Crowdsourcing

Gareth J. F. Jones

Dublin City University, Ireland

NDP National Development Plan 2007 - 2013

sfi science foundation ireland

DCU Dublin City University

UCD DUBLIN University College Dublin

University of Limerick

Trinity College Dublin

Overview

- What is Crowdsourcing?
- Why use Crowdsoucring?
- Amazon Mechanical Turk
- High level issues in Crowdsourcing
- Getting started in Crowdsourcing
- Examples of Crowdsourcing in Information Technology
- Crowdsourcing Example:
  - MediaEval 2011 Rich Speech Retrieval

What is Crowdsourcing?

- Crowdsouring is a form of *human computation*.
- Human computation is a method of having people do things that we might consider assigning to a computing device, e.g. a language translation task.
- A *crowdsourcing system* facilitates a *crowdsourcing process.*
- A crowdsourcing system enlists a crowd of humans to help solve a problem defined by the system owners.
  - Currently best known crowdsourcing system is *Amazon Mechanical Turk*.

What is Crowdsourcing?

- The availability of crowdsourcing services is now making human computation easily available to the research community.
- There is currently significant interest in the exploring the use of crowdsourcing services to support existing research activities in information and data processing technologies;
- ... and how it might be used to open up new research directions, which might be technically innovative or previously have been impractical using other means.

What is Crowdsourcing?

- In doing this, a crowdsourcing system must address the following four issues:
  - How to recruit and retain workers?
  - What contributions can workers make?
  - How to combine worker contributions to solve the target problems?
  - How to evaluate workers and their contributions?

(Doan, Ramakrishnan and Halevy, CACM 2011)

What is Crowdsoucing?

- Collaborations between workers in crowdsourcing environments can be *explicit* or *implicit.*
- Some examples:
  - Wikipedia, Linux: the crowdsoucing system enlists a crowd of workers to *explicitly* collaborate to build a long lasting artefact of use to a larger community.
  - ESP game workers *implicitly* collaborate to label images as a side effect while playing a game.
  - Amazon Mechanical Turk workers collaborate *implicitly*, e.g. enlist workers to find a missing boat in thousands of satellite images.

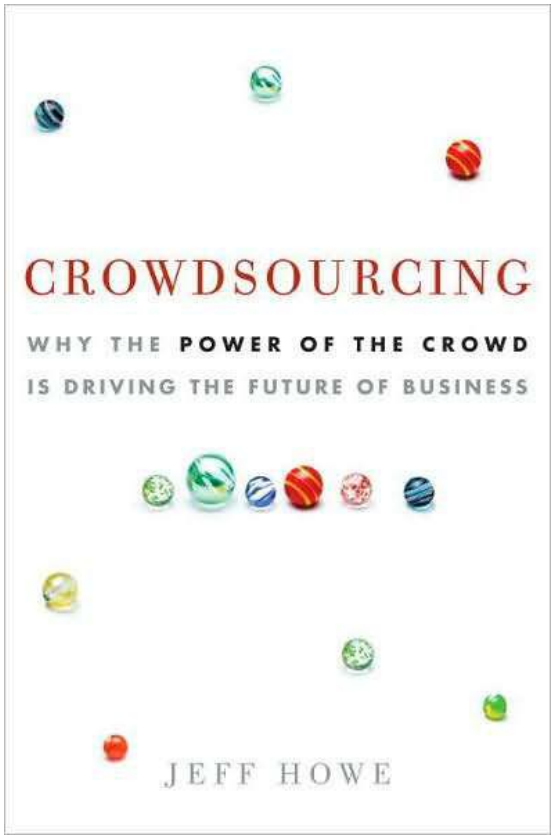
What is Crowdsoucing?

- But not all human-centric systems address these challenges, and such systems do not fall within the scope of crowdsourcing:
  - e.g. crowd management at a sports event does not look to recruit more members of the crowd, if anything in this case it would be preferable for members of the crowd to leave!

Creating a Crowdsourcing System

Basic idea:

- Take a task currently performed by a known agent, e.g. an employee.
- Outsource it to an undefined, often large, group of people, generally referred to as *workers*, via a call for participation.



Why use Crowdsourcing?

- Flexible, cheap, and fast completion of tasks
- Established infrastructure available:
  - MTurk payments, workforce, interface components
- Allows rapid early stage experimentation:
  - Define a task, e.g. gather some sort of data such as relevance assessments in information retrieval.
  - Rapidly develop prototype for workers to undertake the task.
  - Try out ideas straightaway with the available community of workers.
  - Examine output, iterate and try again or scale up to complete the full set of required tasks.

Who are the workers?

- Initially mainly in US, educated, female, bored, money not the key consideration, e.g. mothers at home with their children.
- Increasingly diverse and international, workers in some countries often more interested in money, e.g. student workers in India.

Amazon Mechanical Turk

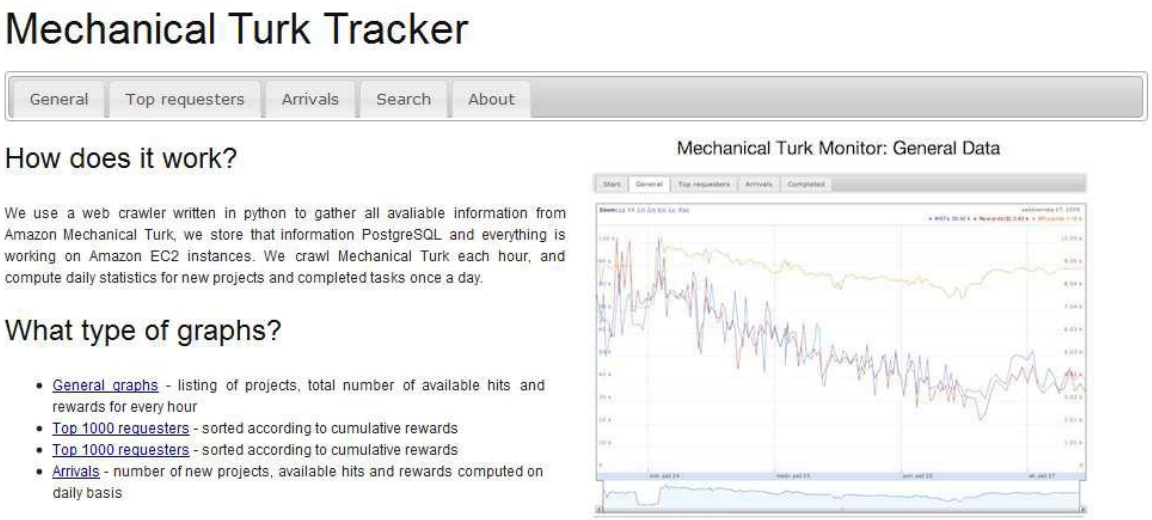
- Amazon Mechanical Turk (AMT) is:
  - “micro-task” crowdsourcing markplace
  - on-demand, scalable, real-time workforce
  - online since 2005
  - available to those with a suitable task to perform via “dashboard” GUI or programmers’ API

Amazon Mechanical Turk

- Tasks are setup and offered by *requesters*.
- A *requestor* creates a *Human Intelligence Task* (*HIT*), which is a web form composed of a number of instructions.
- HITs are undertaken by *workers* or for AMT *"turkers".*
- Requestor can specify the reward which is available for completing the HIT.
- AMT enables:
  - Requesters to get their task done.
  - Workers to make money.

Amazon Mechanical Turk

- AMT is not just a research toy, it is a business...
- From 1/09 – 4/10, 7M HITS from 10K requestors worth $500,000 (a significant underestimate).



- http://www.mturk-tracker.com
- Operation of AMT is supported by deducting a proportion of the requesters payments.

Other crowsdsourcing platforms

- Crowdflower
- CloudCrowd
- DoMyStuff
- Livework
- Clickworker
- Smartsheet
- uTest
- Elance
- oDesk
- vWorker

High Level Issues in Crowdsourcing

Effective use of crowdsourcing to support research requires a number of high level issues to be addressed.

- Relating to the task – research experimental design, then development of crowdsourcing task. Includes:
  - Determining payment level / work incentives
  - Interface and interaction design
- Choosing crowdsourcing platform
- Managing human factors relating to workers:
  - Recruitment - reputation of requester and worker
  - Retention
  - Quality Control / Data Quality
    - Trust, reliability, spam detection, consensus labelling

Recruitment

- Requester makes their task available to the registered workers.
- Worker requests to take the task they want to do.
- A fair requester is likely to be popular with workers, and highly rated workers will be popular with requesters.
- Requesters and workers with established strong reputations are likely to be successful in the recruitment.
  - Requester gets the worker they want.
  - Worker gets chosen for the task they want.
- A reputable requester can often have their pick from among reputable workers applying for their tasks.

Recruitment

- Practical issues may need to be addressed for a specific task, e.g. language translation; workers must have the requisite level of task specific skills.
- Requester can set a qualifying test before agreeing to let the worker take the task, e.g. to transfer and check some sample text.
- Alternatively, the quality of previous tasks completed can be taken proof of the worker's skills.

Recruitment

- The requestor has the option of approving completed work or not, if they don't approve the work, they don't pay the worker!
- Also, they have the option of paying a bonus to individual workers as part of the task specification.
- The requester can approve the work, pay for the work, and then not use, e.g. if it is clear that the worker expended the effort, but that the output is just not useable for some reason.
- The requester's fairness in recognising the genuine effort of workers can have a long term effect on their success in recruiting for subsequent tasks.

Reputation

- Workers can build a reputation within a crowdsoucring system.
- Details of previous tasks undertaken, etc can be available requesters to help them decide which offers to accept.
- This can act as an incentive for a requestor to accept their offer to undertake a task, and to trust the likely quality of their work.
- Requester can also can select a worker based on their own previous experience of the worker.

Reputation

- Reputation of both requesters and workers can be important in crowdsourcing.
- A requester with a reputation for posting well structured and clearly described tasks, with fair and prompt payments is likely to prove popular with regular workers.
- A worker who undertakes tasks in a professional manner is likely to prove popular with requesters.

