

Introductory Course at ESLLI

Bolzano, Italia

August 2016



Crowdsourcing Linguistic Datasets LECTURE 2

Chris Biemann

biem@cs.tu-darmstadt.de

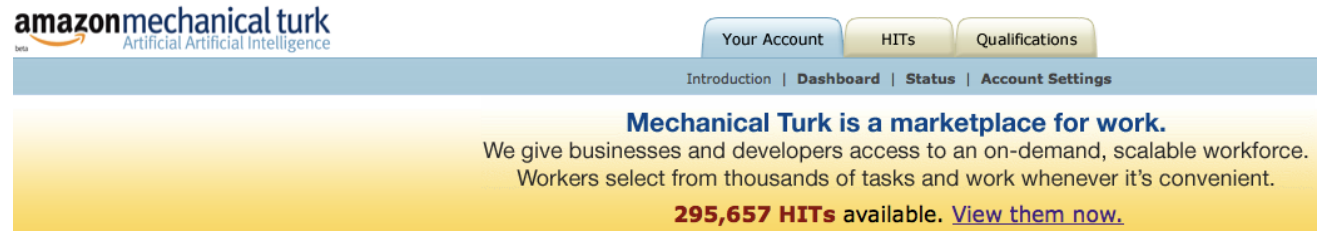
Lesson 2: Crowdsourcing platforms

- Crowdsourcing Platforms for Text Tasks
 - Amazon Mechanical Turk (AMT)
 - Crowdfunder (CF)
- A small sample \$5 project: Emotions and Associations
 - Learn Step by step how AMT and CF work
 - Discussion of Results and Differences
- A more ambitious \$40 project: Paraphrases in Context
 - use of advanced APIs
 - iteration for the evolution of crowdsourcing tasks

Crowdsourcing Platforms

- **Amazon Mechanical Turk** <https://www.mturk.com>

- since 2005
- largest and most well-known platform
- highly flexible



- **Crowdfunder** <http://crowdfunder.com>

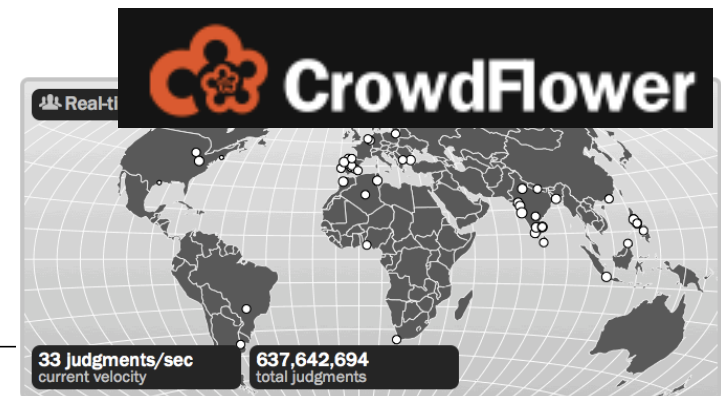
- automated quality control
- more formalized, easier to configure



- **MicroWorkers** <https://microworkers.com/>

- mostly SEO

Most platforms sell 'campaign' packages or specialize in a few 'key' areas. We concentrate on those that allow the free definition of micro-tasks.



Amazon Mechanical Turk: Characteristics

<https://www.mturk.com>

- old and popular platform
- bare-bone platform with very little in-built mechanisms for quality control
- free definition of tasks, can use tasks from your own server
- comparatively low premium of 20%, but at least 1 cent / HIT
- RESTful API
- need a US address for requester role
 - if you don't live in the US, get a forwarding address from services like <http://www.ipsparcels.com>
 - use your credit card with billing address from anywhere
- need to handle taxation bureaucracy on the requester side
 - payouts of more than \$600 / year for a single worker require a US tax form
 - avoid by tracking payouts and revoking qualifications

Summary: Full control (for the better or the worse), low premium, need some trickery for non-US-based requesters, beware the taxation issue.

Crowdfunder: Characteristics

<http://crowdfunder.com>

- established meta-platform, distributes tasks over many platforms
- many mechanisms for quality control, main: gold items
- free definition of tasks, can use java script for complex interfaces
- comparatively high premium of 33%
- RESTful API
- works from anywhere with credit card
- taxation bureaucracy handling included in service
- academic accounts with minimal support, need to make data public
- pro accounts (with private data) need a minimum cash flow of several thousand USD / month!

Summary: Good quality control handling, high premium, high amount of services, access to many platforms, issues with proprietary data, good for business customers esp. outside of the US

\$5 project: Emotion and Association

- Let's say we want to collect emotions for words
- Let's also say we're interested in associated terms
- example: for the target word LOVE, we expect highly positive emotions and associations such as MARRIAGE, ROMANCE, etc.

Applications: Psychology research, MakeYourTextNicer.com, Sentiment Classification, ...

Amazon Mechanical Turk HIT creation I

The screenshot shows the 'Create' tab of the Amazon Mechanical Turk interface. The navigation bar includes 'Home', 'Create', 'Manage', 'Developer', and 'Help'. Below the navigation bar, there are links for 'New Project' and 'New Batch with an Existing Project', and a link for 'Create HITs individually' on the right. The main content area is titled 'Start a New Project' and features a sidebar with various project categories: 'Categorization', 'Data Collection', 'Moderation of an Image', 'Sentiment', 'Survey', 'Survey Link', 'Tagging of an Image', 'Transcription from A/V', 'Transcription from an Image', 'Writing', and 'Other'. The 'Other' category is highlighted with an orange arrow. The main content area displays an 'Example of Other' project with a 'Create Project »' button. Below this, there is a section titled 'Instructions' with the following text: 'Enter instructions for your HIT. It is a good idea to:' followed by a bulleted list of three points: 'Be as specific as possible in your instructions to eliminate ambiguity. For example, when asking workers to extract texts from an image, ask workers to type the texts exactly as shown in the image including capitalization, spaces and punctuations.', 'Include an example of a right answer and a wrong answer. Highlight important information where appropriate.', and 'Clarify what you expect if a HIT is not doable because of missing data or other problems.'

Amazon Mechanical Turk HIT creation I

The screenshot shows the 'Create' tab of the Amazon Mechanical Turk interface. At the top, there are navigation links: Home, Create (active), Manage, Developer, and Help. Below the navigation, there are options: 'New Project' (underlined), 'New Batch with an Existing Project', and 'Create HITs individually'. The main content area is divided into three sections, each with a red border:

- 1. This is an example of a multiple choice question. Select only ONE of the following:**
 - Apple
 - Orange
- 2. This is another example of a multiple choice question. Select only ONE of the following. A drop down list saves more screen real estate than radio buttons but may require more keystrokes for the Worker to select the answer.**
 - select one -
- 3. This is an example of a question where you can select one or more of these options:**
 - Broccoli
 - Cabbage
 - Daikon Radish
 - Eggplant
 - Fennel
 - Ginger

On the right side of the page, there is a blue bar with the text 'It is a good idea to:' followed by a list of instructions: 'in your instructions to eliminate ambiguity. For example, extract texts from an image, ask workers to type the texts manage including capitalization, spaces and punctuations. right answer and a wrong answer. Highlight important information where appropriate.' Below this is a bullet point: 'Clarify what you expect if a HIT is not doable because of missing data or other problems.'

