Artificial artificial intelligence over the web: The Mechanical Turk API

Emiel de Munck
Student Media Technology,
Leiden University
e.de.munck@umail.leidenuniv.nl

Haixiang Zhang
Student Media Technology,
Leiden University
h.zhang.8@umail.leidenuniv.nl

Lianru Zhang
Student Media Technology,
Leiden University
l.zhang.7@umail.leidenuniv.nl

ABSTRACT
This paper explores Mechanical Turk utility on the web, mainly focused around Amazon's Mechanical Turk API. It includes a short history, operating principles, discussion on strengths and weaknesses, the intended and unintended applications that came forth of this technique and an explanation on how to get started using this API. It draws the conclusion that Mechanical Turk API's will be used as long as there are computations in which humans outperform computers.

1.1 Purpose
A Mechanical Turk API (application programming interface) is an interface for software to use human computation instead of other software's computation. Since humans still remain better at various tasks in comparison to computers, such an interface is used to integrate human resources as a relatively easy solution for general purpose computation [1].

1.2 Context and History
* 1770 The origin of Mechanical Turk technology goes back as far as this year, when Wolfgang von Kempelen constructed a chess playing machine to impress the empress of Austria [2]. Inside was a hidden human chess player that operated the machine, making it look like the machine itself could very impressively compute chess movements (Fig. 1).

* 2001 iStockPhoto was the first commercial web-based marketplace for crowdsourced work [3], however this is merely an interface where people can buy or sell photos (possibly on request).

* 2005 A more general-purpose platform was started in by Amazon: Amazon's Mechanical Turk [4]. Here, humans can ask for the help of other humans in order to solve complex computational problems (such as determining the number of dogs in a large series of pictures) that would not be as accurate, affordable or possible to solve with computers.
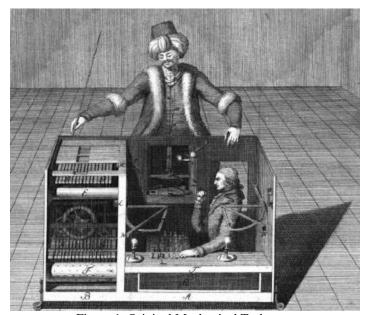

Figure 1, Original Mechanical Turk

2. OPERATING PRINCIPLES

2.1 Concepts and terminology
In this section, different names and terms commonly used in Amazon Mechanical Turk(AMT) will be introduced briefly along with a technical description.

2.1.1 Requesters
The person who posts requests or questions on Amazon Mechanical Turk is called requester. Usually, requesters can use some applications that have been already made by some programmers to submit their questions.But in essence the request is an XML file that is parsed to the AMT website. The statusof their missions is available on Requester Console[6].

2.1.2 Human Intelligence Tasks (HITs)
Any questions shown in the AMT applications are Human

Intelligence Tasks. A HIT should contain all the information of the request such as data regarding to a question, what kind of answer will be considered valid, payment details and number of workers.

Technical operation: within the API the CreateHit operation is called to create such a new task. CreateHit can be given several parameters such as *Question* that fully specifies the contents of the HIT or *HitLayoutId* that contains information about the layout.

2.1.3 Workers

Single person or a group of people trying to finish tasks and answering questions on AMT, who can be considered as the 'artificial intelligence' sitting behind the Mechanical Turk. All workers can find requests on Amazon Turk website. [2]

Sometimes, a worker may have to pass a small test identified by the requester to get qualification to that certain request.

Technical operation: The *AssignQualification* operation is used to build up a qualification. This can be assigned to a worker with *GrantQualification*. The qualification will be of a type that can be specified by the requester with *CreateQualificationType*, because different tasks might require specific qualifications.

2.1.4 Assignments

After a worker accepts a HIT, the AMT will create an assignment to track the process and stall the answer provided by the worker. Assignments are very useful to manage subtasks and sub-questions assigned to workers. When one subtask of a HIT is taken by a worker, the specific assignment will be closed to make sure only one worker is dealing with that question. The assignment will not open again unless the subtask is considered failed. For example, the worker fails to finish it within the time limit. With this kind of task managing solution, there is a significant increase of efficiency on solving HITs.