Payment and Incentives

- Workers undertake tasks for micro-payments – very small payments for individual subtasks.
- Workers may volunteer for an offered tasks because it looks interesting, often they will do so because it looks a god way to earn some money.
- Offering a suitable level of payment is a trade-off between:
  - Underpayment for either or both of worker time or expertise.
  - Overpayment – which may attract workers keen to earn money without undertaking the task properly.
  - Sufficient payment to motivate worker – it can be possible to award bonus payments for excellent work.

Spam Detection

- It is important that workers are actually capable of successfully completing their assigned task.
- Not all tasks require specific expertise or skills, but it is important that the worker should undertake the task to the best of their ability.
- Workers may sign up for a task to earn money (and even pass a qualifying test is required!), and then attempt to get paid without completing the task properly.
- Detecting so called “spam” work is an important issue in quality control.
- Some tasks require all work to be checked, for some tasks checking all work is impractical.

Spam Detection

- One method to combat the potential problem of spam is for requesters to set up “honey pots” with known answers.
- The worker completes these as part of their work, but unbeknownst to them, the requester knows the answer to the honey pot questions and can easily check for faked or poor quality work.
- The requester can then refuse to pay for the work, and bar the worker from undertaking further work for them.
- The worker’s reputation within the system will also fall making it more difficult to get work with other requesters.

Getting Started in Crowdsourcing

- Sign up as a worker and do some HITs.
- Monitor discussion forums.
- Address feedback, e.g. poor guidelines, payments, passing grade, etc. – email exchange.
- Everything counts!
  - The HIT is only as good as the weakest part!

Getting Started in Crowdsourcing

- Have a research problem requiring input from human participants.
- Recruitment for conventional completion of the task is problematic, e.g. limited availability of workers, too difficult to recruit, too expensive, too slow.
- Is the problem crowdsourcable?
  - Can it be split into individual HITs for multiple non-expert workers?

Implementation of an AMT HIT

- Build a mock up and test it locally with your research team, e.g. design working test for practicality of HIT and interpretation of instructions.
- Incorporate feedback, refine task.
- Implement a small test run with AMT with a very small data set:
  - Time the completion of the HIT
  - Do people understand the task?
  - Payment rate for work – too high / too low?

Implementation of an AMT HIT

- Analyze results:
  - Look for spammers.
  - Check for completion times.
  - Check suitability or correctness of the output.
- Iterate and modify accordingly
  - Review comments from turkers
    - are they happy / unhappy?
    - would they work for you again?
  - HIT too easy / too hard?
  - Payment too high / too low?
- If all is in order launch a full *batch* of the HIT.

Implementation of an AMT HIT

- Introduce quality control:
  - Qualification test?
  - Gold answers (honey pots) – trap spammers.
- Adjust qualification passing grade and/or acceptable approval rate of workers if necessary.
- Run experiment with new settings and the same data.
- Other issues:
  - Many tasks active on AMT at any point.
  - If you want the best workers, you need to grab their attention.
- How to schedule:
  - Split a large task into batches; only have one batch in the system at a time.
  - Review feedback from batch *n* before uploading batch *n+1.*

Examples of Crowdsourcing in IT

- Many existing examples of use of crowdsourcing in diverse research areas including: IR, NLP, speech processing, computer vision, user studies, usability testing, psychological studies, surveys, ...
- NLP examples:
  - machine translation (MT) – evaluate quality of the output of a MT system
  - Speech data transcription
- Snow et al (2008): demonstrated high agreement between crowd and experts for 5 NLP based task.
- Emerging examples in social data analytics (Willett et al., 2012).

Why use Crowdsourcing in IR?

- Crowdsourcing can be a useful tool for IR, but it doesn't do the job for you, high quality research is still hard work!
- Effective crowdsourcing is more of art than science, there is no substitute for experience, and you shouldn't rely on getting it right first time.
- Careful experimental design still required.

Why use Crowdsourcing in IR?

- Shown to be effective in major shared task evaluations:
  - ImageCLEF, TREC, INEX, MediaEval
- IR has a number of tasks amenable to crowdsourcing – require “users” to complete non-expert tasks carefully:
  - Relevance assessment
  - Request creation
  - Topic labelling
  - Structuring of content, etc.

Relevance Assessment for IR

- Standard IR test collection development model:
  - Develop user task scenario
  - Identify document set
  - Create evaluation search topics
  - Recruit relevance assessors:
    - Typically small number of assessors, may be biased, may not be sufficiently expert.
    - In many situations insufficient assessors available – researchers would like to do more assessments, but cannot find or maybe pay for sufficient assessors.

Relevance Assessment for IR

- Crowdsourcing model for relevance assessment stage:
  - Split assessment requirements into tasks.
  - Design, implement and publish task.
  - Recruit assessors - potentially immediately available.
    - Removes local assessors constraints, e.g. the number available, expertise limitations.
  - Can be completed quickly without needing to book laboratories, arrange sessions with assessors, etc.

Crowdsourcing Example

- MediaEval 2011 Rich Speech Retrieval task
- Task to try to locate the optimal point to start playback of the video in a single video known to be relevant to a searcher's information need. Referred to as a jump-in point for a known-item search.
  - Models user trying to re-find a previously viewed segment of video.
- Develop test collection for a speech search evaluation task.
- The crowdsourcing task is to find a number of interesting jump-in points, to describe them and to form a search topic statement for each one.

Crowdsourcing Example

- The evaluation task is then to use the topic statement to try to locate the jump-in point.
- Document set 1974 episodes (247 dev, 1727 test) 350 hours of semi-professional video harvested from blip.tv.
- Available for download under a creative commons licence.

Crowdsourcing Example

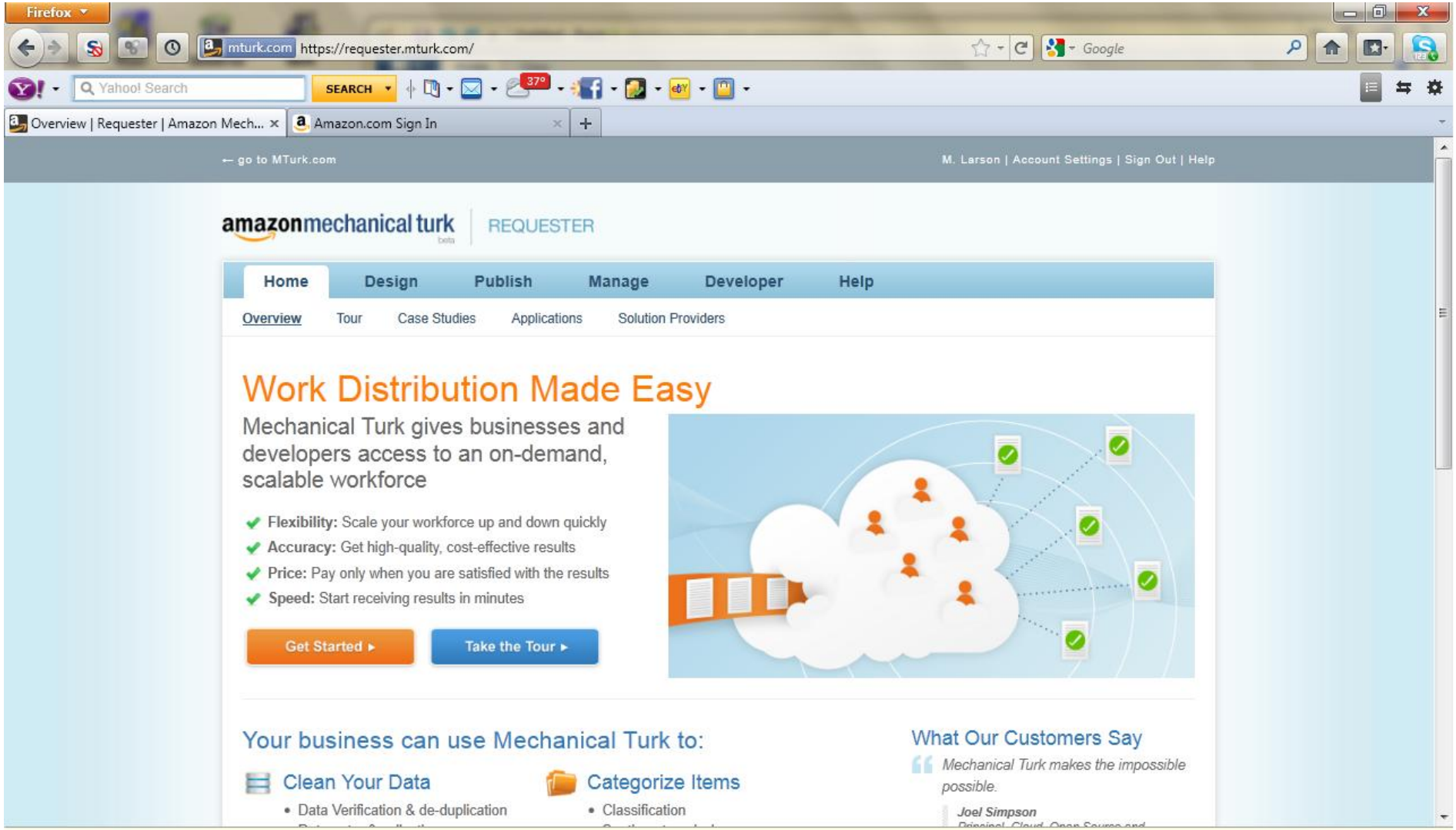
- Task uses five different functions of speech, represented as illocutionary speech acts: `apology', 'definition', 'opinion', 'promise' and 'warning'.
- To carry out the evaluation task participants were required to develop a system that responds to a user query by retrieving a ranked list of jump-in points.
- Create queries and relevant jump-in points for evaluation using an AMT HIT.
- The following slides illustrate the details of the task as well as the design of the HIT with screenshots for the requester and the workers.