At the bottom left, there is a sidebar with three categories: 'Transcription from an image', 'Writing', and 'Other' (highlighted with an orange arrow). On the right side, there is a large orange button labeled 'Create Project »'.

Amazon Mechanical Turk HIT creation I

The screenshot shows the Amazon Mechanical Turk HIT creation interface. At the top, there is a navigation bar with tabs for Home, Create, Manage, Developer, and Help. Below the navigation bar, there are links for 'New Project' and 'New Batch with an Existing Project', and a link for 'Create HITs individually' on the right. The main content area is divided into several sections, each with a numbered instruction and a corresponding form example:

- 1. This is an example of a multiple choice question. Select only ONE of the following:**
 - Apple
 - Orange
- 2. This is another example of a multiple choice question. Select only ONE of the following. A drop down list saves more screen real estate than radio buttons but may require more keystrokes for the Worker to select the answer.**
 - select one -
- 3. This is an example of a question where you can select one or more of these options:**
 - Broccoli
 - Cabbage
 - Daikon Radish
 - Eggplant
 - Fennel
 - Ginger
- 4. This is another example of a question that can have multiple answers.**
 - Apple
 - Banana
 - Cherry
 - Durian
 - Elderberry
- 5. This is an example of an open-ended question. Use this when you only need Workers to enter a line of text.**
 - [Text input field]
- 6. This is another example of an open-ended question. Use this when you need Workers to enter multi-line texts.**
 - [Multi-line text input field]

At the bottom of the interface, there is a blue button that says 'You must ACCEPT the HIT before you can submit the results.'

Annotations in the image include:

- A red box around the first three examples (1, 2, and 3).
- A red box around the fourth example (4).
- A red box around the fifth and sixth examples (5 and 6).
- An orange box around the 'Other' button in the left sidebar.
- A blue box around the text 'information where appropriate' and the bullet point 'Clarify what you expect if problems.'

Amazon Mechanical Turk HIT creation II

1 Enter Properties

2 Design Layout

3 Preview and Finish

Project Name:

This name is not displayed to Workers.

Describe your HIT to Workers

Title

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

Give more detail about this task. This gives Workers a bit more information before they decide to view your HIT.

Keywords

Provide keywords that will help Workers search for your HITs.

This project may contain potentially explicit or offensive content, for example, nudity. [\(See details\)](#)

Amazon Mechanical Turk HIT creation III

Setting up your HIT

Reward per assignment

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

How many unique Workers do you want to work on each HIT?

Time allotted per assignment

Maximum time a Worker has to work on a single task. Be generous so that Workers are not rushed.

HIT expires in

Maximum time your HIT will be available to Workers on Mechanical Turk.

Auto-approve and pay Workers in

This is the amount of time you have to reject a Worker's assignment after they submit the assignment.

Amazon Mechanical Turk HIT creation IV

1 Enter Properties 2 Design Layout 3 Preview and Finish

Project Name: This name is not displayed to Workers.

Frame Height: 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** *I_x* [List Icons] [Undo] [Redo] Source

Instructions: Emotion and Association

Please look at the target word. then:

- 1. select the emotion direction that best describes your feeling about this word. For example, you might "strong positive" for LOVE, or "neutral" for "hairbrush"
- 2. please enter the first word that comes to your mind. For example, enter something like TREE or PEAR for APPLE. only enter one word.
- 3. if you encounter any problems with this hit or want to explain more, please use the optional comment field

Target: **$\${TARG}$**






use variables to feed
your own data to the
template

Amazon Mechanical Turk HIT creation IV

1 Enter Properties 2 Design Layout 3 Preview and Finish

Project Name: This name is not displayed to Workers.

Frame Height: Height in pixels of the frame your HIT will be displayed in

Format Font **U** **I** **B** **A** **I_x**     

Instructions: Emotion and Association

Please look at the target word. then:

- 1. select the emotion direction that best describes your feeling for the target word. For example, "neutral" for "hairbrush"
- 2. please enter the first word that comes to your mind. For example, "brush" for "hairbrush"
- 3. if you encounter any problems with this hit or want to exit the HIT, click on the "Report Problem" button.

Target: $\${TARG}$

What emotion do you have regarding $\${TARG}$?

strong positive

positive

neutral

negative

strong negative

What word came to your mind first when reading $\${TARG}$?

Do you have any comments? (optional)

use variables to feed your own data to the template

Amazon Mechanical Turk HIT creation V

- Download a sample csv file and fill it
- Upload to define a batch

Column name is variable name

	A
1	TARG
2	LOVE
3	HATE
4	GARBAGE
5	WALL
6	LAPTOP
7	TABLE
8	PEACH
9	KEY
10	HELICOPTER
11	GRASS
12	
13	

Publish Batch ×

Choose a .csv file with the variables you specified in your project. ([learn more](#))

File validation completed successfully

Validated

File Name: input.csv
File Validated: Yes
File Size: .07 KB
Line Count: 11

Don't have a data file? [Download a sample .csv file.](#)

Amazon Mechanical Turk HIT creation VI

HIT Preview

Instructions: Emotion and Association

Please look at the target word. then:

- 1. select the emotion direction that best describes your feeling about this word. For example, you might "strong positive" for LOVE, or "neutral" for "hairbrush"
- 2. please enter the first word that comes to your mind. For example, enter something like TREE or PEAR for APPLE. only enter one word.
- 3. if you encounter any problems with this hit or want to explain more, please use the optional comment field

Target: LOVE

What emotion do you have regarding LOVE?

- strong positive
- positive
- neutral
- negative
- strong negative

Deliberate redundancy –
always expect
maximally sloppy workers

What word came to your mind first when reading LOVE ?

Do you have any comments? (optional)

Amazon Mechanical Turk HIT creation VII

Batch Properties

Title:	Emotions and Associations
Description:	For a word, select its emotional value and provide associations
Batch expires in:	7 Days
Results are auto-approved and Workers are paid after:	1 Hours
Master Qualification:	Masters

HITs

Number of HITs in this batch:	10
Number of assignments per HIT:	x 10
Total number of assignments in this batch:	<u>100</u>

Ouch: 50% premium!

