**How to use Amazon Mechanical Turk to create Project**

**Working with Mechanical Turk**

To be able to create HITs and retrieve the results you need a few things:

1. Amazon account.
2. AWS account.
3. Mechanical Turk linked to your AWS account.
4. Amazon balance.
5. AWS key & AWS secret.

**Creating Your Amazon Account**

You possibly have an Amazon account already but if not, simply go to <https://www.amazon.co.uk/> and create one from the home page.

**Other Accounts**

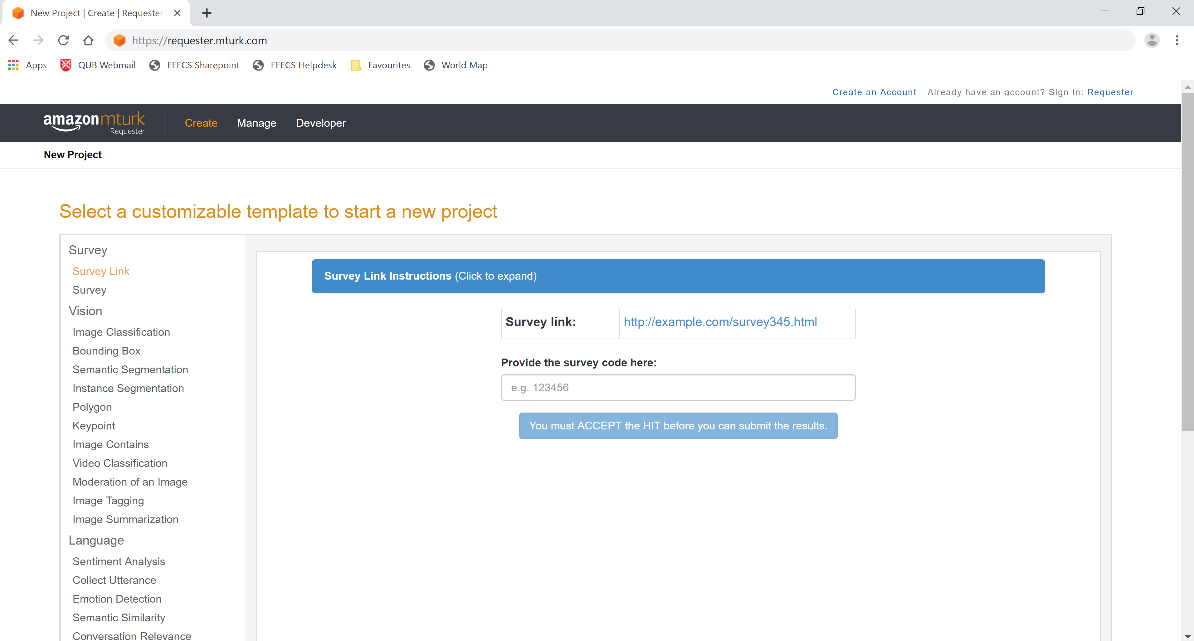
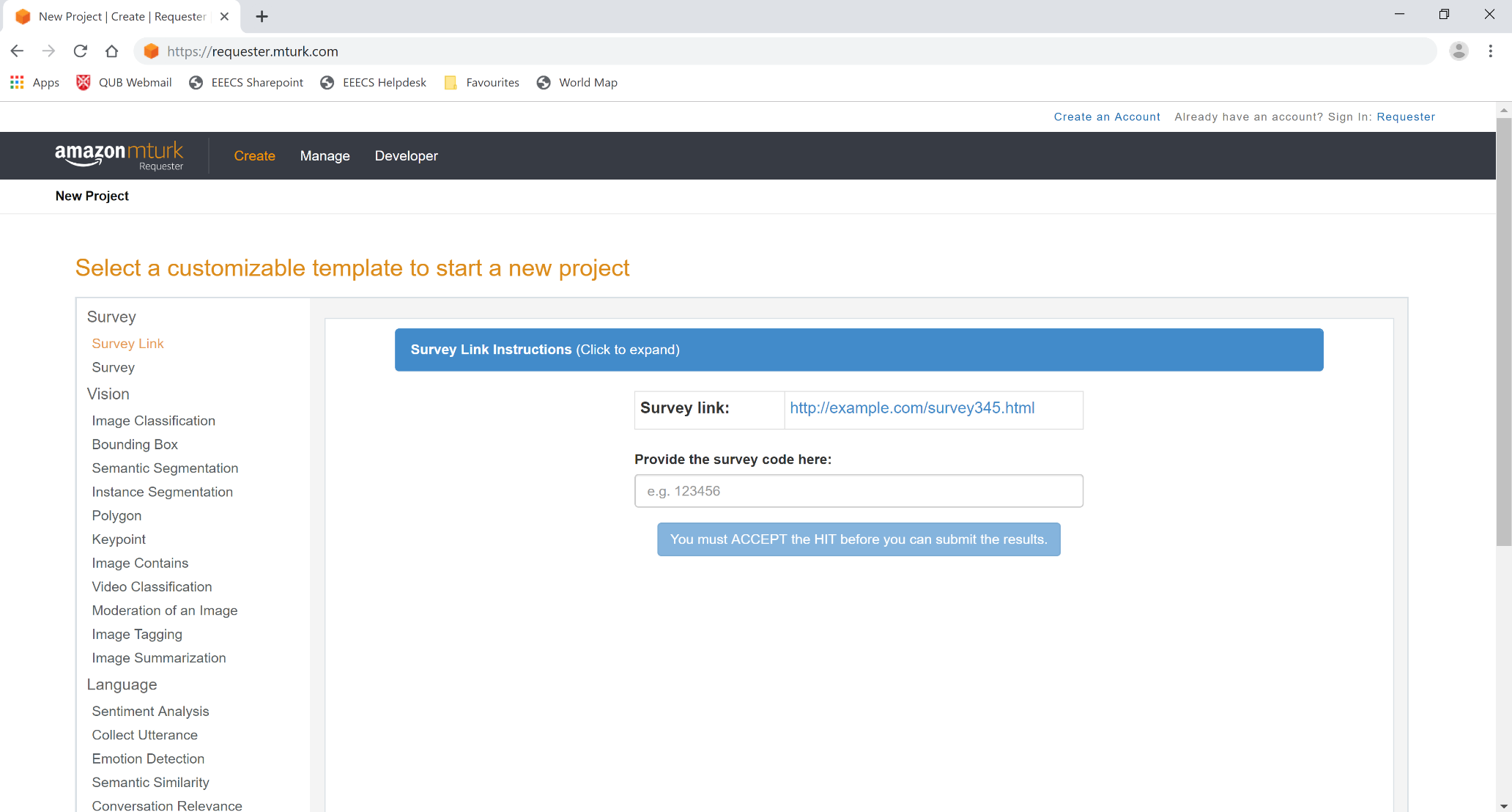
To setup your AWS account, mturk, and mturk sandbox accounts simply follow this guide: <https://docs.aws.amazon.com/AWSMechTurk/latest/AWSMechanicalTurkGettingStartedGuide/SetUp.html>