Technical operation: The *GetAssignmentsForHIT* operation will retrieve completed assignments for a HIT. It can be given the *AssignmentStatus* parameter to determine which set of assignments are returned for a HIT.

2.1.5 Approval and Payment

The requester will get assignments with answer data when all of the questions in the HITs was answered or the expiration date has passed. The requester has the right to approve or reject the answers provided by the Mechanical Turk. If the answers are approved, the payment will be transferred to the specific worker(s) automatically via Amazonaccounts.

Technical operations: the *ApproveAssignment* operation approves the result of an assignment that is completed. It automatically initiates two payments from the requesters Amazon account: a payment to the worker who completed

the task and a fee for Amazon itself.

2.1.6 Qualifications and Quality Control

AMT provides the requesters a qualification method as a quality control tool. The qualification consists of a name and a number value that will be assigned to workers, with which the requesters can filter some workers who are not qualified enough on certain tasks. A qualification test can be included in a qualification type. The test answers by workers can be checked manually by requesters or according to a preset answer key.

To workers, it is also possible to search HITs in some specific areas by Qualification types.

Detailed information about Qualifications can be found at [6]

Technical operation: UpdateQualificationScore and UpdateQualificationType can be used to dynamically keep in control of what qualification score or what type of qualification is needed for a certain HIT respectively.

2.2 Amazon Mechanical Turk Workflow

In this section, a diagram shows how the Amazon Mechanical Turk works with the concepts described in the last section.
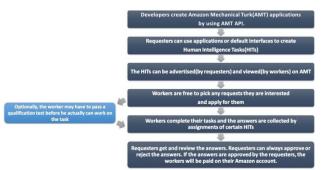

Figure2. The Amazon mechanical Turk workflow

3. STRENGTHS AND WEAKNESSES

3.1 Strengths

Mechanical Turk is very suitable,convenient and efficient for the requesters to outsourceburstytasks which have to be done in limited time with numerous workers [8]. In other words, the requesters can build up one application which includes a simple description of the task and then get results from the workers in several days. Mechanical Turk intends to complete uncomplicated tasks (for humans). At the same time, the requesters do not need to take responsibilities for the Mechanical Turk workers as they would for regular employees. Mechanical Turk Web, therefore, is a good option for individuals or companies, when they temporarily need a large number of people to assist them on certain projects.

So money and time can be the strengths of this approach. Especially on the short run, often making use of human computation will be a lot cheaper compared to investing the time and money to develop a software-based algorithm for the same tasks. In addition, people can increase their income with relative ease by serving as workers on Mechanical Turk. Regardless of their geographical position, workers can complete their tasks even with limited equipment and skills.

In short, the strengths of Mechanical Turk API's are "easy access to a large, stable, and diverse subject pool, the low cost of doing experiments, and faster iteration between developing theory and executing experiments." [1]

3.2 Weaknesses
On the other hand, Mechanical Turk is not as suitable for long-term tasks. The main reason is the availability of workers. Workers usually regard this job as part-time or they use it to spend their spare time. Therefore, they do not feel as much responsibility for these jobs as for their primary job.

Other weaknesses of using human resources are the inconsistency of results (both within and amongst individual workers), as well as the costs in terms of money and time. For every task completion human participants will need to invest time, and thus compensation in the form of money is often required.
However, the income from this Mechanical Turk job is relatively low; the reservation wage of workers is only $1.38 per hour.[9]. When this is compared to the US average minimum wage, the federal(non- Mechanical Turk) employers should pay $2.13 per person with tips[22] or $7.25 per person without tips[23] to their employees.The Workers appear to be mainly from the UnitedStates (46.8%) [24].
In addition, if the tasks are too complicated, the workers are less likely to accept this task.

