

Be a requester on Amazon Mechanical Turk (Part 2)

CS590-CSC (Crowdsourcing and Social Computing), Fall 2018, Purdue University

Background

HTML, CSS and Javascript are three key components of web programming. Requesters who wish to post tasks through crowdsourcing platforms need to first create web pages using these three elements to present their tasks to crowd workers.

On a high level (quote from <https://blog.hubspot.com/marketing/web-design-html-css-javascript>):

- HTML provides the basic structure of sites, which is enhanced and modified by other technologies like CSS and JavaScript.
- CSS is used to control presentation, formatting, and layout.
- JavaScript is used to control the behavior of different elements.

Here is a fun piece to read for better understanding the relationship between HTML, CSS, and Javascript: <https://blog.codeanalogies.com/2018/05/09/the-relationship-between-html-css-and-javascript-explained/>.

HTML tutorials

- HTML 101: <https://websitesetup.org/html-tutorial-beginners/>
- A more detailed HTML Tutorial: <https://www.w3schools.com/html/>

CSS tutorials

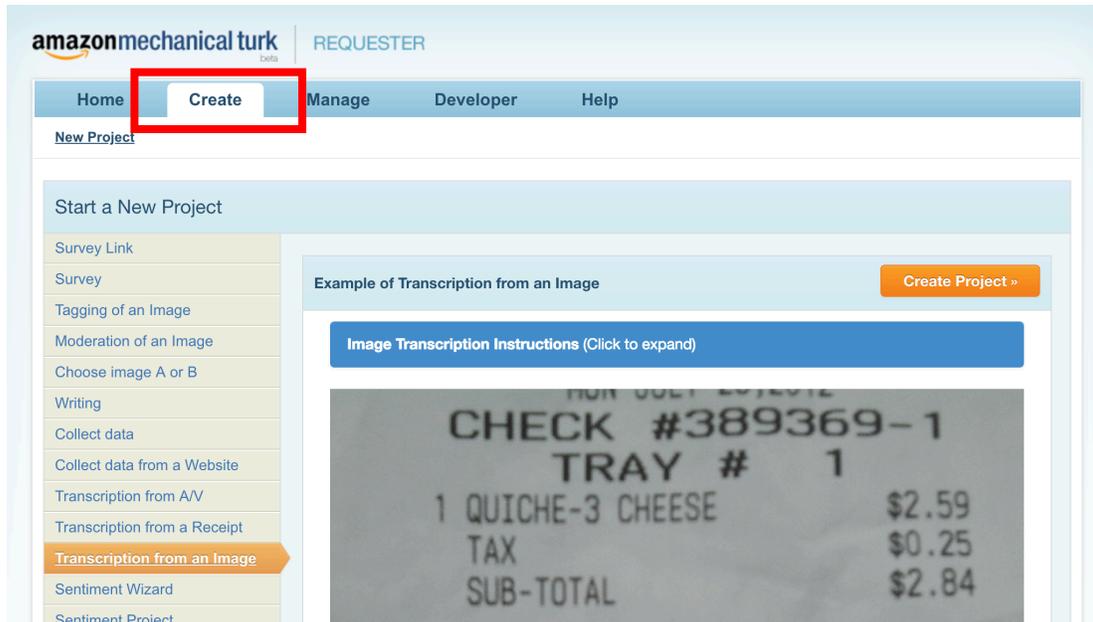
- CSS 101: <https://medium.freecodecamp.org/get-started-with-css-in-5-minutes-e0804813fc3e>
- A more detailed CSS Tutorial: <https://www.w3schools.com/css/>

Javascript tutorials

- Javascript 101: <https://webdesignfromscratch.com/html-css/js101/>;
<https://medium.com/@christinatrung/javascript-101-again-6f5c3de74165>
- A more detailed Javascript Tutorial: <https://www.w3schools.com/js/default.asp>

Publish a HIT using a template

MTurk provides some templates for typical types of tasks. You can find all of these templates on the “Create” page.



Let's create a receipt image annotation HIT (in sandbox) using the template! Here are some receipt image URLs prepared for you in this toy HIT:

- <http://cdn.arn.com.au/media/8341541/subway-receipt-hero.jpg>
- https://sunnymoney.weebly.com/uploads/1/9/6/4/19645963/veggie-grocery-receipt_orig.jpeg
- <https://media-cdn.tripadvisor.com/media/photo-s/0c/ee/40/74/receipt.jpg>

Tutorial on publishing HITs using templates:

https://requestersandbox.mturk.com/help/getting_started.html

- Key point: Understand how to set a variable in the HIT template using the syntax of `${VARIABLE_NAME}`

Publish a HIT through an API

Other than publishing a HIT through the requester website using existing templates, you can also publish a HIT using the API provided by MTurk.

A few tutorials on how to get started with API:

- MTurk API tutorial blog: <https://blog.mturk.com/tutorials/home>
- Get started with Java API: <https://blog.mturk.com/tutorial-using-the-mturk-api-with-the-aws-sdk-for-java-48b4826f9122>; <https://blog.mturk.com/tutorial-publishing-hits-with-mturk-and-the-java-programming-language-5db99eda0f2a>

- Get started with Python API: <https://blog.mturk.com/tutorial-a-beginners-guide-to-crowdsourcing-ml-training-data-with-python-and-mturk-d8df4bdf2977>

If you want to design your tasks using HTML forms without having to host it on your own website, you will need to use the HTMLQuestion data structure:

- https://docs.aws.amazon.com/AWSMechTurk/latest/AWSMturkAPI/ApiReference_HTMLQuestionArticle.html

But you can also host your tasks on your own website. In this case, you will need to use the ExternalQuestion data structure:

- External Question:
https://docs.aws.amazon.com/AWSMechTurk/latest/AWSMturkAPI/ApiReference_ExternalQuestionArticle.html
- The external question schema URL (i.e., the value of xmlns in the XML file) is:
<http://mechanicalturk.amazonaws.com/AWSMechanicalTurkDataSchemas/2006-07-14/ExternalQuestion.xsd>