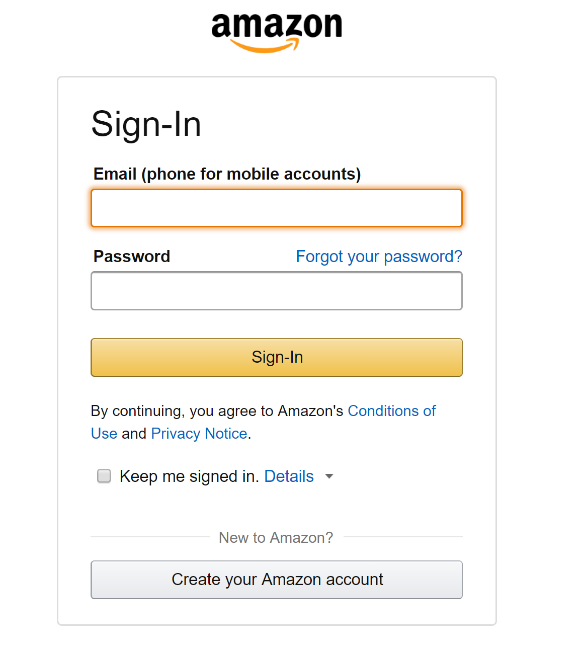
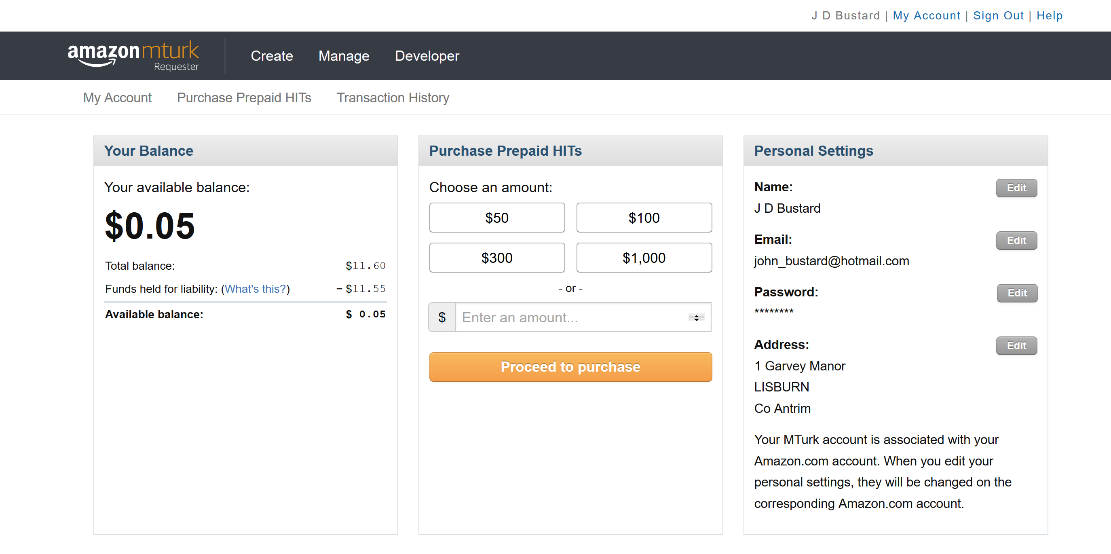
Following steps 1-5 should have you covered.

