

Amazon Mechanical Turk

“ Artificial Artificial Intelligence”

Joe Kadhim
Vanessa Crittenden

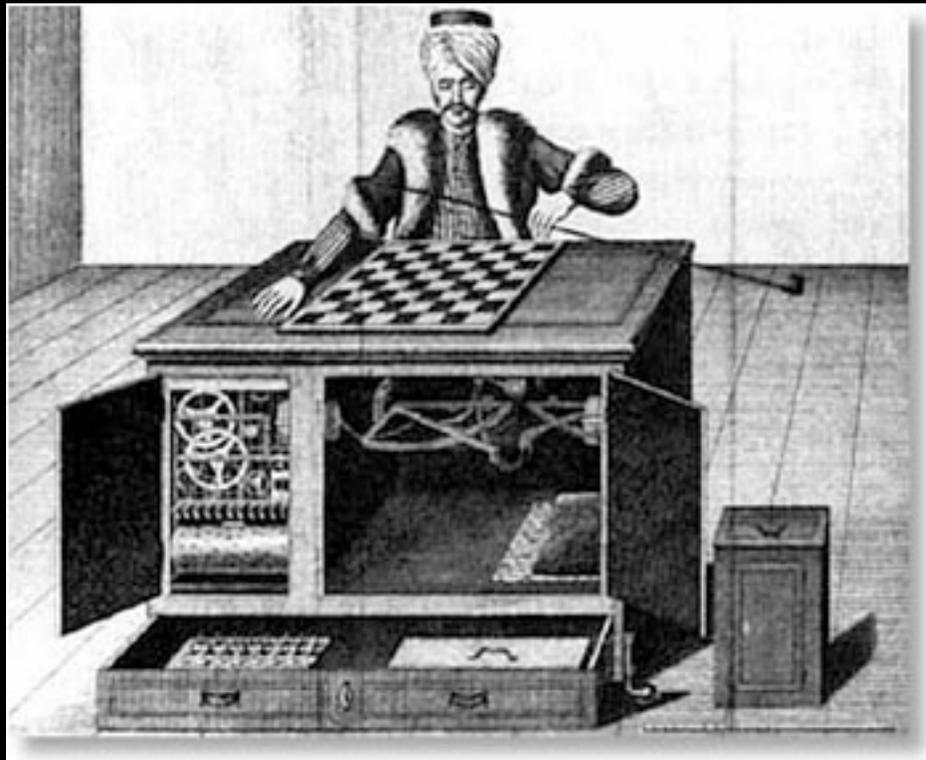
Agenda

- Why the name?
- Overview of the service
- History
- Terminology
- How it works - API
- Pros and Cons
- Related Services
- Conclusion / Future Plans

Why the name?

- In 1769, Wolfgang von Kempelen created a chess-playing machine called “The Turk”
- He traveled around Europe with “The Turk” claiming it was unbeatable
- The Turk beat many challengers, including Napoleon and Ben Franklin
- Turns out, a chess expert was hidden inside the machine

The First Mechanical Turk



Mturk, The Service

- Allows computers to perform tasks that are better executed by humans
- Provides an API for computers to integrate “artificial intelligence” into their processing by making requests of humans.
- Developers use the web service to submit tasks to the Amazon Mechanical Turk web site, approve completed tasks, and incorporate the answers into their software applications
- Individuals (workers) can then browse among existing tasks (HITs) and complete them for a fee offered by the developer (requester)

History

- Launched November 2, 2005
- In beta
- Initially created to solve some 'in house' issues at Amazon that required human judgment and intelligence
- Soon realized that this was a unique service that could be shared as a web service

Terminology

- HIT
- Requester
- Worker
- Qualification

HIT

- Human Intelligence Task
- Any task that can be completed using a computer connected to the internet
- Examples
 - Amazon
 - A9 BlockView -- match the best image to a business name
 - Extract Product Description -- Summarize
 - Select Japanese Text Direction -- Vertical or Horizontal
 - Identify Performers on a CD cover

HIT

- Examples cont...
 - External
 - Language Translation
 - Web Site Review
 - Spam Identification
 - Consumer Research
 - Add Keywords to Images
 - Identify Objects in Photographs
 - The list goes on....