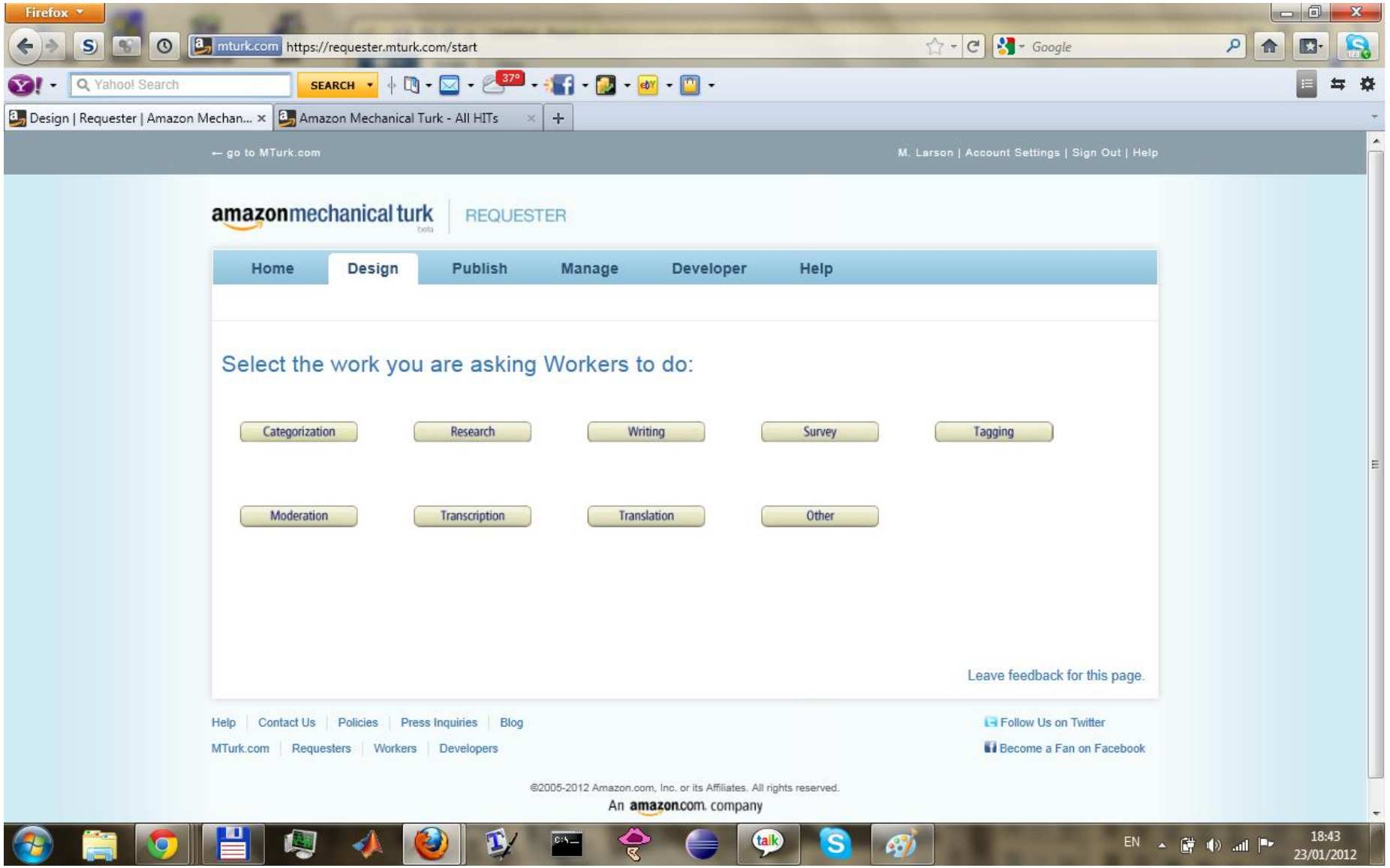
MediaEval 2011 RSR AMT Example

- First step is for the requester to develop a task for the worker to undertake the work.
- An AMT HIT is designed on an HTML form.
- Several options are available to obtain an HTML form for a HIT:
  - Use one of the provided templates.
  - Download one of the templates, edit and upload.
  - Write your own.
  - Select an existing template that you uploaded previously.

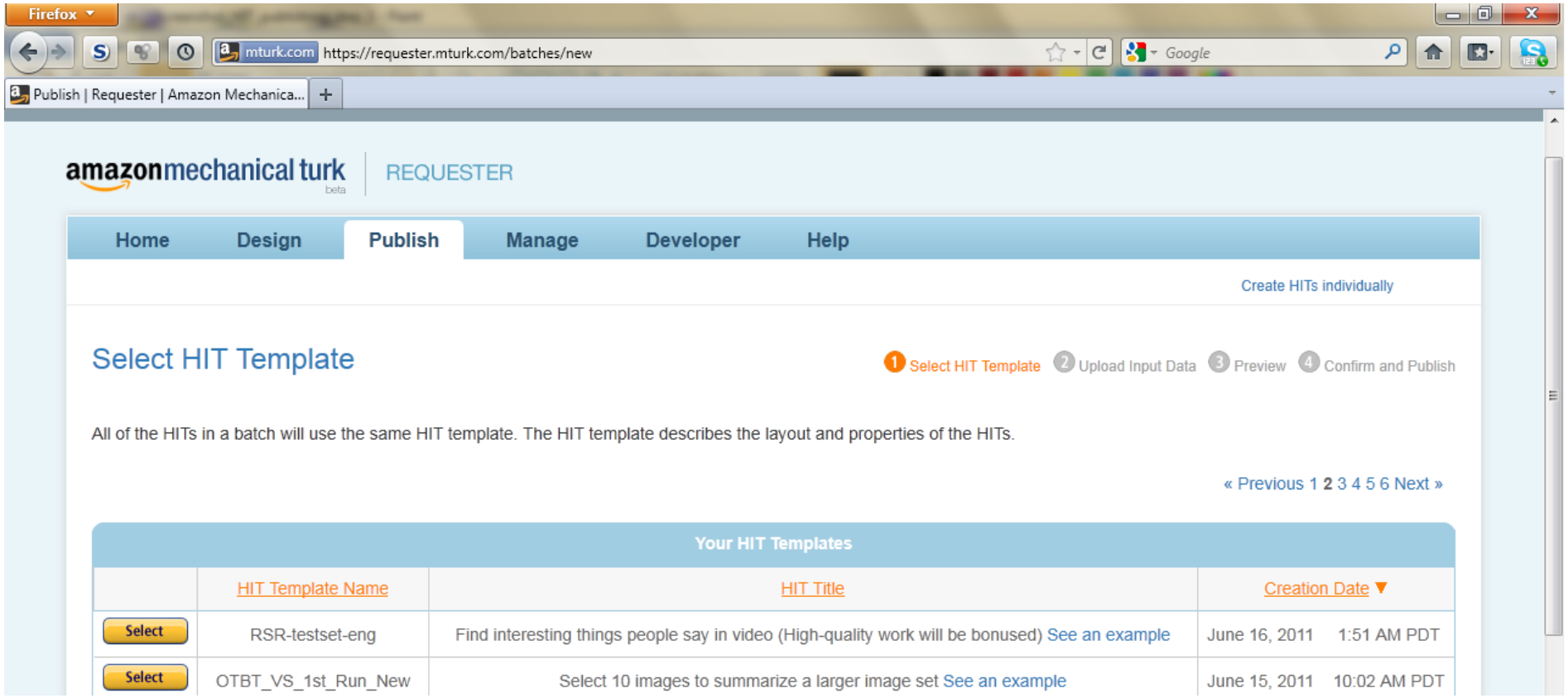
MediaEval 2011 RSR MTurk Example

- The next slides show the requester entry point into AMT.
- Develop the task, AMT HIT, by first selecting the *Design* option.
- Requester can:
  - Select from the available alternative HIT template examples for different types of work task,
  - Select from their previously uploaded HIT templates
  - Upload a new HIT template.


Firefox
mturk.com https://requester.mturk.com/
Google
Yahoo! Search
SEARCH
Overview | Requester | Amazon Mech...
Amazon.com Sign In
← go to MTurk.com
M. Larson | Account Settings | Sign Out | Help
amazonmechanical turk
REQUESTER
Home
Design
Publish
Manage
Developer
Help
Overview
Tour
Case Studies
Applications
Solution Providers
Work Distribution Made Easy
Mechanical Turk gives businesses and developers access to an on-demand, scalable workforce
Flexibility: Scale your workforce up and down quickly
Accuracy: Get high-quality, cost-effective results
Price: Pay only when you are satisfied with the results
Speed: Start receiving results in minutes
Get Started ▸
Take the Tour ▸
Your business can use Mechanical Turk to:
Clean Your Data
Data Verification & de-duplication
Categorize Items
Classification
What Our Customers Say
Mechanical Turk makes the impossible possible.
Joel Simpson