3.4 Conclusion

The use of Mechanical Turk is restricted to simple, short-term tasks.However, the task completion might be relatively costly or time consuming if they can also be done by software computation. In short one could say that when no efficient software solution is available, a Mechanical Turk API can become a viable option.

4. INTENDED APPLICATIONS

Mechanical Turk usually provides professional business services for their requesters, with robust platforms.

Take AM Mechanical Turk for example. It offers multiple application sub-websites and resources to support users creating their own unique projects or work on it. It provides the five main business items: Data Cleaning, Categorization, Business feedback, Create & Moderate Content and Translation. These categories can therefore be seen as the main intended forms of applications for this platform.

For instance, if one company wants to improve their own websites through investigations regarding design and usability from the average end-user, they can select DomainPolish in AM Mechanical Turk website to do this work. The company can view their results through the web dashboard of DomainPolish. It provides a low-cost but high- quality feedback from the countless reviewers to the company in a short period.

4.1 Data Cleaning
It helps requesters clean their data quickly and efficiently via the three methods below.[10]
●Data Verification, Entry, and Collection
● Data Cleansing, Normalization, and De-duplication
● Algorithm Training
Relevant APP: Channel Intelligence, Acxiom, Crowd4Self

4.2 Categorization
This service can help requesters find the best classification for a product. For instance, to judge the sentiment of a blog post as positive, negative, or neutral. [11]
● Categorize Images and Content
● Classify Images
Relevant APP: Tagasauris, Stealth Requester, Aol, TagCow

4.3 Feedback
Provides reliable feedback to companies who want to get specific databases from their market. For instance workers are requested to test a new product, or to do some research on a new market the companies want to develop or invest in, etc.[12]
● Test Search Relevancy
● Product usability Testing
● Research
Relevant APP: Channel Intelligence,DomainPolish, EasyUsability, Feedback Army, PickFu, TryMyUI

4.4 Create & Moderate Content
It helps the requesters moderate and managecontent in their media websites, which include photos, videos, articles, etc. The goal can be to increase the visitors for the requesters' website.[13]
● Moderate Photos and Content
● Create and Edit Content
● Transcription
● Sentiment
● Tagging

4.5 Translation

It can be used to translate thousands documents every day from one language to another.[14]

Relevant APP: DARPA (Defense Advanced Research Projects Agency), SpeechPad, ReceiptLoader

5. UNINTENDED APPLICATIONS

Except official personal and corporate business, people also use Mechanical Turk to help them deal with their daily problems.

For instance,one person wants to change his hairstyle and then finds there are several styles he all prefers. He wants to figure out which one is his final choice. He can post a HIT on Mechanical Turk Web with those hairstyles and ask workers' suggestions. After the vote from the workers, he could decide which one to select according to the workers' feedback.

Furthermore, some scientists use Mechanical Turk to assist them improving their works. Recently, scientists used photographs to track animals, along with algorithmic pattern recognition.They post the task to ask Mechanical Turk Workers to help them identify animals in the images collected.[15]

6. GETTING STARTED

6.1 To requesters

In this part, there will be shown a brief introduction of how to create a Human Intelligence Task(HIT) , publish and manage the request. Here are a few steps to follow:

6.1.1 Making accounts

First of all, an Amazon Web Services (AWS) account should be created to enable access to all the services provided by Amazon. One can make an account at the AWS web site.[16]
Second, the requester should also have a registered Requester account. The account can be created at the Requester web site.[5]

6.1.2 Installing programming tools

To create a HIT, one can either choose a command line tool or SDK (Software Development Kit) where the SDK is designed for developers to make their own AMT applications.
1) Command line tools
There are two versions of command line tools available from Amazon: Windows and Unix based(i.e. Mac OS X, Linux, etc.). Go to Amazon Mechanical Turk Command Line Tools[17] and download the certain version relevant to the operating system.
For Unix based users: After downloading the Command Line Tool, go to [Command Line Tools Installation Directory]\directory and open file named "Overview.html". Follow "Overview.html" to install the command line tool.
For Windows users: Run "mech-turk-setup.exe" after downloading it.
For both users, an access key ID and secret access key are required for the final step of the setup.
2) SDKs(Software Development Kit)
If one wants to develop AMT applications, there are official SDKs for C#, JAVA, Perl and a library for Ruby. Just go to Sample Code & Libraries[18] website and find the corresponding programming SDK or library one wants to use and follow the instructions to get started.