Requester

- The individual, business or organization that creates the HIT
- Can specify worker qualifications
- Must deposit money into an Amazon account
- Submit Hits to the Turk Web Site (more on this later)
- Can choose to approve or reject a completed task
- Retrieve results

Worker

- Anyone with access to the web that wants to complete a task
- <http://www.mturk.com/mturk/welcome>
- Must complete qualification quiz if required
- Over time, stats about the workers develop -- allowing requesters to see how many HITs a worker has completed and the % of those that were approved
- A weekly \$\$ award is given to the top workers that have completed the most tasks with the highest approval ratings

Qualification

- The requester can post a 'quiz' that a worker is required to pass in order to do a task
- Can be multiple choice or in the form of an essay/text.

How it Works

- A requester can submit tasks in two ways:
 - Via the Mturk Requester site
<http://requester.mturk.com/mturk/welcome>
 - Via the web service API
 - If a developer wants to integrate human intelligence into new and existing application
 - If a business or organization needs to submit multiple HITs and automatically approve the results

API

- SOAP or REST
- How a developer might use it:

```
Read(photo);  
photoContainsHuman = callMechanicalTurk(photo);  
If(photoContainsHuman == TRUE)  
acceptPhoto;  
else  
rejectPhoto;
```

API cont..

- CreateHIT
- DisableHIT
- DisposeHIT
- ExtendHIT
- CreateQualificationType
- GrantQualification
- ApproveAssignment
- RejectAssignment

SOAP sample request

```
<?xml version="1.0" encoding="UTF-8" ?>
<soapenv:Envelope
  xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <soapenv:Body>
    <GetHIT
      xmlns="http://mechanicalturk.amazonaws.com/AWSMechanicalTurkRequester/2005-10-01">
      <AWSAccessKeyId>0PAP1H1P8JJVZEXAMPLE</AWSAccessKeyId>
      <Timestamp>2005-10-10T00:00:00.000Z</Timestamp>
      <Signature>[...]</Signature>
      <Request>
        <HITId>123RVWYBAZW00EXAMPLE</HITId>
        <ResponseGroup>Minimal,HITDetail</ResponseGroup>
      </Request>
    </GetHIT>
  </soapenv:Body>
</soapenv:Envelope>
```

SOAP sample response

```
<GetHITResponse>
  <OperationRequest>
    <RequestId>XA5TETQ3G6QF7EXAMPLE</RequestId>
  </OperationRequest>
  <HIT>
    <Request>
      <IsValid>>true</IsValid>
    </Request>
    <HITId>123RVWYBAZW00EXAMPLE</HITId>
    <CreationTime>2005-10-10T23:59:59.99Z</CreationTime>
    [... other fields of the HIT ...]
  </HIT>
</GetHITResponse>
```

Pros

- Get Internet scale work done quickly, efficiently
- Very cost effective
- Can monitor and manage quality of work
- New way of thinking about getting work done

Cons

- Virtual Sweatshop
- Requesters don't have to file tax forms, but workers are required to report their self-employment income
- Can be misused
 - Homework assignments

Related Services

- Google Answers
 - Cancelled last month
- Collaborative Human Interpreter
 - First implemented by Mturk
- MyCroft
 - Project from UC Berkeley School of Information
- Stardust@home
 - Distributes the huge task of tracking interstellar dust tracks

Conclusion

- Solves two problems
 - Human intelligence can be incorporated into web applications
 - Businesses can accomplish very big jobs at a low cost
 - People can make money (not really)
- Future plans
 - Worker API
 - Integration with IDEs

References

- <http://www.mturk.com/mturk/welcome>
- http://www.amazon.com/Mechanical-Turk-AWS-home-page/b/ref=sc_fe_l_2
- <http://developer.amazonwebservices.com/connect/kbcategory.jspa?categoryID=7>
- http://en.wikipedia.org/wiki/Amazon_Mechanical_Turk
- http://www.businessweek.com/the_thread/techbeat/archives/2005/1/1/amazons_mechani.html
- http://www.wired.com/wired/archive/14.06/crowds.html?pg=4&topic=crowds&topic_set