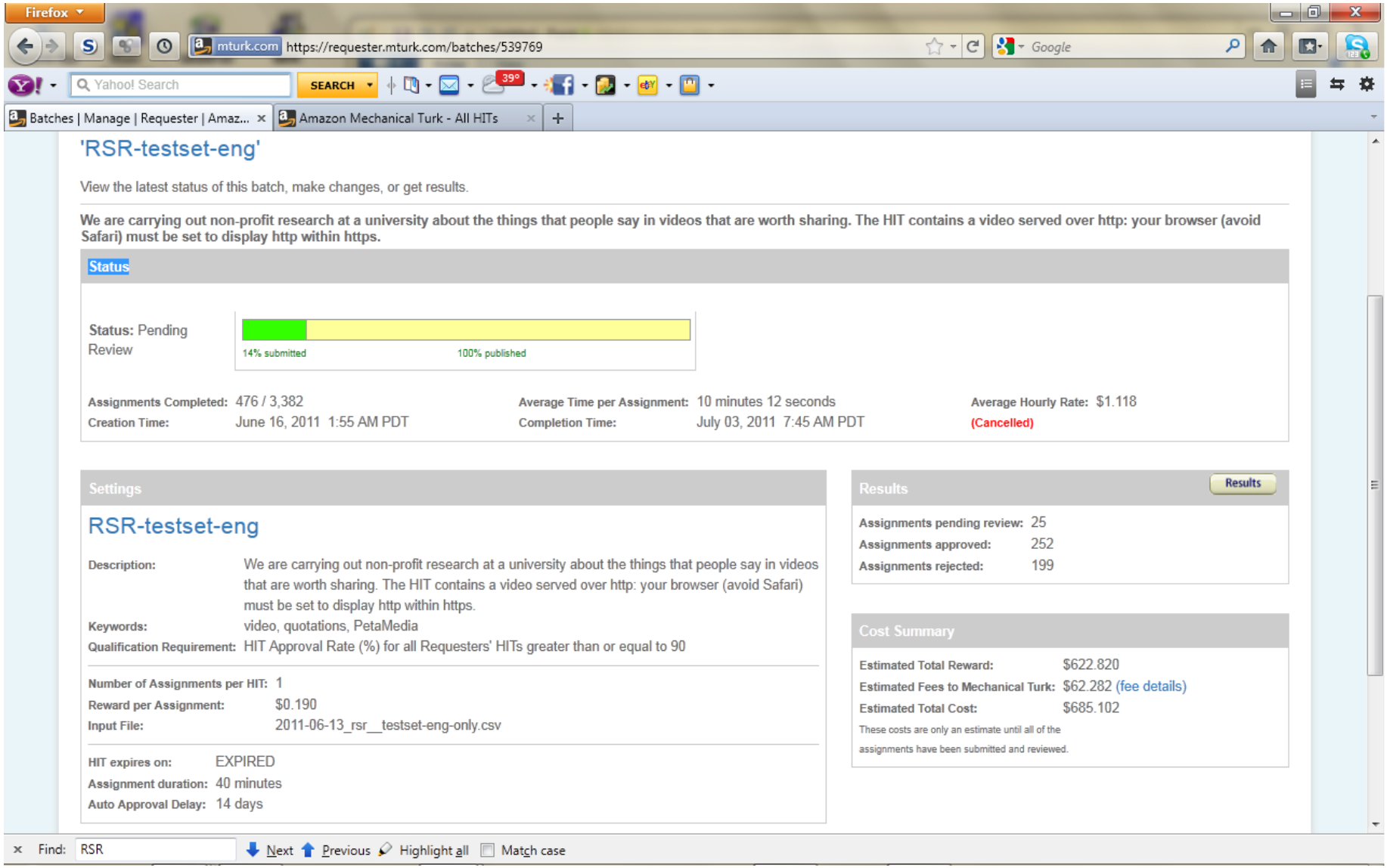

Firefox
mturk.com https://requester.mturk.com/start
Google
Yahoo! Search
SEARCH
Design | Requester | Amazon Mechan...
Amazon Mechanical Turk - All HITs
← go to MTurk.com
M. Larson | Account Settings | Sign Out | Help
amazonmechanical turk
REQUESTER
Home
Design
Publish
Manage
Developer
Help
Select the work you are asking Workers to do:
Categorization
Research
Writing
Survey
Tagging
Moderation
Transcription
Translation
Other
Leave feedback for this page.
Help | Contact Us | Policies | Press Inquiries | Blog
MTurk.com | Requesters | Workers | Developers
Follow Us on Twitter
Become a Fan on Facebook
©2005-2012 Amazon.com, Inc. or its Affiliates. All rights reserved.
An amazon.com company
EN
18:43
23/01/2012

Requester can select from their HITS



MediaEval 2011 RSR MTurk Example

- Once launched, the requester can monitor the progress of the current batch of instance of the HIT.
- The following screen shows the percentage of the requested task completed so far by the workers.
- This also shows the average time to complete a HIT and the average rate of pay for the work.
  - This information is useful to the requester in assessing the effectiveness of the HIT and to potential workers to decide whether to apply for the HIT.


Firefox
mturk.com https://requester.mturk.com/batches/539769
Google
Yahoo! Search
SEARCH
Batches | Manage | Requester | Amaz...
Amazon Mechanical Turk - All HITs
'RSR-testset-eng'
View the latest status of this batch, make changes, or get results.
We are carrying out non-profit research at a university about the things that people say in videos that are worth sharing. The HIT contains a video served over http: your browser (avoid Safari) must be set to display http within https.
Status
Status: Pending Review
14% submitted
100% published
Assignments Completed: 476 / 3,382
Creation Time: June 16, 2011 1:55 AM PDT
Average Time per Assignment: 10 minutes 12 seconds
Completion Time: July 03, 2011 7:45 AM PDT
Average Hourly Rate: $1.118
(Cancelled)
Settings
RSR-testset-eng
Description: We are carrying out non-profit research at a university about the things that people say in videos that are worth sharing. The HIT contains a video served over http: your browser (avoid Safari) must be set to display http within https.
Keywords: video, quotations, PetaMedia
Qualification Requirement: HIT Approval Rate (%) for all Requesters' HITs greater than or equal to 90
Number of Assignments per HIT: 1
Reward per Assignment: $0.190
Input File: 2011-06-13_rsr__testset-eng-only.csv
HIT expires on: EXPIRED
Assignment duration: 40 minutes
Auto Approval Delay: 14 days
Results
Results
Assignments pending review: 25
Assignments approved: 252
Assignments rejected: 199
Cost Summary
Estimated Total Reward: $622.820
Estimated Fees to Mechanical Turk: $62.282 (fee details)
Estimated Total Cost: $685.102
These costs are only an estimate until all of the assignments have been submitted and reviewed.
Find: RSR
Next
Previous
Highlight all
Match case

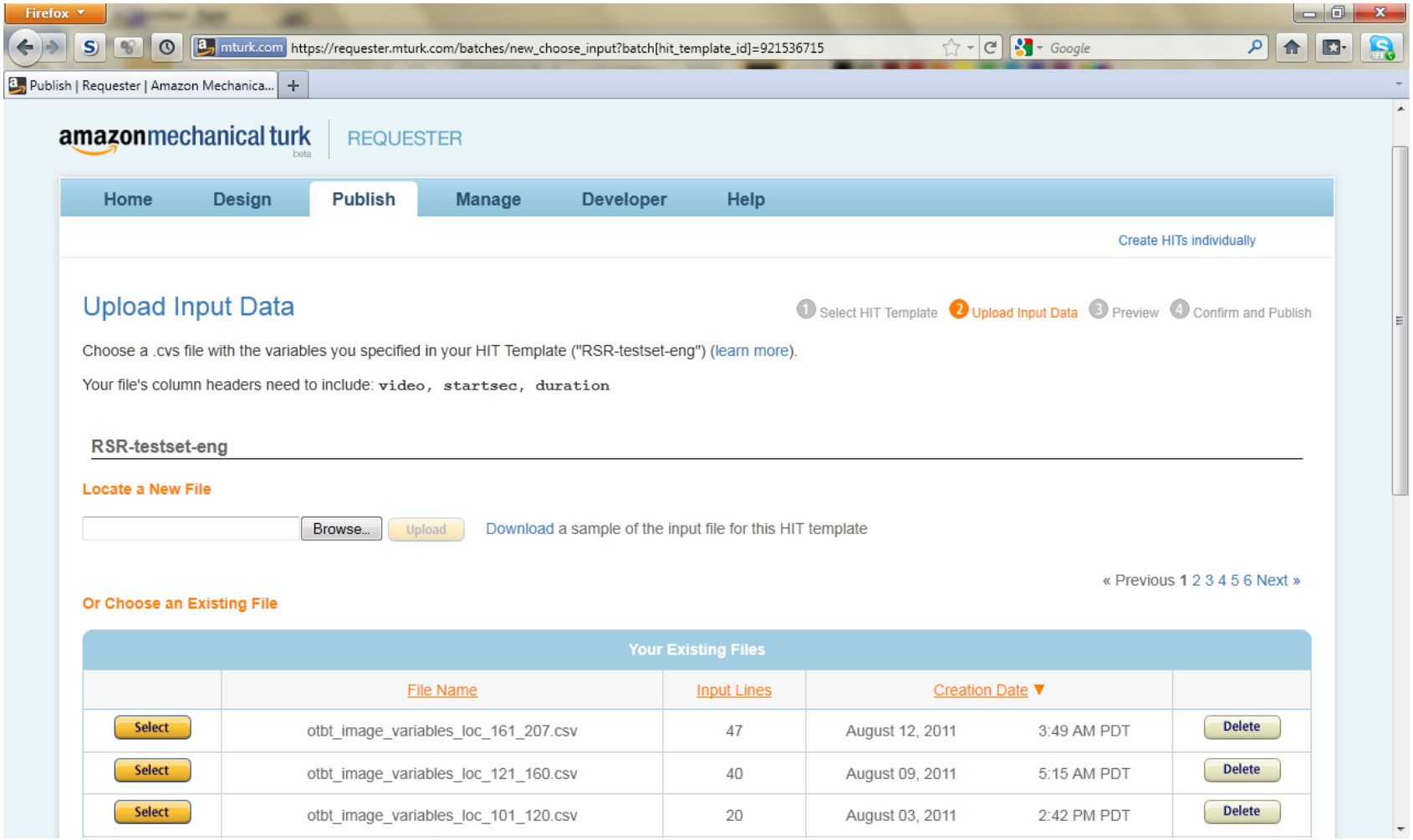
MediaEval 2011 RSR MTurk Example

- The MediaEval 2011 RSR HIT was written from scratch and uploaded to AMT.
- Completion of the overall task required multiple workers to complete the HIT by viewing different video file, the page required input variables to this HIT page to specify the video to be viewed in this instance of the HIT.
- The variables gave details of the path to the server where the video to be viewed in the HIT was stored.
- The code also specified the video player to be used and the video to be played.

Setup Batch for HIT

- The batch of instances of the HIT are defined in a csv file uploaded by the requester.
- The values of the variables (the names of the videos in this case) are specified in the file.
- During execution of the batch AMT keeps track of which HITs have been completed.
- Workers are assigned HITs until either the batch has been completed, or the requester stops execution of further HITs in this batch.