Cost

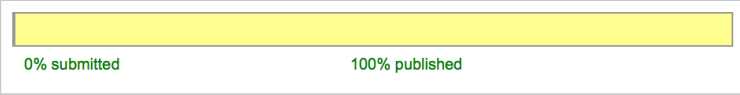
Reward per Assignment:	\$0.02
	x 100 <i>(total number of assignments in this batch)</i>
Estimated Total Reward:	<u>\$2.00</u>
Estimated Fees to Mechanical Turk:	+ \$1.00 <i>(fees paid to Mechanical Turk) (fee details)</i>
Estimated Total Cost:	<u>\$3.00</u> <i>(this is the amount that will be deducted from your Available Balance when you click "Publish HITs")</i>

Amazon Mechanical Turk HIT running!

For a word, select its emotional value and provide assoiations

Status Cancel this batch

Status: In Progress



0% submitted 100% published

Assignments Completed: 0 / 100 **Average Time per Assignment:** Not yet available **Average Hourly Rate:** Not Yet Available

Creation Time: October 26, 2015 10:40 AM PDT **Estimated Completion Time:** Not yet available

Settings

EmoAssoc1

Description: For a word, select its emotional value and provide assoiations

Keywords: word, emotion, association

Qualification Requirement: Masters has been granted

Number of Assignments per HIT: 10

Reward per Assignment: \$0.02

Input File: input.csv

HIT expires on: Not yet available

Assignment duration: 10 minutes

Auto Approval Delay: 1 hour

Results Results

Assignments pending review: 0

Assignments approved: 0

Assignments rejected: 0

Cost Summary

Estimated Total Reward: \$2.00

Estimated Fees to Mechanical Turk: \$1.00 ([fee details](#))

Estimated Total Cost: \$3.00

These costs are only an estimate until all of the assignments have been submitted and reviewed.

Amazon Mechanical Turk HIT still running!

For a word, select its emotional value and provide associations

Status
Cancel this batch

Status: In Progress

25% submitted
100% published

Assignments Completed: 25 / 100
Average Time per Assignment: 16 seconds
Average Hourly Rate: \$4.50

Creation Time: October 26, 2015 10:40 AM PDT
Estimated Completion Time: October 26, 2015 11:07 AM PDT (TODAY)

EmoAssoc1-1

Customize View Filter Results Upload CSV Download CSV

10 of 10 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Approve
Reject

	<u>HIT ID</u> ▲	<u>Worker ID</u>	<u>Lifetime Approval Rate</u>	<u>Input.Targ</u>	<u>Comment</u>	<u>Q1 Answer</u>	<u>Associations</u>
<input type="checkbox"/>	33Q5P9PUSPMH8731G8T1FAR1XATCZN	A18WKVB9OP0FAR	0% (0/0)	PEACH	{}	+1	juicy
<input type="checkbox"/>	33Q5P9PUSPMH8731G8T1FAR1XATCZN	A2WNW8A4MOR7T7	0% (0/0)	PEACH	{}	+1	yum
<input type="checkbox"/>	3A520CCNWN0V4XF707JXO50V91ZAED	A18WKVB9OP0FAR	0% (0/0)	GARBAGE	{}	-1	smelly
<input type="checkbox"/>	3MQY1YVHS3K8SR8JD1OSHFBCUGU2BY	A18WKVB9OP0FAR	0% (0/0)	TABLE	{}	0	dinner
<input type="checkbox"/>	3PN6H8C9R4QQCR174C8DINIA9PYADK	A18WKVB9OP0FAR	0% (0/0)	HATE	{}	-2	violence
<input type="checkbox"/>	3UV0D2KX1MJ88MZMMHTCTKZVTSHF4O	A2WNW8A4MOR7T7	0% (0/0)	KEY	{}	0	door
<input type="checkbox"/>	3VEI3XUCZRXR05QEBWLUD709NCORPM	A18WKVB9OP0FAR	0% (0/0)	LOVE	{}	+1	romance
<input type="checkbox"/>	3XUSYT70IT1UBE3PCABKTH6X03MD0W	A18WKVB9OP0FAR	0% (0/0)	HELICOPTER	{}	-1	rescue
<input type="checkbox"/>	3YGY136419RXKE1VH829JENXONRN	A18WKVB9OP0FAR	0% (0/0)	WALL	{}	0	room
<input type="checkbox"/>	3ZXNP4Z39RLRR1R1AQUHR9OPHQU7L6	A18WKVB9OP0FAR	0% (0/0)	GRASS	{}	+1	fields
	<u>HIT ID</u> ▲	<u>Worker ID</u>	<u>Lifetime Approval Rate</u>	<u>Input.Targ</u>	<u>Comment</u>	<u>Q1 Answer</u>	<u>Associations</u>

Approve
Reject

Amazon Mechanical Turk HIT results

- Check for scammers, reject their HITs (use with care!) and block the most blatant culprits from further HITs
- Download, and post-process

	RequesterArn	AssignmentId	AutoApproved	ExpirationTime	NumLifes	Assigner	Worker	AssignmentStatus	AcceptedTime	SubmittedTime	AutoApproved	ApprovalTime	Answer	AnswerScore	AnswerText
1	BatchId:2136	600	3600	Mon Nov 02 09:4	3A9AA9	A2VE5I	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	-1	FANTASY
2	BatchId:2136	600	3600	Mon Nov 02 09:4	3JZQSN	A18WK	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	1	romance
3	BatchId:2136	600	3600	Mon Nov 02 09:4	3RRCEF	A2HR7	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	happy
4	BatchId:2136	600	3600	Mon Nov 02 09:4	3STRJB	A1SRI5	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	happy
5	BatchId:2136	600	3600	Mon Nov 02 09:4	3SUWZ	A1S88V	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	Francis
6	BatchId:2136	600	3600	Mon Nov 02 09:4	3VSOLA	A2WNV	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	Francis
7	BatchId:2136	600	3600	Mon Nov 02 09:4	3YDTZA	A3KOV	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	wife
8	BatchId:2136	600	3600	Mon Nov 02 09:4	3YGXW	ADAGU	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	wife
9	BatchId:2136	600	3600	Mon Nov 02 09:4	34S9DK	A3KOV	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	wife
0	BatchId:2136	600	3600	Mon Nov 02 09:4	3A7Y0F	A2VE5I	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	wife
1	BatchId:2136	600	3600	Mon Nov 02 09:4	3C5W7	ADAGU	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	affection
2	BatchId:2136	600	3600	Mon Nov 02 09:4	3C5W7	ADAGU	Approved	Mon C	Mon C	Mon C	Mon C	Mon C	LOVE	2	children
													HATE	-2	brutal
													HATE	-1	HOT
													HATE	-2	anger

Crowdfunder HIT creation I

What would you like to do?



