Total cost</b>	=	<b>\$0.348</b>

Your available balance **\$20.00**Your balance after work is completed **\$19.65**

Specify how many workers are to code each item, and how much you'll pay them  
Ability to filter quality based on "Masters" option; or based on history of a coder