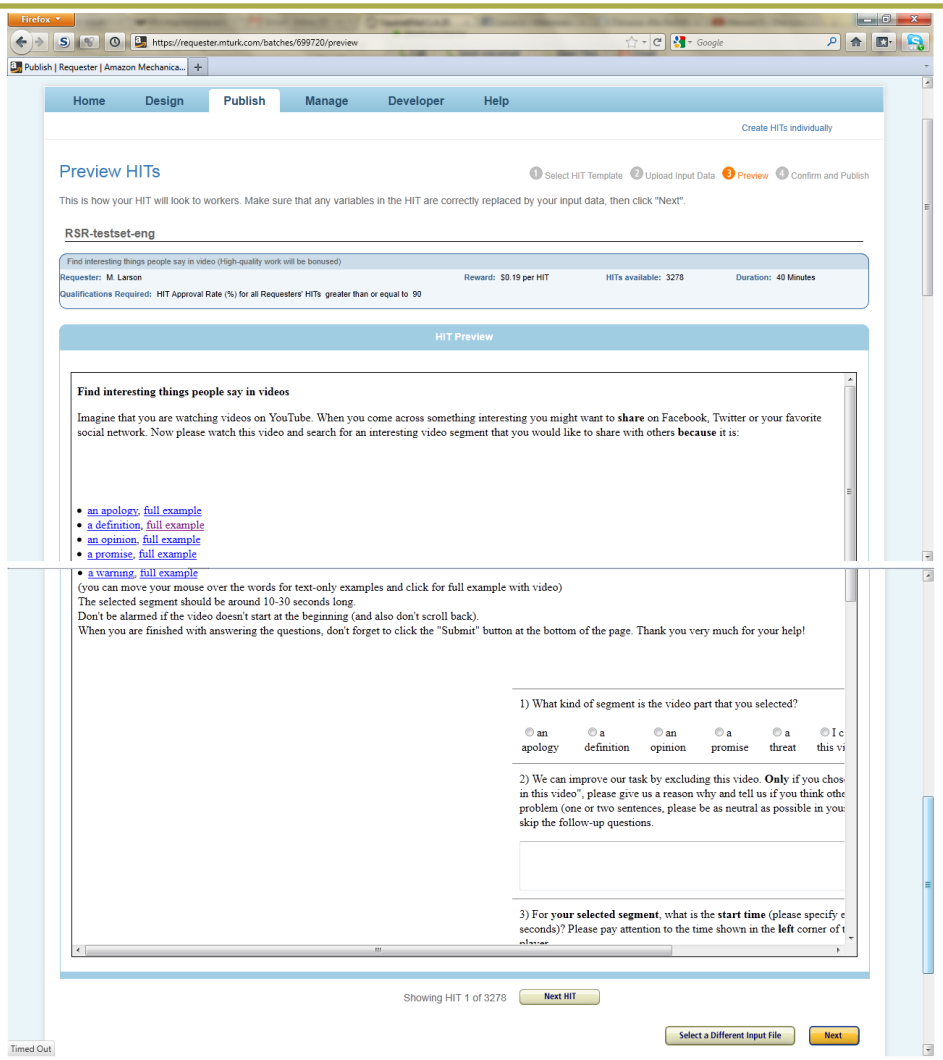
Upload Input Data for HIT



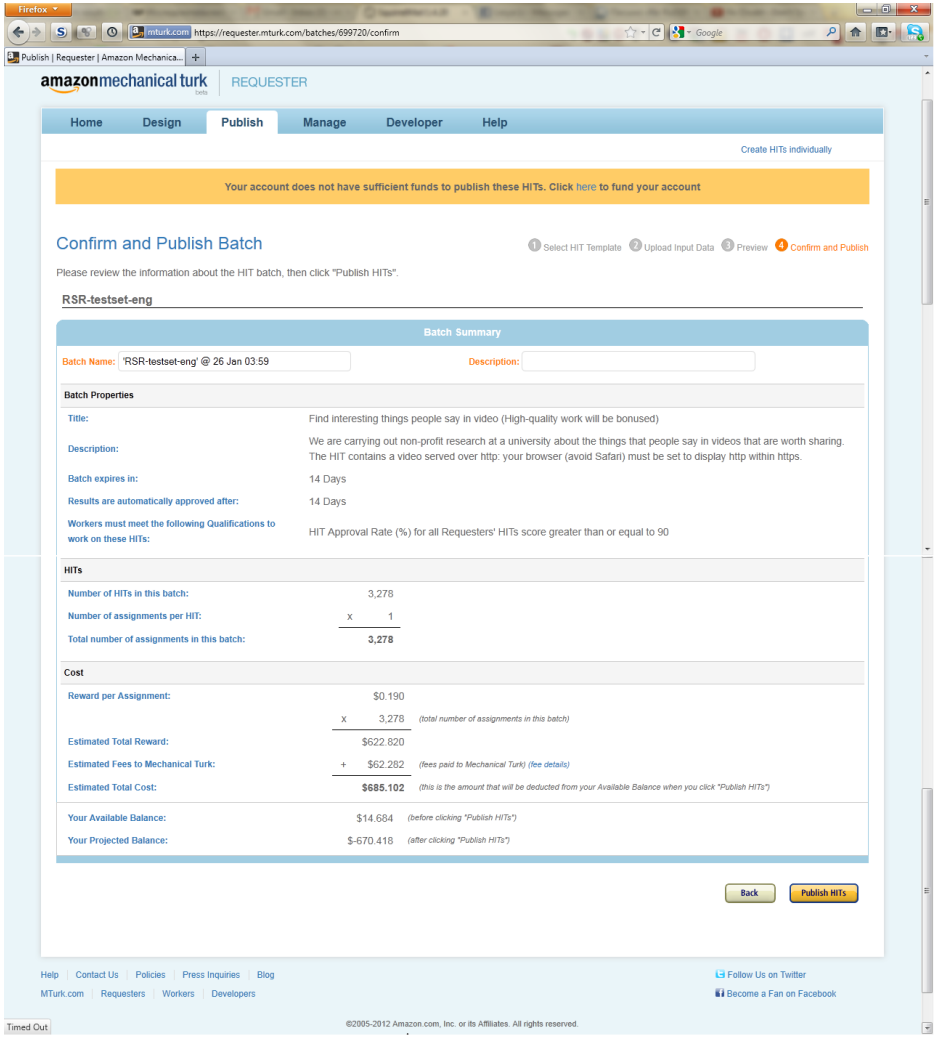
Preview and Publish HIT

- Once uploaded the requester can preview the HIT as it will be seen by the worker.
- When the requester is satisfied with the HIT, they can publish it to make it available to the workers.
- Note that for AMT, the requester must have sufficient credit registered in the system to pay for completion of the batch.

Preview the HIT



Publish the HIT



Details of the MediaEval 2011 RSR HIT

Find interesting things people say in videos

Imagine that you are watching videos on YouTube. When you come across something interesting you might want to **share it** on Facebook,Twitter or your favourite social network.

Now please watch this video and search for an interesting video segment that you would like to share with others **because** it is:

- an apology, full example
- a definition, full example
- an opinion, full example
- a promise, full example
- a warning, full example

(you can move your mouse over the words for text-only examples and click for full example with video)

Details of the HIT

- The selected segment should be around 10-30 seconds long. **Don't be alarmed if the video doesn't start at the beginning (and also don't scroll back).**
- When you are finished with answering the questions, don't forget to click the "Submit" button at the bottom of the page. Thank you very much for your help!

Details of HIT

- 1) What kind of segment is the video part that you selected?
  - an apology
  - a definition
  - an opinion
  - a promise
  - a warning
  - I can't find anything like this in this video

- 2) We can improve our task by excluding this video. Only if you chose "I can't find anything like this in this video", please give us a reason why and tell us if you think other people will have the same problem (one or two sentences, please be as neutral as possible in your description), and you should skip the follow-up questions.

Details of HIT

- 3) For **your selected segment** (in 1) above, what is the **start time** (please specify exactly in minutes and seconds)? Please pay attention to the time shown in the **left** corner of the bottom line of the video player.
  - Minute
  - Second
- 4) For **your selected segment** (in 1) above, what is the **end time** (please specify exactly in minutes and seconds)? Please pay attention to the time shown in the **left** corner of the bottom line of the video player.
  - Minute
  - Second

Details of HIT

- 5) What was said during **your selected segment**? Please write down the **exact words** the speaker is saying (please transcribe precisely). If you are not sure what the exact word was, please write down what your think the word was and mark it with a star (for example, 'French president *Sarkosie was saying ...' if you are not sure how to spell the name 'Sarkozy' properly)

Details of HIT

- 6) When sharing this particular part of the video (your selected segment) on a social network, what **comment** would you add to the video to make sure that your friends have an idea what the video segment is about?

  Please do not use informal internet language (such as '4 u‘ instead of 'for you').

  Be as objective as possible when describing the video segment and do not express your personal opinion/attitude, either positive or negative.

- 7) Imagine you would like to search for **similar video segments** using a search engine (such as Google, Bing, Yahoo) what would you put in the search box?

Details of HIT

- We understand that this work requires a lot of your time and concentration, so we would like to bonus the high-quality of your results.
- Please tell us your opinion about the size of bonus you deserve. Choose and justify your choice. Please keep in mind that we are carrying out non-profit university research (we can afford a maximum of 21 cents bonus, but only for really excellent responses).
- When making our decision on your bonus level we create a compromise between our budget and your request.