Sentiment Analysis



Search Relevance



Data Categorization



Data Collection & Enrichment



Data Validation



Image Annotation



Transcription



Content Moderation



Start from Scratch

Crowdfunder HIT creation II

Step 1: Upload your data

Delete

Add More Data

Split column

Convert Uploaded Test Questions

Download

<input type="checkbox"/>	UNIT ID	STATE	JUDGMENTS	AGREEMENT	TARG
<input type="checkbox"/>	815612655	new	0		LOVE
<input type="checkbox"/>	815612656	new	0		HATE
<input type="checkbox"/>	815612657	new	0		GARBAGE
<input type="checkbox"/>	815612658	new	0		WALL
<input type="checkbox"/>	815612659	new	0		LAPTOP
<input type="checkbox"/>	815612660	new	0		TABLE
<input type="checkbox"/>	815612661	new	0		PEACH
<input type="checkbox"/>	815612662	new	0		KEY
<input type="checkbox"/>	815612663	new	0		HELICOPTER
<input type="checkbox"/>	815612664	new	0		

File upload successfully processed
10 units were added.

Crowdfunder HIT creation III

Step 2: Design your job

Title

Association And Emotion

Content

DATA | {{targ}}

LOVE

FORM ELEMENT | multiple choice

What emotion do you have regarding [[targ]]??

- strong negative
- negative
- neutral
- positive
- strong positive

Cancel

Update element



Multiple Choice

General Options

Question Text

What emotion do you have regarding [[targ]]??

Options

LABEL	VALUE	
strong negative	-2	
negative	-1	
neutral	0	
positive	+1	
strong positive	+2	

Add Choice

A response is required

Yes No

Additional options ▾

► Conditional Logic

Crowdfunder HIT creation IV

Step 3: Create test questions

Questions

Test Questions

Reports ▾

Test Questions (aka, Gold) are rows with known answers that are regularly inserted throughout your job. They drastically improve results by:

1. training contributors -- contributors see why they got the test question wrong if they fail it;
2. removing underperformers -- we remove the contributor and all of their answers from the job when they fail too many test questions.

Create Test Questions

Crowdflower HIT creation V

DATA
DESIGN
QUALITY
LAUNCH
MONITOR

Settings

Basic Contributors Quality Control Sharing / Visibility API

Internal Contributors

Off **Share your job with team members**
 If you want your team members to work on this job, we will provide you with a link to share with your teammates

External Contributors

Share your job with CrowdFlower Contributors
 When enabled, we'll allow our online contributors to work on your job

Job may be offered to Unleveled contributors

No Job contains explicit content

No NDA Channels Only ([more info](#))

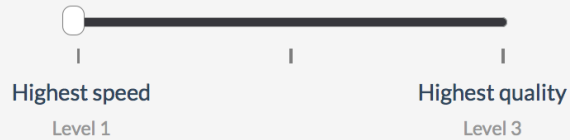
Hide Contributor Channel Table

CONTRIBUTOR CHANNEL NAME	ENABLED	ACTIVITY ?	TOP COUNTRIES	TRUSTED / UNTRUSTED ?	ADULT CONTENT?
BitcoinGet	<input checked="" type="checkbox"/>	<div style="width: 100%; height: 5px; background-color: green;"></div>	USA, GBR, DEU	<div style="width: 100%; height: 15px; background-color: green; border: 1px solid red;"></div>	Yes
Bituro	<input checked="" type="checkbox"/>				Yes
Boizu	<input checked="" type="checkbox"/>				No
Borders (Philippines)	<input checked="" type="checkbox"/>				No
? ClixSense	<input checked="" type="checkbox"/>	<div style="width: 100%; height: 5px; background-color: green;"></div>	USA, PHL, IND	<div style="width: 100%; height: 15px; background-color: green; border: 1px solid red;"></div>	Yes
Coin Tasker	<input checked="" type="checkbox"/>	<div style="width: 10%; height: 5px; background-color: green;"></div>	USA	<div style="width: 100%; height: 15px; background-color: green; border: 1px solid red;"></div>	Yes
? CoinWorker.com	<input checked="" type="checkbox"/>	<div style="width: 100%; height: 5px; background-color: green;"></div>	USA, CAN, TWN	<div style="width: 100%; height: 15px; background-color: green; border: 1px solid red;"></div>	No
Content Runner	<input checked="" type="checkbox"/>				No
Crowd Guru	<input checked="" type="checkbox"/>	<div style="width: 100%; height: 5px; background-color: green;"></div>	DEU, AUT, ITA	<div style="width: 100%; height: 15px; background-color: green; border: 1px solid red;"></div>	Yes
CrowdFlower Elite	<input checked="" type="checkbox"/>				Yes
Crowdsdom	<input checked="" type="checkbox"/>				No
DDD (Cambodia)	<input checked="" type="checkbox"/>				No

Crowdfunder HIT creation 6

Performance formerly levels

Select the ratio of speed to quality that best reflects your needs – when only our highest-rated contributors are given access to your job, it takes longer to complete



Judgment Details

Set the number of contributors who will answer a given question

Judgments per Row

Set the number of rows displayed on each page

Rows per Page

Save

Payment Details

Set the payment amount per page
See our [Job Costs FAQ](#) item for more details

¢ per Page

100 rows for your new job.

Launch Your Team's Job

Set your quantity and you are ready to launch your job

Rows to launch

of **10** total rows (100.00%)

+ 1 Test Questions

Set the payment amount per page

¢ per Page ⓘ

This is what we expect your job to cost: ⓘ

Crowd costs \$2.55

Transaction Fee (20%) \$0.51

Cost buffer ⓘ \$6.94

Crowdfunder HIT running !

Monitor

[Dashboard](#) [Alerts](#) [Advanced Analytics](#) [Contributors](#)

STATUS
Running Pause Job

JOB ID 791545
Association And Emotion

- Settings
- Preview
- Copy

READ THE DOCS

- [Dashboard Guide](#)
- [Troubleshooting Guide](#)
- [Auto-paused Jobs](#)

0% Complete	\$0 Cost	1 Active Test Questions	10 Rows
0 Judgments By Hour	0 Trusted Judgments	0 Untrusted Judgments	0 Pending Judgments

Judgments By Hour

Estimated Completion: **Unknown**

Crowdfunder HIT running !

Monitor

STATUS **Running** [Pause Job](#)

JOB ID 791545
Association And Emotion

- Settings
- Preview
- Copy

READ THE DOCS

- [Dashboard Guide](#)
- [Troubleshooting Guide](#)

[Dashboard](#) [Alerts](#) [Advanced Analytics](#) [Contributors](#)

0%
Complete

\$0
Cost

1
Active Test Questions

10
Rows

0
Judgments By Hour

0
Trusted Judgments

0
Untrusted Judgments

0
Pending Judgments

Judgments By Hour

[Dashboard](#) [Alerts](#) [Advanced Analytics](#) [Contributors](#)

We found 1 possible issue with your job.