6.1.3 Creating a HIT

In this section, a short instruction and a simple example of how to create and manage a HIT will be shown. First, here is the workflow of how a HIT is beingcreated, tested and managed in the whole process(Fig. 3).
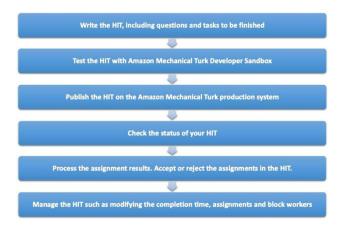

Figure 3, Workflow for requesters

The following example shows how to make a HIT (in this case a movie survey) using command line interface.
1) Make a tab-delimited text as the input file which could be named moviesurvey.input and populated it with the variable data for the HIT.
2) A question template file should be made and the text of your template should be added as well. It should be XML format which contains placeholders for the field defined in the input file and conforms to the QuestionForm[19]schema.

```
<?xml version="1.0" encoding="UTF-8"?>
<QuestionForm xmlns="http://mechanicalturk.amazonaws.com/AWSMechanicalTurk
DataSchemas/2005-10-01/QuestionForm.xsd">
  <Question>
   <QuestionIdentifier>1</QuestionIdentifier>
   <QuestionContent>
     <Text>${question}</Text>
   </QuestionContent>
   <AnswerSpecification>
     <FreeTextAnswer/>
   </AnswerSpecification>
  </Question>
</QuestionForm>
```

Figure 4, question template

3) Create a HIT properties file and populate it with the properties of the HIT.

```
#######################################
##  HIT Properties
#######################################

title:Movie Survey
description:This is a survey to find out how many movies you have watched
recently.
keywords:movie, survey
reward:0.05
assignments:100
annotation:sample#command

#######################################
## HIT Timing Properties
#######################################

# this Assignment Duration value is 60 * 60 = 1 hour
assignmentduration:3600

# this HIT Lifetime value is 60*60*24*3 = 3 days
hitlifetime:259200

# this Auto Approval period is 60*60*24*15 = 15 days
autoapprovaldelay:1296000
```

Figure5.properties file

4) Go to your [Command Line Tools Installation Directory]\bin directory and run commands.
5) Before you can publish your HIT, you would better test it in sandbox by including argument -sandbox with loadHITs command. You can modify your HIT freely until you get satisfied.
6) To publish the HIT to production system, open a command prompt, navigate to [Command Line Tools Installation Directory]\bin directory and run the following command.

```
loadHITs -input c:\moviesurvey\moviesurvey.input -question  c:\moviesur
vey\moviesurvey.question -properties c:\moviesurvey\moviesurvey.properties
```

Figure 6, Publish the HIT

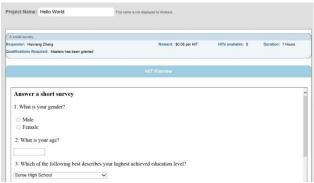7) Make a final check of the published HIT.(Figure 7.)

Figure 7, Published HIT

Amazon Mechanical Turk Command Line Reference[20] shows all commands that can be used to create and manage HITs.
It is also very convenient to create and publish HITs using other tools such as SDK. Go to Amazon Mechanical Turk Developer Guide[21] or Amazon Mechanical Turk Sample Code and Libraries[18] to get more information.

6.2 To workers
It is much easier for a worker to get started on AMT.
1) After one create Amazon Web Services (AWS) account, one can sign in Amazon Mechanical Turk Homepage[3] as a worker and start to search for some requests by just clicking the button called "Find HITs Now".See the appendix for screenshots of how your screen should look like at the end of these steps. Then a list of available HITs with their basic information such as HIT Expiration Date, Time Allotted and Reward will be shown to workers.(Figure 8.)