  0 cents

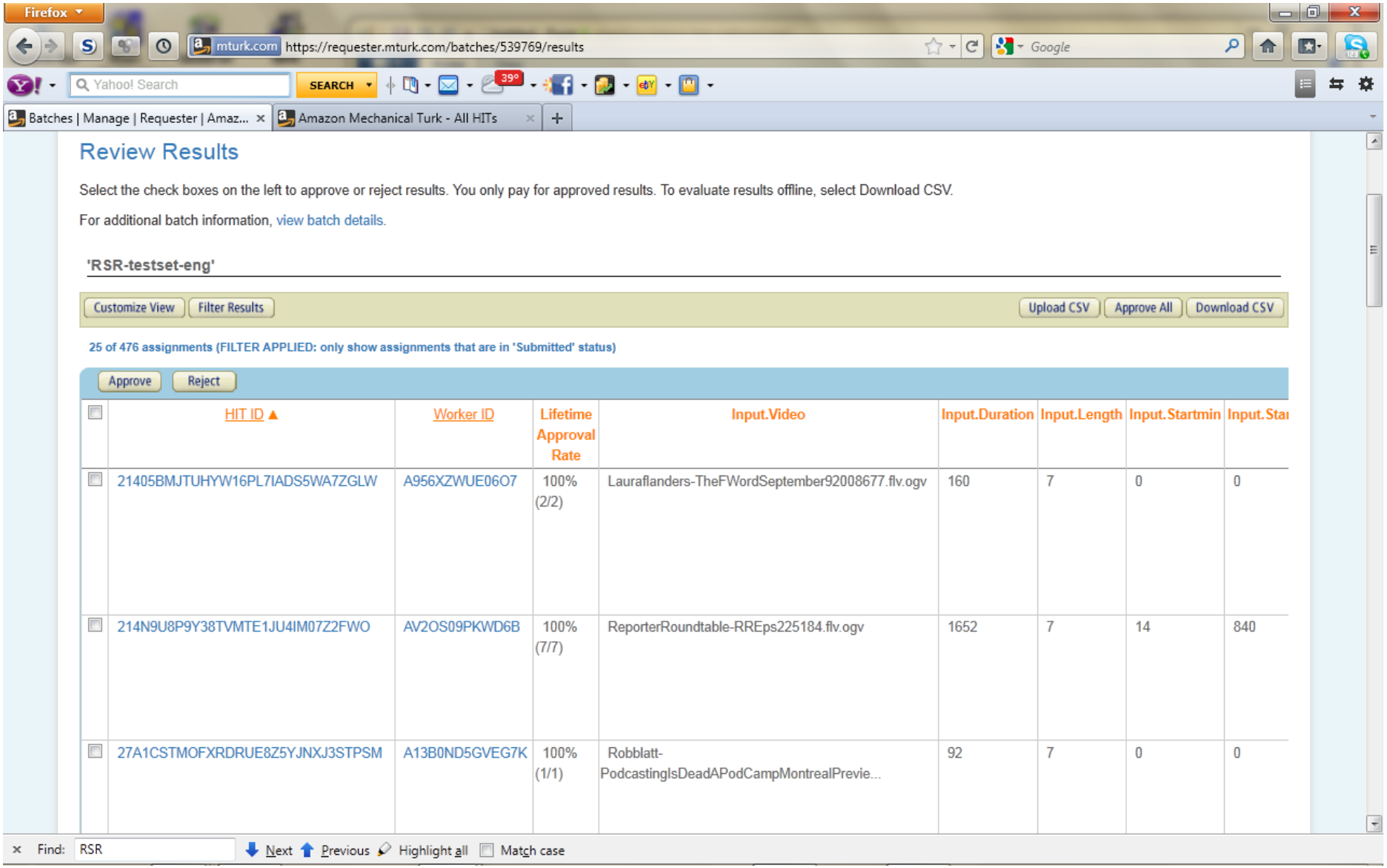
  7 cents

  11 cents

  21 cents

Details of Completed HITs

- The requester can view a summary of the workers who have selected and completed the HIT and their activities, shown in the next slide.
- A more detailed version can be downloaded in a csv file.
- The requester can also review the HIT exactly as seen by the worker, the next slide.
  - In the case of the RSR HIT, this means that the requester can see the specific video viewed by the worker. This enables to check and confirm the details of the work delivered.



Firefox

mturk.com https://requester.mturk.com/batches/539769/results

Google

Yahoo! Search SEARCH

Batches | Manage | Requester | Amaz... × Amazon Mechanical Turk - All HITs ×

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.

'RSR-testset-eng'

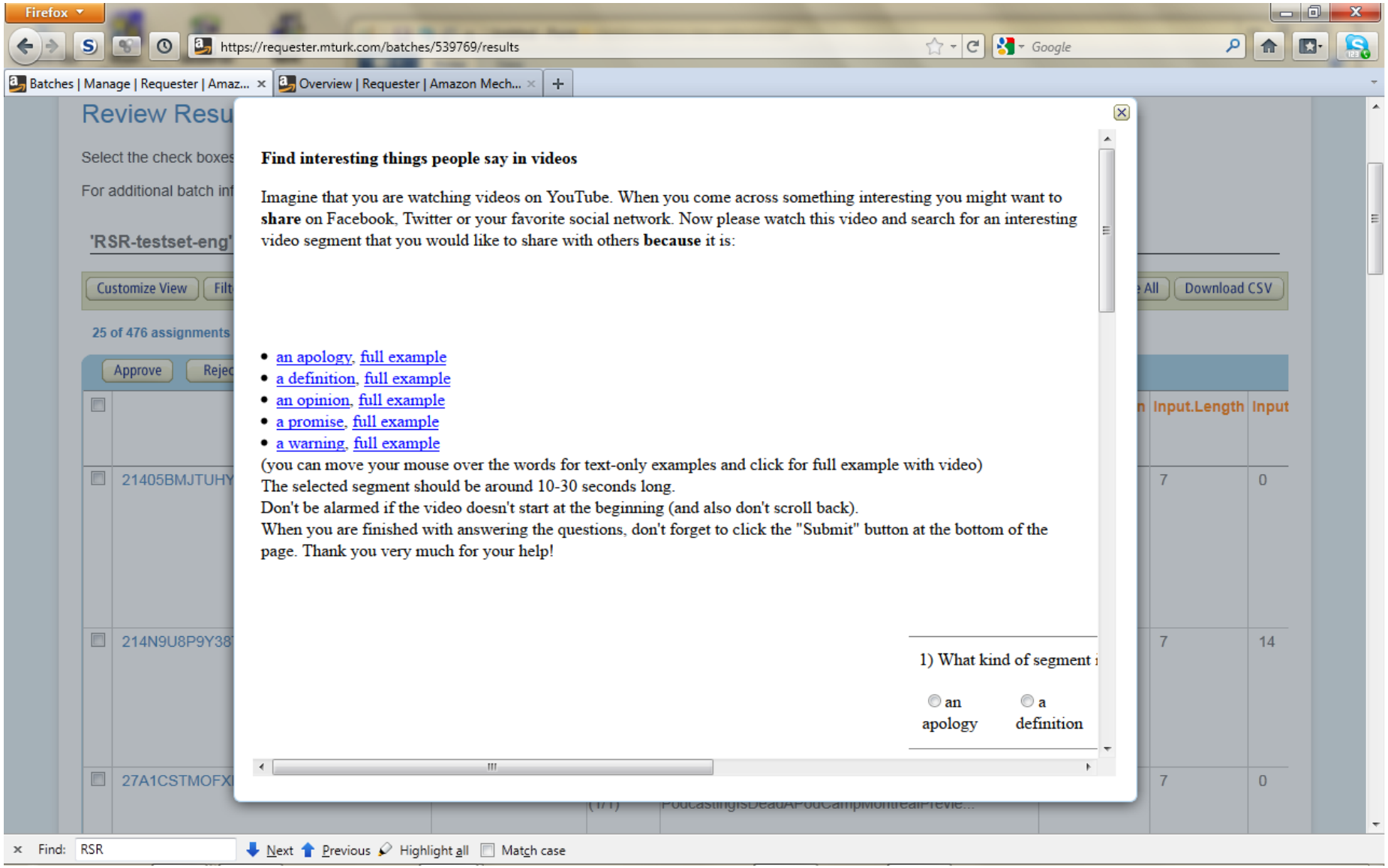
Customize View | Filter Results | Upload CSV | Approve All | Download CSV

25 of 476 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

Approve | Reject

HIT ID ▲	Worker ID	Lifetime Approval Rate	Input.Video	Input.Duration	Input.Length	Input.Startmin	Input.Sta
21405BMJTUHYW16PL7IADS5WA7ZGLW	A956XZWUE06O7	100% (2/2)	Lauraflanders-TheFWordSeptember92008677.flv.ogv	160	7	0	0
214N9U8P9Y38TVMTE1JU4IM07Z2FWO	AV2OS09PKWD6B	100% (7/7)	ReporterRoundtable-RREps225184.flv.ogv	1652	7	14	840
27A1CSTMOFXRDRUE8Z5YJNXJ3STPSM	A13B0ND5GVEG7K	100% (1/1)	Robblatt-PodcastingIsDeadAPodCampMontrealPrevie...	92	7	0	0

× Find: RSR Next Previous Highlight all Match case



Find interesting things people say in videos

Imagine that you are watching videos on YouTube. When you come across something interesting you might want to **share** on Facebook, Twitter or your favorite social network. Now please watch this video and search for an interesting video segment that you would like to share with others **because** it is:

- an apology, full example
- a definition, full example
- an opinion, full example
- a promise, full example
- a warning, full example

(you can move your mouse over the words for text-only examples and click for full example with video)
The selected segment should be around 10-30 seconds long.
Don't be alarmed if the video doesn't start at the beginning (and also don't scroll back).
When you are finished with answering the questions, don't forget to click the "Submit" button at the bottom of the page. Thank you very much for your help!

1) What kind of segment

an apology a definition

Notes on the MediaEval 2011 RSR HIT

- There was no qualification requirement for the HIT. However, workers had to have a 90% acceptance rate by requesters for their previous work.
- A small scale initial run received negative feedback because the review was judged to be too low for the work required. Increasing reward for subsequent batch fixed this problem.
- Also, indicating that we were working for a non profit organisation meant that workers were more willing to accept the level of payment on offer.