TEST QUESTIONS

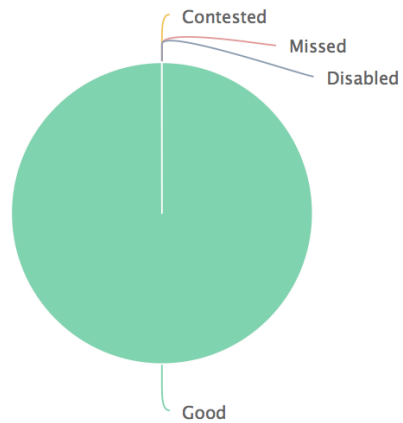
ALERT	FIX
You have created some test questions for your job. However, since you have not completed the recommended number of test questions, your job may run slowly, you may not get back high-quality data, or both.	Please create the recommended number of test questions. Fix it now

Crowdfunder HIT running II

Contributor Funnel i

No contributors found.

Test Questions i



Total: 1

[View Missed Test Questions](#)

Contributor Satisfaction i




Crowdfunder HIT running III

Monitor

[Dashboard](#) [Alerts](#) [Advanced Analytics](#)


65%

Complete 

\$2

Cost 

3

Judgments By Hour 

55

Trusted Judgments 

Full Report [← Back to Results](#)

[Filter](#) [Options](#) [Reset](#) [Pivot](#) [Download](#)

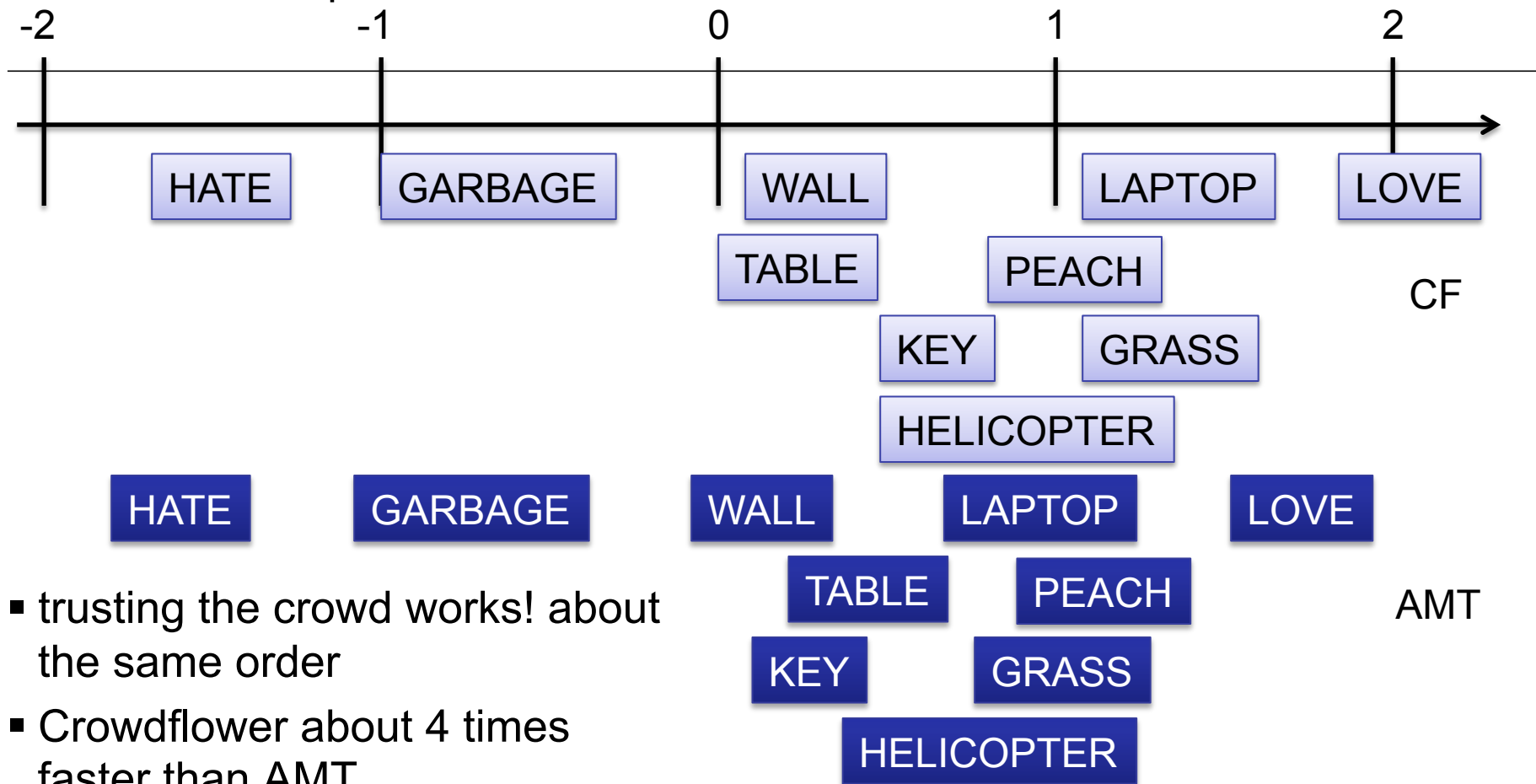
id	assoc	comments	what_emoti...	targ
17827565...	brick		0	WALL
17827566...	concrete		0	WALL
17827566...	Finish		-1	WALL
17827560...	Internet		+2	LAPTOP
17827568...	movie		+2	LAPTOP
17827569...	computer		0	LAPTOP
17827575...	acer	i have acer	+1	LAPTOP
17827579...	needed		+1	LAPTOP
17827581...	entertainm...		+2	LAPTOP
17827556...	wood		0	TABLE
17827557...	nothing		0	TABLE
17827560...	eat		0	TABLE
17827565...	neutral		0	TABLE

Crowdfunder HIT results

- scammers and untrusted judgments are removed automatically (via test questions)
- can choose between aggregated or full data; aggregation weighted by worker trust
- revisit your test questions, convert HTI results to future test questions
- download and post-process

id	assoc	comment	what_emc	targ
1782755650	love		2	LOVE
1782755765	passion		1	LOVE
1782756083	happy		2	LOVE
1782756550	positive		2	LOVE
1782756566	kiss		2	LOVE
1782756620	my fiance		2	LOVE
1782756667	wife		2	LOVE
1782797837	family		2	LOVE
1782805561	girl		2	LOVE
1782817990	novia		1	LOVE

Results on \$5 Emotions and Associations



- trusting the crowd works! about the same order
- Crowdfunder about 4 times faster than AMT
- Crowdfunder a little cheaper because of 1c – minimum rule

Some associations

Crowdflower:

GRASS – green (3), **happy** (2), nature (2), patio, park, **Scent**

PEACH – fruit (3), color, taste, **neutral**, **white flag**, eat, **rico**, "**fruit, sweet**"

TABLE – wood (3), **neutral**, **nothing**, eat, food, **coffe**, chair, **comer**

AMT:

GRASS – clouds, **fresh cut**, life, mow, **SUNNY**, fields, grow, outside

PEACH – **Georgia** (2), yum, sweet, fruit, **JUICE**, juicy, food

TABLE – food, dinner, eat, dinner, **STURDY**, chair, sit, furniture

- errors occur in both platforms, but different ones!
- use this to improve guidelines!