**IMPORTANT:** Make sure you take note of the Access Key and Secret Key for the user that you create through the AWS interface. These are required to create HITs programmatically.

**Topping Up**

This is not necessary if you only plan on using the sandboxed Mechanical Turk, but if you want to push to the production endpoint, you need to prepay for your HITs.

1. Go to ‘requester.mturk.com’
2. Click on ‘Sign in: Requester’
3. Sign in and navigate to ‘https://requester.mturk.com/account’



* Sign in as a Requester on mturk.com
* Click create – new project
* Follow <https://docs.aws.amazon.com/AWSMechTurk/latest/RequesterUI/CreatingYourBatchofHITs.html>
* Select a customizable template to start a new project:
  + There is a list of different templates you can work off –
    - ***Video Classification*** could be useful for labelling scenes, but you would need to supply them with the URLs instead of them going out and looking for ones
    - ***Image Tagging/Summarization*** could be useful for labelling a character’s physical appearance in an image, i.e. you could get their permanent traits;
    - ***Sentiment Analysis*** could be useful for labelling dialogue from scripts and/or subtitle files to get the emotion behind it. Could also use ***Emotion Detection*** here;
    - ***Intent Detection*** could be useful for labelling actions e.g. fight scenes etc in scripts/scenes to see the motivation behind it and what they hope the outcome will be;
    - ***Semantic Similarity*** could be used to compare scenes/ dialogue to generalise into specific themes;
    - ***Data Collection*** could be one of the most useful for collecting metadata, though you may have to add specific links you want looked at;
    - ***Other*** could be the best selection as it is a simple/basic template which you can alter easily.
  + Then click Create Project
* **On the Edit Project page, you must specify the properties that are common for all of the tasks created using this project**
* **Enter Properties:**
  + *Describe your task to Workers:*
    - Enter a Title – you must be as specific as possible when describing the task to Workers
    - Enter a Description – give more details about the task
    - Enter Keywords – this helps Workers search for your tasks, so the more the better
  + *Setting up your task – this section is about time and money:*
    - How long the Workers can work on the task: maximum time a Worker has to work on a single task – be generous so that Workers don’t feel rushed;
    - How much they are paid: how much they are paid for completing an assignment – consider the difficulty of the task;
    - The number of assignments per task: how many unique Workers do you want to work on each task?
    - How long until the task expires: maximum time your task will be available o Workers on Mechanical Turk;
    - The options for auto-approving and paying workers – amount of time oy have to reject a Worker’s assignment after they submit it.
  + *Workers requirements:*
    - Do you require Masters to do your tasks? – Masters are Workers who have demonstrated excellence across a wide range of tasks – basically you are guaranteed to get reliable workers who give accurate results. However, there is an additional 5% fee when using Masters Workers;
    - You can add additional requirements too (up to 5):
    - There are premium requirements (which you have to pay extra for) and three others which are free:
      * HIT Approval Rate (%) for all Requesters HITs
      * Location
      * Number of HITs Approved
    - As mentioned below, a combination of HIT Approval Rate (e.g. greater than 95%) and Number of HITs Approved (e.g. greater than 5000) has a significant effect on the overall quality of the results.
    - There is also a checkbox for if the project may contain potentially explicit or offensive content, for example, nudity. If this is checked, you must add a content warning to the title of your project, and it must be either Private or Hidden Visibility.
  + *Task Visibility:*
    - Public - All Workers can see and preview my tasks (unsuitable for project containing adult content)
    - Private - All Workers can see my tasks, but only Workers that meet all Qualification requirements can preview my tasks (suitable for project containing adult content)
    - Hidden - Only Workers that meet my Qualification requirements can see and preview my tasks (suitable for project containing adult content)
* When happy with your details, click Save, and then click Design Layout to move to the next step.
* **Design Layout:**
  + Here you can use the HTML editor to design the layout of your task, which Workers will complete along with any HTML, CSS, or JavaScript to customize your layout. Otherwise, you can use Crowd HTML Elements to simply design your layout - <https://blog.mturk.com/tutorial-using-crowd-html-elements-b8990ec71057>.
  + You can also use your own code editor of your choice.
  + This layout is common for all of the tasks created with this project. You can define variables (e.g. ${variable\_name}) for data that will vary from task to task based on your CSV input file (<https://blog.mturk.com/using-csv-files-to-create-multiple-hits-in-the-requester-ui-22a25ec563dc>).
  + If using Crowd-HTML format, you must include the script tag that appears at the top of each project template on Requester;
  + Any JavaScript/CSS library can also be added;
  + Make sure to keep checking the Preview window to ensure your changes are working.
* **Publishing a Batch of HITs:**
  + When you believe your task is ready to be published, I would recommend switching to the [MTurk Requester Sandbox account](https://requestersandbox.mturk.com/" \t "_blank)— this looks identical to the Requester account, but is created in the “Sandbox” which is an environment where you can try publishing tasks without spending any money to ensure they work correctly;
  + This removes the risk of publishing a task that Workers will not be able to access properly for example;
  + To switch over, you can copy and paste the Design Layout and then proceed with *Publish Batch:*
    - A pop-up window will appear asking you to choose your .csv input file for the variables you specified in your project, for example if you had a different video clip for each HIT, you would have added a variable name like $(video\_url) in place of the source link;
    - You can create your own input file, or click the link to download the sample file;
    - All that’s left to do is to add in the links you require for each HIT, and when ready, save;
    - Now you can search for and select the required input file to continue with publishing the batch;
    - When the file has been validated, click *Upload*;
    - Now you can preview the tasks you have created – this will display the project you have created, along with the variables contained in the input file
      * You can click through to see every HIT to ensure the URLs work (NB if you are looking to embed a YouTube video, make sure the link is correct);
      * When happy with the tasks, click *Next*
    - The last step is to confirm and publish the batch
      * This page provides you with the *Batch Properties* you have chosen, including the worker qualifications etc
      * The number of tasks and subsequent cost summary is included here
      * The very last section has the payment method where you prepay for the cost summarised above for the price you pay for each task being submitted
      * After you have added/chosen a payment method, you can click *Purchase & Publish*;
  + Now you have published your first Batch, you can go to the Manage tab and see all batches in progress, along with the results;

Qualification Options:

* Requesters frequently say that the combination of HIT Approval Rate (e.g. greater than 95%) and Number of HITs Approved (e.g. greater than 5000) has a significant effect on the overall quality of the results.
* Premium qualifications have addition fees;
* Can create new qualifications e.g.