Notes on the MediaEval 2011 RSR HIT

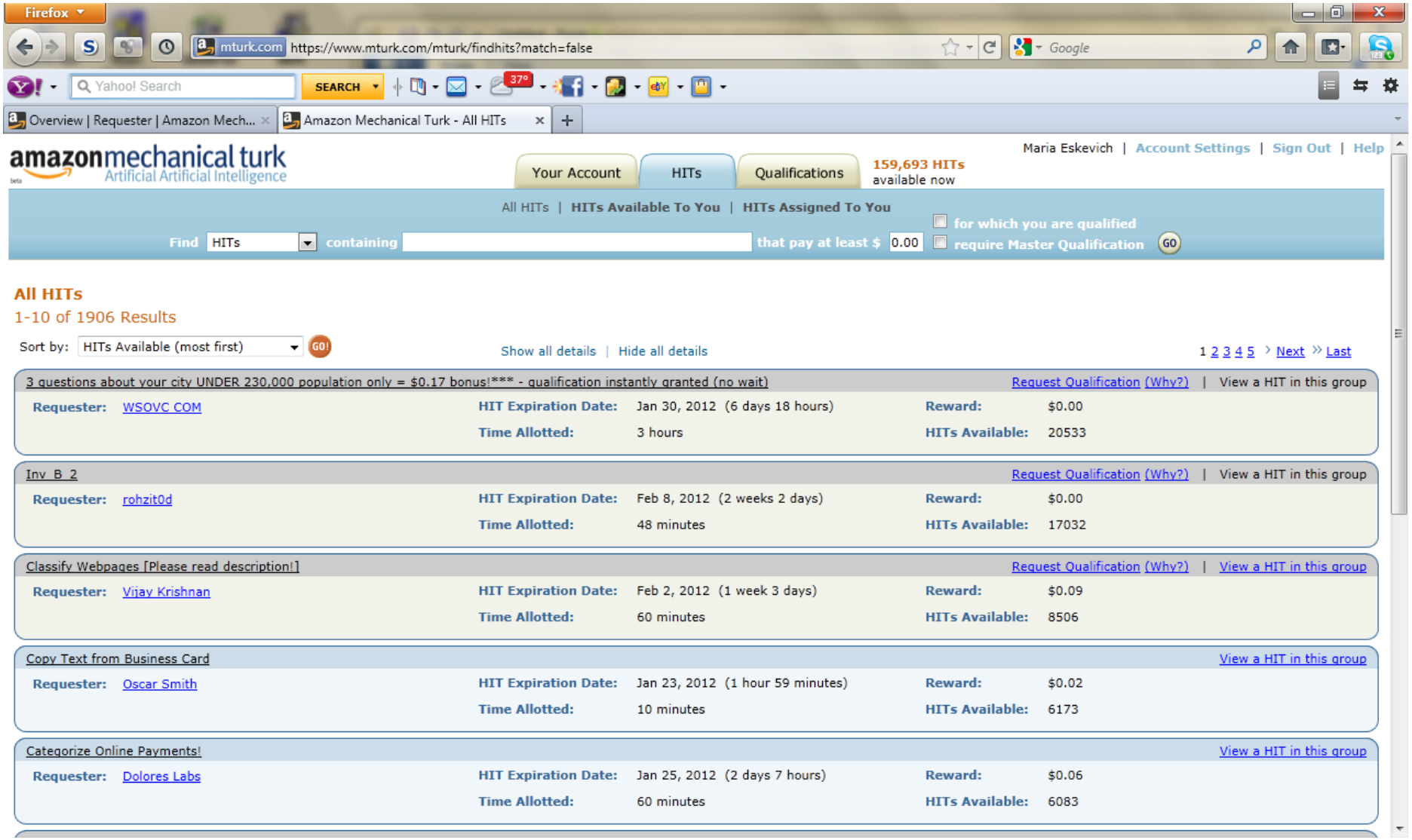
- Workers were allowed to select their own bonus from several available to reflect the quality of their work. Workers were generally found to be honest and good judges of the bonus their work deserved.
- Some problems were encountered due to the need to play an external video:
  - Issues with the worker's browser.
  - Issues with their equipment, e.g. audio playback.
  - Issues with bandwidth required to play the video.

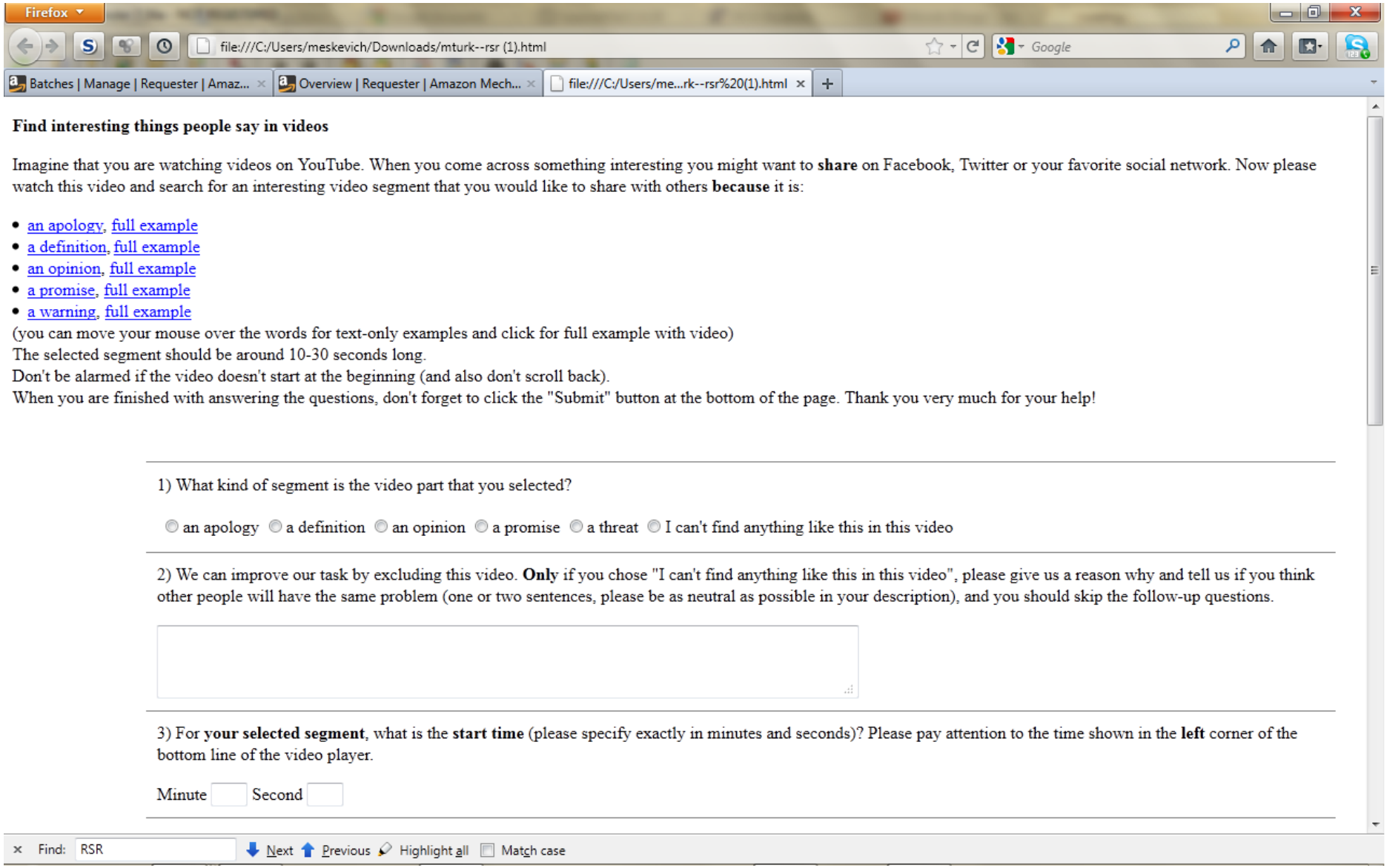
Notes on the MediaEval 2011 RSR HIT

- Workers were allowed to indicate if they were unable to find one of the target speech acts in the video they were given:
  - In which case they did not need to complete the HIT form.
  - Subsequent checking showed that they were generally correct in their judgement.
- Some spamming of the HIT was found. In these cases workers were clearly not attempting to complete the HIT properly, and they were not paid in these cases.
- In some cases workers had completed the HIT, but there work was not found to be useable in the RSR task.

Selecting and Completing a HIT

- When deciding whether to select a HIT, a worker can view the currently available HITs.
- This shows:
  - a brief summary of the HIT,
  - whether there is a qualification requirement,
  - the time allotted for completing it,
  - the available reward for the HIT,
  - no of HITs available for the batch,
  - and for partially completed batches, the payment per hour made to workers so far.
- Requesters may also provide potential workers with a sample HIT so that they know what they will need to do.


Firefox
mturk.com https://www.mturk.com/mturk/findhits?match=false
Google
Yahoo! Search
SEARCH
Overview | Requester | Amazon Mech...
Amazon Mechanical Turk - All HITs
amazonmechanical turk
Artificial Artificial Intelligence
Maria Eskevich | Account Settings | Sign Out | Help
Your Account
HITs
Qualifications
159,693 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
All HITs
1-10 of 1906 Results
Sort by: HITs Available (most first)
GO!
Show all details | Hide all details
1 2 3 4 5 Next Last
3 questions about your city UNDER 230,000 population only = $0.17 bonus!*** - qualification instantly granted (no wait)
Request Qualification (Why?) | View a HIT in this group
Requester: WSOVC COM
HIT Expiration Date: Jan 30, 2012 (6 days 18 hours)
Reward: $0.00
Time Allotted: 3 hours
HITs Available: 20533
Inv_B_2
Request Qualification (Why?) | View a HIT in this group
Requester: rohzit0d
HIT Expiration Date: Feb 8, 2012 (2 weeks 2 days)
Reward: $0.00
Time Allotted: 48 minutes
HITs Available: 17032
Classify Webpages [Please read description!]
Request Qualification (Why?) | View a HIT in this group
Requester: Vijay Krishnan
HIT Expiration Date: Feb 2, 2012 (1 week 3 days)
Reward: $0.09
Time Allotted: 60 minutes
HITs Available: 8506
Copy Text from Business Card
View a HIT in this group
Requester: Oscar Smith
HIT Expiration Date: Jan 23, 2012 (1 hour 59 minutes)
Reward: $0.02
Time Allotted: 10 minutes
HITs Available: 6173
Categorize Online Payments!
View a HIT in this group
Requester: Dolores Labs
HIT Expiration Date: Jan 25, 2012 (2 days 7 hours)
Reward: $0.06
Time Allotted: 60 minutes
HITs Available: 6083



Find interesting things people say in videos

Imagine that you are watching videos on YouTube. When you come across something interesting you might want to **share** on Facebook, Twitter or your favorite social network. Now please watch this video and search for an interesting video segment that you would like to share with others **because** it is:

- an apology, full example
- a definition, full example
- an opinion, full example
- a promise, full example
- a warning, full example

(you can move your mouse over the words for text-only examples and click for full example with video)
The selected segment should be around 10-30 seconds long.
Don't be alarmed if the video doesn't start at the beginning (and also don't scroll back).
When you are finished with answering the questions, don't forget to click the "Submit" button at the bottom of the page. Thank you very much for your help!

---

1) What kind of segment is the video part that you selected?

○ an apology ○ a definition ○ an opinion ○ a promise ○ a threat ○ I can't find anything like this in this video

---

2) We can improve our task by excluding this video. **Only** if you chose "I can't find anything like this in this video", please give us a reason why and tell us if you think other people will have the same problem (one or two sentences, please be as neutral as possible in your description), and you should skip the follow-up questions.