Some comments (all on AMT except last one)

HELICOPTER (+1) – rescue : I always think of helicopters as rescuing people and flying them to hospitals or saving them from places rescue people can't get to, like mountains or lakes, so it's a positive word.

LOVE (+2) – Francis : Pope Francis really made an impression on me watching his visit on tv, and how he inspired love and how much that love seemed to make the world a little better place for awhile.

GARBAGE (-2) – cans : I have to take out the garbage cans every week, and they are heavy, and in the winter it;s hard, which is why I put strong negative.

LAPTOP (+1) – acer : I have acer

Lessons Learned from This Experiment

- Specify Guidelines!
 - non-English “comer”
 - capitalization
 - verbs vs. nouns vs. proper names
 - add redundancy: “put SINGLE WORD in lowercase”
- People mix up the two tasks “neutral”
- Small batch: almost no scammers despite lack of quality control

Quality Control – a few words

- This will be covered in depth *later* – here only mechanisms for two platforms
- Crowdfunder:
 - test items, automated worker quality assessment
 - can select channels, can access language skills
- Amazon Mechanical Turk
 - standard levels: masters, photo categorization masters, ..
 - qualifications: platform-wide stats such as approval rate, number of approved hits
 - custom qualifications

MTurk Qualification creation I

Home Create **Manage** Developer Help

Results Workers Qualification Types

Manage Qualification Types

Below is a list of your Qualification Types and the corresponding number of Workers.

Create New Qualification Type ▶

You have not created any Qualification Types. Please use the button above to create a Qualification Type.

Create New Qualification Type



This will create a new Qualification Type that can be assigned to Workers.

Friendly Name:

Description:

The Qualification Type you are creating does not require Workers to take a Qualification test. You can create a Qualification Type that requires Workers to take a test using the Developer Tools. [Get more info.](#)

or [Cancel](#)

MTurk Qualification creation II

[Home](#) [Create](#) [Manage](#) [Developer](#) [Help](#)

[Results](#) [Workers](#) [Qualification Types](#)

Manage Qualification Types

Below is a list of your Qualification Types and the corresponding number of Workers.

[Create New Qualification Type](#)

Qualification Types				
Name ▼	ID	Workers who have this Qualification	Creation Date	Description
⊗ Myqual	3ZA92V7SPMSVP47W3MGLFXA84NVC9M	0	2015-10-28 07:27:02 -0700	use this qualification to work on my German HITs

MTurk Qualification assignment

Home Create **Manage** Developer Help

Results Workers Qualification Types [We're Hiring! Learn More](#)

[Manage Workers](#) > [Manage Individual Worker](#)

Worker Details

Worker ID: A1ETW1FPX2RS71

Worker Status for your work	Never Blocked	WORKER ACTIONS: Bonus Worker Block Worker
APPROVAL RATE ON YOUR ASSIGNMENTS		
Lifetime	100% (40/40)	
Last 30 days	0% (0/0)	
Last 7 days	0% (0/0)	

[Assign Qualification Type](#)

Worker's Qualifications For Your Work

MTurk Qualification assignment

The image shows a screenshot of the MTurk interface with a modal dialog box titled "Assign Qualification Type" overlaid. The dialog box has a close button (X) in the top right corner. The main text inside the dialog reads: "Select which of your Qualification Types to assign to this Worker. You will be prompted to provide a score between 0 and 100. (Assign up to 5 Qualification Types at a time.)". Below this text, there is a checked checkbox labeled "Myqual". Underneath the checkbox, the word "Score:" is followed by a text input field containing the number "10". At the bottom right of the dialog, there is a yellow "Assign" button followed by the text "or Cancel".

Home

Results **Wor**

Manage Workers

Worker De

Worker ID: A1E

Worker Stat

APPROVAL RA

Lifetime ...

Last 30 da

Last 7 day

Worker's C

Assign More

NS:

Qualification Type

Assign Qualification Type

Select which of your Qualification Types to assign to this Worker. You will be prompted to provide a score between 0 and 100. (Assign up to 5 Qualification Types at a time.)

Myqual

Score:

10

Assign or Cancel

Advanced options for restricting workforce

Advanced Worker requirements «

Worker requirements:

Customize Worker Requirements...

Specify ALL the qualifications Workers must meet to work on your HITs:

Masters [remove](#)

HIT Approval Rate (%) for all greater than or equal to 95 [remove](#)

Number of HITs Approved greater than or equal to 1000 [remove](#)

Myqual greater than 3 [remove](#)

[\(+ Add another criterion](#) (up to 5)

Only Workers who qualify to do my HITs can preview my HITs.

Yes No

Using AMT from your own programs

Choosing the Right Tool

In addition to the Mechanical Turk Requester website, we offer a robust [Application Program Interface \(API\)](#) and [command line tools \(CLTs\)](#) to enable you to programmatically distribute tasks that require human intelligence to a widely distributed, on-demand workforce.



The web interface lets you quickly create and manage thousands of Human Intelligence Tasks (HITs). The Requester website lets you build templates that you can reuse for future projects. You can also manage all of your results, including accepting and rejecting work through the web.



The API provides full flexibility and direct integration into your application. We offer multiple [SDKs](#) to make it easier for you to build solutions leveraging Mechanical Turk.



The Command Line Tools require more experience with technology but give you more features and control than the Requester website.

- Crowdfunder also has a RESTful API

Using AMT from your own programs

Choosing the Right Tool

In addition to the AMT CLI, you can also use the AMT API to create and manage HITs from your own programs. The AMT API is widely distributed and can be used in a variety of ways.



	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	✓	✓	
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			✓

- Crowdflower also has a RESTful API

\$40 project: Learning to Paraphrase in Context

Goal: A semantic writing aid that:

- provides **paraphrase detection** and **prediction** capability
- Improves **readability** of documents
- Provides **varied writing** styles by suggesting semantically equivalent paraphrases
- Removes **redundant** or **overused** words or phrases.