Figure 8. A list of HITs that are available to workers

2) If one finds some HITs interesting, they can be accepted and could be done by following the instructions. When the HITs are finished, they should be submitted.
3) If one's work is accepted by the requester, the payment will be transferred automatically to the worker's Amazon Payments account.

7. FINAL THOUGHTS

Although the core principle is hundreds of years old, Mechanical Turk implementations are livelier than ever thanks to the web. At first, replacing computer computation with human computation might not seem technical at all, but it is the communication between human requesters and workers that needs advanced web applications with automated processes. As long as there will be tasks more easily computed by humans than computers, and for the foreseeable future there will, Mechanical Turk technology will continue to see implementations.

REFERENCES

[1]   W. Mason, S. Suri. *Conducting behavioral research on Amazon's Mechanical Turk*.In *Psychonomic Society*, 44:1-23, 2012.

[2]   Wood, G. 2003. *Edison's Eve: A Magical History of the Quest for Mechanical Life*. New York: Bantam Doubleday Dell Publishing Group. ISBN:1400031583

[3]   Howe, J. 2008. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business.* New York: Crown Publishing Group. ISBN:0307449327

[4]   Amazon's Mechanical Turk platform. https://www.mturk.com

[5]   Amazon Mechanical Turk Requester website https://requester.mturk.com/

[6]   Amazon Mechanical Turk Creating and Managing Qualifications http://docs.aws.amazon.com/AWSMechTurk/latest/AWSMechanicalTurkRequester/Concepts_QualificationsArticle.html

[7]   The blog of A Computer Scientist in a Business School by PanosIpeirotis. http://www.behind-the-enemy-lines.com/2012/02/mturk-vs-odesk-my-experiences.html

[8]   Chilton, L. B., Horton, J. J., Miller, R. C., &Azenkot, S. (2010).Task search in a human computation market. In Proceedings of theACM SIGKDD Workshop on Human Computation(pp. 1–9).New York: ACM)

[9]   Amazon's Mechanical Turk platform. https://requester.mturk.com/tour/data_cleansing

[10]   Amazon's Mechanical Turk platform.https://requester.mturk.com/tour/categorization

[11]   Amazon's Mechanical Turk platform.https://requester.mturk.com/tour/business_feedback

[12]   Amazon's Mechanical Turk platform. https://requester.mturk.com/tour/content_creation_moderation

[13]   Amazon's Mechanical Turk platform.https://requester.mturk.com/case_studies/translation)

[14]   Amazon's Mechanical Turk Blog. http://mechanicalturk.typepad.com/blog/2013/05/mechanical-turk-workers-help-identify-endangered-species.html

[15]   Amazon Mechanical Turk Requester website https://requester.mturk.com/

[16]   Amazon Web Service website https://aws-portal.amazon.com/gp/aws/developer/registration/index.html

[17]   Amazon Mechanical Turk Command Line Tools website http://aws.amazon.com/developertools/Amazon-Mechanical-Turk/694

[18]   Amazon Mechanical Turk Sample code and Libraries http://aws.amazon.com/code/Amazon-Mechanical-Turk

[19]   Amazon Mechanical Turk Question From http://docs.aws.amazon.com/AWSMechTurk/latest/AWSMturkAPI/ApiReference_QuestionFormDataStructureArticle.html

[20]   Amazon Mechanical Turk Command Line Reference http://docs.aws.amazon.com/AWSMechTurk/latest/AWSMturkCLT/Welcome.html

[21]   Amazon Mechanical Turk Developer Guide http://docs.aws.amazon.com/AWSMechTurk/latest/AWSMechanicalTurkRequester/Welcome.html

[22]   eLaws. United States Department of Labor.Retrieved 5 November 2012.

[23]   United States Department of Labor.Retrieved 5 November 2012.

[24]   Iperirotis, P. 2010. *Demographics of Mechanical Turk*.