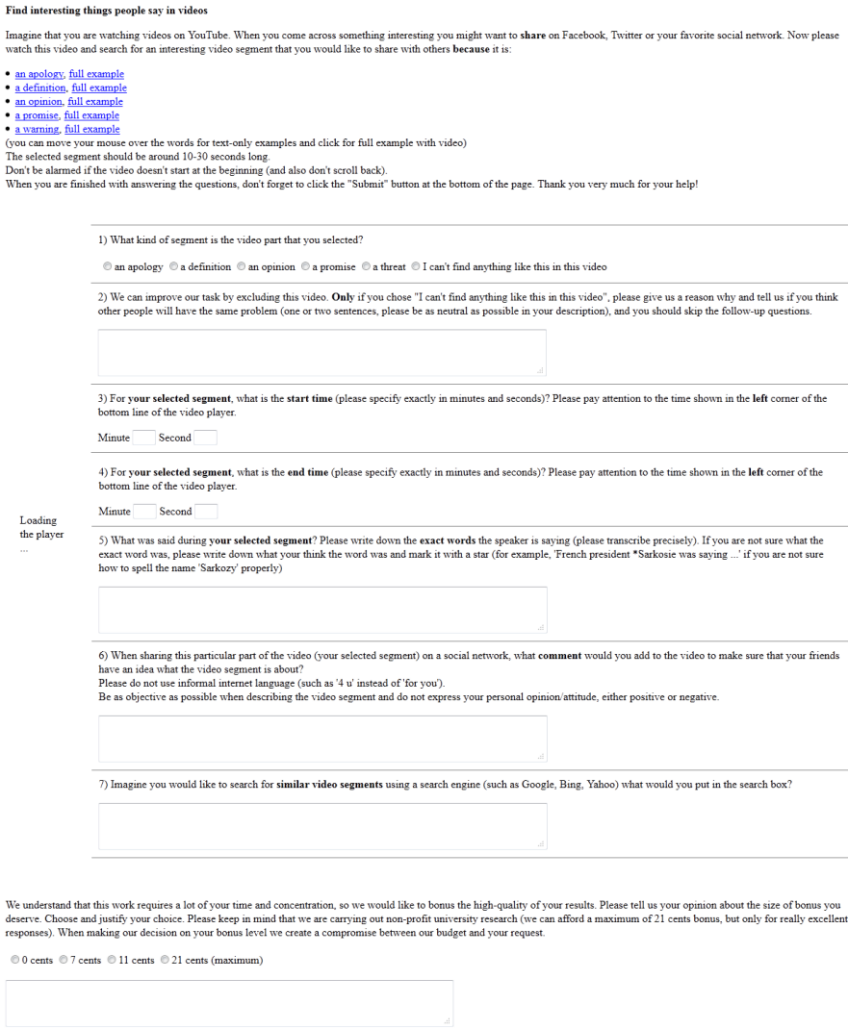
---

3) For **your selected segment**, what is the **start time** (please specify exactly in minutes and seconds)? Please pay attention to the time shown in the **left** corner of the bottom line of the video player.

Minute Second

Display of HIT on a single page

- The HIT is viewed in a single tab.
- The worker scrolls to complete the various fields.



Find interesting things people say in videos

Imagine that you are watching videos on YouTube. When you come across something interesting you might want to **share** on Facebook, Twitter or your favorite social network. Now please watch this video and search for an interesting video segment that you would like to share with others **because** it is:

- an apology, full example
- a definition, full example
- an opinion, full example
- a promise, full example
- a warning, full example

(you can move your mouse over the words for text-only examples and click for full example with video)
The selected segment should be around 10-30 seconds long.
Don't be alarmed if the video doesn't start at the beginning (and also don't scroll back).
When you are finished with answering the questions, don't forget to click the "Submit" button at the bottom of the page. Thank you very much for your help!

Loading the player ...

1) What kind of segment is the video part that you selected?

an apology a definition an opinion a promise a threat I can't find anything like this in this video

2) We can improve our task by excluding this video. **Only** if you chose "I can't find anything like this in this video", please give us a reason why and tell us if you think other people will have the same problem (one or two sentences, please be as neutral as possible in your description), and you should skip the follow-up questions.

3) For **your selected segment**, what is the **start time** (please specify exactly in minutes and seconds)? Please pay attention to the time shown in the **left** corner of the bottom line of the video player.

Minute Second

4) For **your selected segment**, what is the **end time** (please specify exactly in minutes and seconds)? Please pay attention to the time shown in the **left** corner of the bottom line of the video player.

Minute Second

5) What was said during **your selected segment**? Please write down the **exact words** the speaker is saying (please transcribe precisely). If you are not sure what the exact word was, please write down what your think the word was and mark it with a star (for example, 'French president *Sarkosie was saying ...' if you are not sure how to spell the name 'Sarkozy' properly)

6) When sharing this particular part of the video (your selected segment) on a social network, what **comment** would you add to the video to make sure that your friends have an idea what the video segment is about?
Please do not use informal internet language (such as '4 u' instead of 'for you').
Be as objective as possible when describing the video segment and do not express your personal opinion/attitude, either positive or negative.

7) Imagine you would like to search for **similar video segments** using a search engine (such as Google, Bing, Yahoo) what would you put in the search box?

We understand that this work requires a lot of your time and concentration, so we would like to bonus the high-quality of your results. Please tell us your opinion about the size of bonus you deserve. Choose and justify your choice. Please keep in mind that we are carrying out non-profit university research (we can afford a maximum of 21 cents bonus, but only for really excellent responses). When making our decision on your bonus level we create a compromise between our budget and your request.

0 cents 7 cents 11 cents 21 cents (maximum)

References and Resources

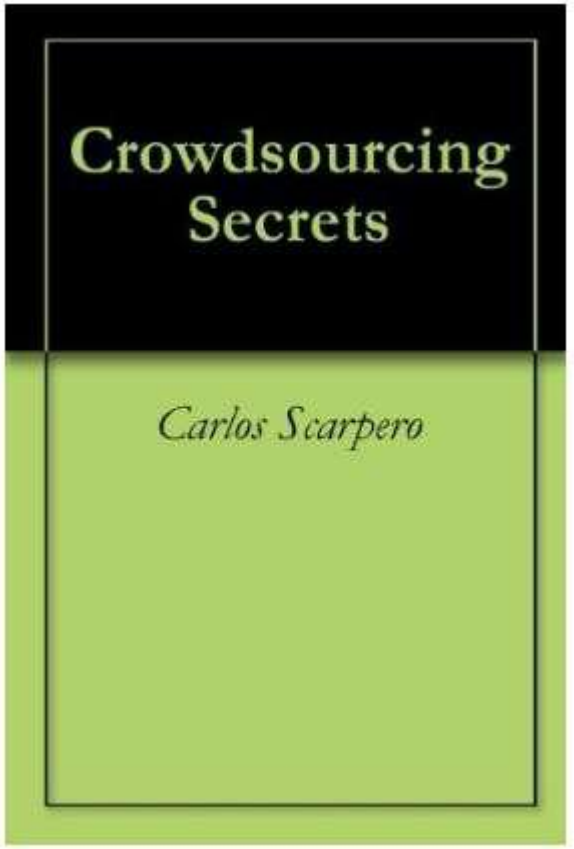
- Anhai Doan, Raghi Ramakrishnan and Alon Y. Halevy, *Crowdsourcing Systems on the Word Wide Web,* Caommunications of the ACM, 45(4): 86-96, 2011.
- Rion Snow, Brendan O'Connor, Daniel Jurafsky, and Andrew Y. Ng. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2008), pages 254–263, Stroudsburg, PA, USA. ACL, 2008.
- Chris Callison-Burch and Mark Dredze. Creating speech and language data with Amazon's Mechanical Turk. In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT 2010), pages 1–12, Stroudsburg, PA, USA. ACL, 2010.
- Keelan Evanini, Derrick Higgins, and Klaus Zechner. Using Amazon Mechanical Turk for transcription of non-native speech. In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT 2010), pages 53–56. ACL, 2010.
- Catherine Grady and Matthew Lease. Crowdsourcing document relevance assessment with Mechanical Turk. In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT 2010), pages 172–179. ACL, 2010.

References and Resources

- Ian Lane, Alex Waibel, Matthias Eck, and Kay Rottmann. Tools for collecting speech corpora via Mechanical-Turk. In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT 2010), pages 184–187. ACL, 2010.
- Matthew Marge, Satanjeev Banerjee, and Alexander I. Rudnicky. Using the Amazon Mechanical Turk for transcription of spoken language. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), pages 5270– 5273. IEEE, 2010.
- Wesley Willett, Jeffrey Heer, Maneesh Agrawala. Strategies for Crowdsourcing Social Data Analysis. Human Factors in Computing Systems (CHI 2012), ACM , 2012
- Martha Larson, Maria Eskevich, Roeland Ordelman, Christoph Kofler, Sebastian Schmiedeke, and Gareth J. F. Jones. Overview of Mediaeval 2011 Rich Speech Retrieval task and genre tagging task. In Proceedings of the MediaEval 2011 Workshop, volume 807. CEUR-WS.org, 2011.
- Maria.Eskevich, Gareth .J.F.Jones, Martha Larson and Roeland.Ordelman, Creating a Data Collection for Evaluating Rich Speech Retrieval, In Proceedings of the Eighth international conference on Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, 2012.

Resources and References

- July 2010, *kindle*-only
- “This book introduce you to the top Crowdsourcing sites and outlines step by step with photos the exact process of get started as a requester on Amazon Mechanical Turk.”



Resources and References

- 2012 book: Omar Alonso, Gabriella Kazai and Stefano Mizzaro, *Crowdsourcing for Search Engine Evaluation: Why and How,* .
- Forthcoming Special Issue on Crowdsourcing of *Information Retrieval Journal*
- Large number of resources available from Matt Lease at: http://ir.ischool.utexas.edu/crowd
  - Slides from conference and workshop tutorials and keynotes.
  - Slides contain lists of recommended readings.

Acknowledgements

- Some of the background material on crowdsourcing for information retrieval draws on presentations by Matt Lease and Omar Alonso, and is used with permission.
- Many thanks to Maria Eskevich for providing the screen shots for the MediaEval 2011 Rich Speech Retrieval task crowdsouring example.

MediaEval 2012

MediaEval

multimedia benchmarking initiative

- www.multimediaeval.org
- Please complete the MediaEval 2012 task questionnaire!