The semantic writing aid tool is envisioned to improve readability of documents and provide varied writing styles by suggesting semantically equivalent paraphrases and removing redundant words or phrases.

Paraphrases

transparency

legibility

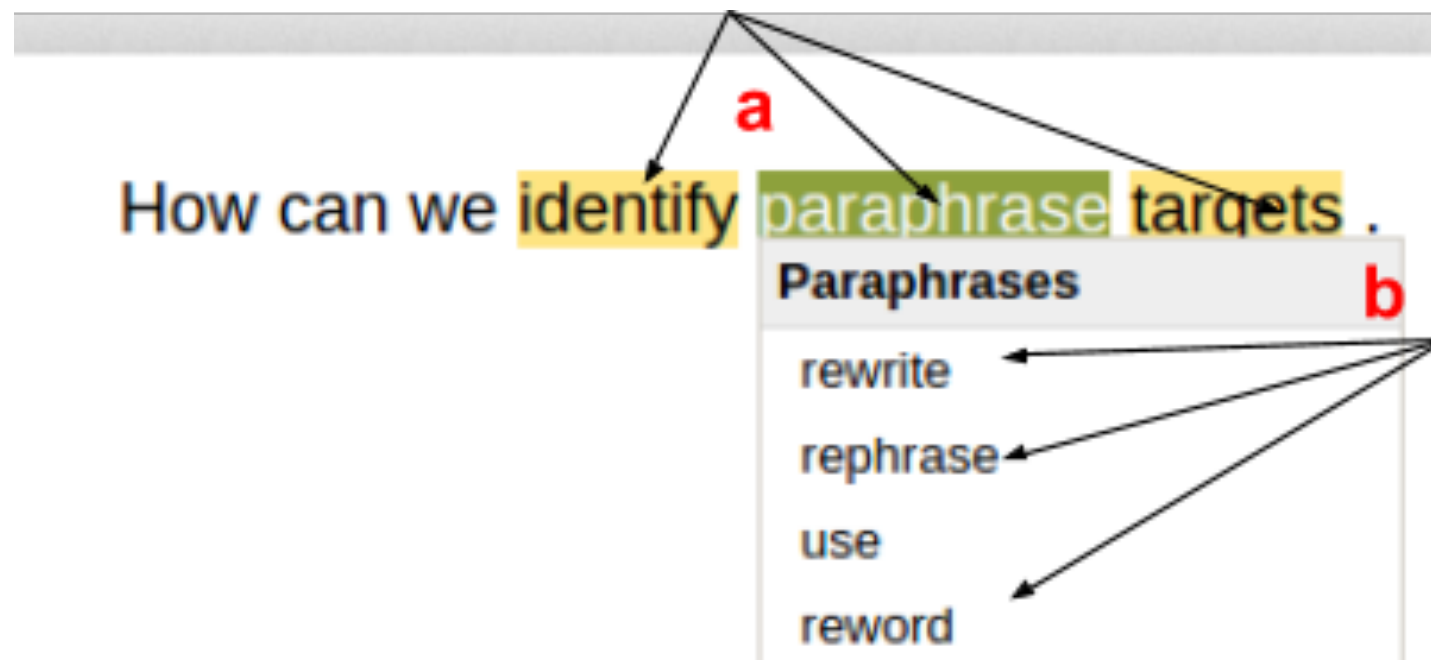
literacy

visibility

clarity

Data preparation for crowdsourcing

- Paraphrase target detection



Paraphrase targets (**a**) and Paraphrase candidates (**b**)

Paraphrase target detection

- Use our baseline system to detect targets
- Unigram tokens
 - PPDB
 - WordNet
 - JoBimText distributional thesaurus
- Multiword expression
 - WordNet
 - STREUSLE
 - Wiki50
 - Multiword expression project
 - A total of 79349 MWEs

Using your own GUI on AMT

Using an externally hosted page in a HIT

- **Kathleen's** blog
- All URLs should be **https**
- **Boto** is a Python package that provides interfaces to Amazon Web Services for embedding external pages in AMT

<http://kaflurbaleen.blogspot.de/2012/05/in-which-i-battle-mturk-external-hits.html>
<http://kaflurbaleen.blogspot.de/2014/06/in-which-i-battle-mturk-external-hits.html>
<https://github.com/boto/boto>

Boto Python Code Example

```
import boto.mturk.connection
import boto.mturk.question
sandbox_host = 'mechanicalturk.sandbox.amazonaws.com'
real_host = 'mechanicalturk.amazonaws.com'
mturk = boto.mturk.connection.MTurkConnection(
    aws_access_key_id = 'XXXXX',
    aws_secret_access_key = 'XXXXX',
    host = sandbox_host,
    debug = 1 # debug = 2 prints out all requests. but we'll just keep it at 1
)
print boto.Version
print mturk.get_account_balance()
URL = "https://maggie.lt.informatik.tu-darmstadt.de/semsch/lt.html"
title = "Semantic Writing Aid"
description = "Replace the highlighted words with term for similar meaning"
keywords = ["essay", "paraphrase", "survey", "English", "writing"]
frame_height = 500 # the height of the iframe holding the external hit
amount = .05
questionform = boto.mturk.question.ExternalQuestion( URL, frame_height )
create_hit_result = mturk.create_hit(
    title = title,
    description = description,
    keywords = keywords,
    question = questionform,
    reward = boto.mturk.price.Price( amount = amount),
    response_groups = ( 'Minimal', 'HITDetail' ), # I don't know what response groups are
)
HIT = create_hit_result[0]
assert create_hit_result.status
print '[create_hit( %s, $%s )]: %s]' % ( URL, amount, HIT.HITId )
```

Paraphrasing Tool embedded in AMT

The screenshot displays the Amazon Mechanical Turk (AMT) interface. At the top, the logo for Amazon Mechanical Turk is visible, along with navigation links for 'Your Account', 'HITS', and 'Qualifications'. A notification indicates '204,837 HITS available now'. Below this, there are filters for 'All HITS', 'HITS Available To You', and 'HITS Assigned To You'. A search bar allows finding HITS containing a specific text, with a minimum payment of \$0.00. A timer shows '00:00:33 of 7 days'. The main task area includes buttons for 'Submit HIT' and 'Return HIT', and a checkbox for 'Automatically accept the next HIT'. The task details show 'Semantic Writing Aid' by requester 'Seid', with a reward of '\$0.05 per HIT' and no qualifications required. The task content is a paragraph of text with several words highlighted in yellow. A dropdown menu is open over the word 'spend', showing a list of paraphrases: 'pay', 'use', 'dedicate', 'devote', 'move', 'Ignore this error', and 'Rule implementation...'. At the bottom of the task area, there are buttons for 'Submit HIT' and 'Reload', with a tooltip that says 'Finish and Submit HIT'.

amazonmechanical turk
Artificial Artificial Intelligence

Your Account HITS Qualifications 204,837 HITS available now

All HITS | HITS Available To You | HITS Assigned To You

Find HITS containing [] that pay at least \$ 0.00 for which you are qualified require Master Qualification GO

Timer: 00:00:33 of 7 days

Finished with this HIT? Let someone else do it?

