A Fun and Engaging Interface for Crowdsourcing Named Entities

Kara Greenfield, Kelsey Chan, Joseph P. Campbell

4 Oct 2016

This work is sponsored by the Department of the [Air Force and/or other appropriate department(s)] under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the United States Government.
Named Entity Recognition (NER)

- Automatically extract named mentions of entities from natural language text
- Ontology of named entity types
  - Person
  - Organization
  - Location
Don’t NER Corpora Already Exist?

- Domain adaptation is still a huge problem
  - New genre
  - New language
  - New ontology

![Graph showing F1 score vs. Number of training documents](image)
Standard NER Annotation Tools

- Applicable to many tasks, not just NER
- Require linguistic knowledge
- Designed for experts

BRAT Rapid Annotation Tool
(Stenetorp, et al., 2012)

Callisto
(MITRE, 2013)
Early NER Crowdsourcing Interfaces

- Easy for the researcher to create human intelligence tasks
- Difficult for the worker to read the passage
- Limited to short spans of text
- Requires the worker to annotate multiple entity types simultaneously

Twitter NER Annotation System

(Finin et al, 2010)
Early NER Crowdsourcing Interfaces

- Allows worker to annotate an entire entity mention at once instead of word by word
- Presents document in a more natural way
- Interface helps workers to distinguish named mentions from nominals at the expense of slightly more work

Span-based NER Annotation

(Lawson et al, 2010)
• Make it easy for the worker to accurately complete the task
• Recognize when the worker didn’t accurately complete the task
• Make the worker want to complete additional HITs
• Make it easy for the worker to accurately complete the task
  • Recognize when the worker didn’t accurately complete the task
  • Make the worker want to complete additional HITs
Help the Workers

• Minimize the effort that a worker must go through to annotate a document
• Minimize the mental burden on a worker
Clear Instructions Matter

• The use of a small set of pilot experiments to tune the examples included in the instructions greatly eliminated requests for clarification in further runs

• Quickly responding to worker requests for clarification resulted in workers completing tasks accurately and completing more HITs
MITLL NER Annotation Interface

• Make it easy for the worker to accurately complete the task
• Recognize when the worker didn’t accurately complete the task
• Make the worker want to complete additional HITs
De-incentivizing Spam Workers

- Workers who annotate too little
  - **Financial**: incentive based pay structure where the monetary reward increases with the number of annotations (Lawson et. Al. 2010)
  - **Psychological**:
    - Asking workers who didn’t annotate anything on a HIT if they were sure there were no entity mentions
    - Threats of lowered approval ratings

- Workers who annotate too much
  - **Psychological**:
    - Threats of lowered approval ratings
Identifying Worker Fatigue

- Compared the responses of a worker who was less likely to be fatigued with one who was more likely to be fatigued
- Document sizes based on preliminary experimentation with time workers took to complete a HIT and not ending documents mid-sentence
Identifying Worker Fatigue

• Compare workers who may be fatigued to automated system output using MITIE (MIT Information Extraction system)

• Used a third worker to verify that system false positives (as compared to the first 2 workers) were in fact false positives
MITLL NER Annotation Interface

• Make it easy for the worker to accurately complete the task
• Recognize when the worker didn’t accurately complete the task
• Make the worker want to complete additional HITs
Easy to Use Interface

• An easy to use interface can overpower a lack of strong financial incentives

Morning :) Just some friendly advice :) I have done about 140 of your hits. I really like the names ones. I am guessing your account is a new, based on the # of reviews it has on the workers Turkopticon sight. I also noticed that it seems like your batches are not really being worked as fast as you likely hope, and I wanted to offer some advice on that. Though I really enjoy your hits (and the interface I must say is really fantastic! Kudos!), the pay does leave something to be desired.
Conclusions

• Strongest motivator seems to be an implicit threat of lowered approval ratings

• Most workers with high approval ratings didn’t suddenly switch to deliberately performing poorly on our task, but did sometimes make honest mistakes due to fatigue or lack of comprehension

• An easy to use interface can be a stronger motivational factor than financial incentives
Future Work

• More rigorously determine the relationship between financial incentives and ease of use incentives

• Develop similar easy to use interfaces for more complicated annotation tasks
References

• Finin, T. et al., 2010. Annotating Named Entities in Twitter Data with Crowdsourcing. s.l., s.n., pp. 80-88.

• Geyer, K., Greenfield, K., Mensch, A. & Simek, O., 2016. Named Entity Recognition in 140 Characters or Less. s.l., s.n.

• King, D., n.d. MITLL/MITIE. [Online] Available at: https://github.com/mit-nlp/MITIE


