Annotation Guidelines v 0.0.9 – Emy

Hello team,

WHAT: Our objective is to describe the everyday products we can see in the photos provided in the .   
  
In each photo, you will see an object (or objects) with a box drawn over it, this is **the mystery object**.   
  
Our objective is to describe **the mystery** **object** in as much detail as possible, using only what we can observe in the photo.

WHY: The combined image & text data inform our where2BuyIt feature, so our users around the world can find where to buy nearly anything, just by snapping a pic with their phone.

HOW:

Identify the **mystery object**(s) in the photo, and describe these in detail into the form

* **What is it?** 
  + What the object made to do? What is the common name we would call it?
    - For example: table, tray, stand
* **How does it look**?
  + Can you describe the color? Pattern? Style?
  + Are some parts blue and others yellow?
  + If the **brand** is visible, please do record that.
* Please describe **only what can be observed** in the photo.

*Main assumptions*:

* We are using text annotations to inform computer vision classification models.
* The unlabeled dataset is already curated to be relevant to the annotation task.

*Questions for the team:*

* What are the main challenges in getting quality annotations?
* What are the advantages of contracting a team to work specifically on annotation project vs. [an annotation service vendor.](https://becominghuman.ai/10-best-text-annotation-services-and-tools-e9a12fbb8172)

*How long until 1 million annotations?*

**Annotations per hour** – this is expected to be a pretty quick writeup, but there may be a high variance of output based on process experience, individual ability+motivation+momentum, shift in expectations/culture of the team.

* Low range: 6 – one annotation every ten minutes
* High range: 12 – one annotation every five minutes

**Work hours per day & week**

* Nominally we expect 8 work hours. Demanding more time per day for menial tasks will lead to decreased quality, morale, increased turnover. In reality, we have to expect some lost output due to everyday human circumstances, so my estimate gives 7 hours per worker per day.

My ballpark formula is :

1,000,000 annotations / ( 35 *productive* *hours per week per contributor* \* 100 *contributors* \* *annotations per hour*)  
That gives us a ballpark range of 24-48 weeks to 1,000,000 annotations, depending on average annotations per hour.  
Perhaps we want to get multiple annotations (from different annotators) per picture for maximum coverage of details & fail-over. Perhaps we would find a few team members provide much better quality of work vs. others and would decide to prefer their annotations at a later point in the process. This strategy would slow the output but may lead to better quality.

*Ensuring Quality:*

Collect & curate BEST & WORST annotations as documentation for the Annotation Team, and request Annotation Team to join recurring (monthly or biweekly) reviews of BEST & WORST annotations (anonymous authorship) so that team can shift towards best output.