Submit HIT Return HIT

Automatically accept the next HIT

Semantic Writing Aid
Requester: Seid
Qualifications Required: None
Reward: \$0.05 per HIT

I was already afraid you wouldn't write to me at all any more. I will in any case urge that each year we spend a whole month together, so that you see that your father who is fond of you and who loves you, has achieved through such a lot of strenuous work, to be there for strangers but especially for my own. I am very pleased that you find joy with the piano. My carpentry are in my opinion for your age the best.

Paraphrases

- pay
- use
- dedicate
- devote
- move
- Ignore this error
- Rule implementation...

Submit HIT Reload

Finish and Submit HIT

But, really, how do you crowdsource paraphrasing?

- Directly embedding paraphrase tool: no control over
 - how many phrases are paraphrased
 - whether people edit the text freely
 - quality assurance
- Remember lessons learned:
 - keep it simple
 - avoid free text whenever possible
 - make sure to have a handle on test items

AMT interface for paraphrasing in context

- 5 targets per hit, 5 candidates per target
- Extra field for additional input
- JQuery to control user inputs: enforce 'at least one'

Select alternatives for the highlighted terms/phrases.

(minimum 1 and maximum 3 alternatives possible!!!)

The alternatives have to be **natural** in the context. You can choose **as many as three** for each term. Different forms (singular or plural nouns, present or past form of verbs) are accepted. If none apply, provide an alternative in the "Other" textbox.

Code for the patrolling behavior might **look like** this (note that part is on a schema basis) : For the task-switching in the patrolling behavior , assume some variable " age " that Antie has been keeping track of using the CPU 's clock pulses to track the time since its last bootup . The integers nestpatrol_age and trailpatrol_age will represent the ages at which it is appropriate to switch to these **tasks** . The final behavior , foraging , will involve one extra **subroutine** that will employ an **algorithm** designed by R. Andrew Russell to localize an **object** through sand based on a chemical plume .

look like	tasks	subroutine	algorithm	object
<input type="checkbox"/> re like <input type="checkbox"/> the air <input type="checkbox"/> seem like <input type="checkbox"/> sound like <input type="checkbox"/> s like Other: <input type="text" value="Type here"/>	<input type="checkbox"/> capability <input type="checkbox"/> assignment <input type="checkbox"/> subject <input type="checkbox"/> challenge <input type="checkbox"/> job Other: <input type="text" value="Type here"/>	<input type="checkbox"/> subprogram <input type="checkbox"/> quorum <input type="checkbox"/> function <input type="checkbox"/> user <input type="checkbox"/> recursive Other: <input type="text" value="Type here"/>	<input type="checkbox"/> identifier <input type="checkbox"/> method <input type="checkbox"/> technique <input type="checkbox"/> calculus <input type="checkbox"/> tool Other: <input type="text" value="Type here"/>	<input type="checkbox"/> objeto <input type="checkbox"/> subjectmatter <input type="checkbox"/> isomorphism <input type="checkbox"/> article <input type="checkbox"/> entity Other: <input type="text" value="Type here"/>

Experiment and Analysis

- 5% control hits (unrelated paraphrase candidates)
- Experiment conducted in smaller chunks
- Price was adjusted in every iteration (the higher the price the fastest)
- More than 4000 paraphrase instances collected with an observed agreement of 80%

Rounds	Number of HITs	Time/HIT	Total Time (h:m:s)	Price/ HIT	Total Price(\$)
1	15	2 m 8 sec	1:13:00	0.10c	8.40
2	20	2 m 37 sec	6:21:00	0.05	5.10
3	30	3 m 11 sec	14:35:00	0.05	7.20
4	39	3 m 35 sec	19:14:00	0.05	8.60
5	39	2 m	30:40:00	0.03	7.80

Experiment and Analysis

- Lessons learned
 - If the AMT template is used, the **name** attribute is required to get answers
 - By default, tasks created with AMT template sets qualifications to **Masters** only
 - **Location** of workers matters
 - Workers should not be rejected if the HIT submitted has an error (example attribute name missed – lost \$2)
 - The smaller the task, the easier to manage – can automatize.
 - Approve/reject
 - Group by Worker ID
 - Check for control questions
 - If we wait for the whole task to be completed, we get more spammers
 - Keywords matter for the task's visibility

Post-Hoc Machine Learning Evaluation: directly check the output of your system

- Crowdsourcing task for paraphrase scoring (without context)
 - Use the British Academic Writers Corpus (BAWC) for this task
- Compute correlation with PPDB2 results
- Repeat experiment (with context)

Score phrases / words based on their similarity in meaning.

For every word or phrases, you have given five options below it where your task is to judge how much each one of the options are **similar in meaning** to the given word or phrases. If they are **perfect match**, give a score of five (5). If they are **some how similar** but not exactly the same, give a score of three (3). If they are not similar at all, give a score of zero (0). You can have more than one options having the same score.

For example: highlight->address=5, highlight-> emphasise =5, highlight-> identify=4, highlight -> raise =2, highlight ->submit =0

instance

occasion

exception

allegedly

case

mention

conjunction

alliance

link

connection

relation

combination

highlights

news

explore

talk

describe

clip

parallelism

comparison

distinction

randomness

storage

analysis

repeated

constant

use

perform

frequent

persistent

Considerations for the requester



- Mostly US workers
 - More flexibility with designing HITs, e.g. with Javascript
 - Quality control largely up to the requester (although not entirely)
 - Qualification tests
 - Custom qualifications
 - Master Workers
 - DYI embedded controls/ majority voting scheme
- Better international coverage
 - HIT designs are mostly plug-and-play for requesters
 - Built in mechanisms for quality control = more quality, less control
 - Proprietary EM-style method for determine quality
 - Prices can be hard to predict
 - Limited to tasks that can be QCed this way

(slide by Chris Callison-Burch)

In a Nutshell: Learned in Lesson 2

- Two crowdsourcing platforms: Amazon Mechanical Turk and Crowdfunder
- Step-by-step instructions for conducting a simple task
- First steps with “Emotions and Associations”
- Advanced Options for restricting the workforce
- Using custom interfaces in AMT