Hybrid Human-Machine Information Systems

Gianluca Demartini
gianlucadademartini.net
Gianluca Demartini

- B.Sc., M.Sc. at U. of Udine, Italy
- Ph.D. at U. of Hannover, Germany
  - Entity Retrieval
- Worked at the eXascale Infolab U. Fribourg (Switzerland), UC Berkeley (on Crowdsourcing), Yahoo! (Spain), L3S Research Center (Germany)
- Senior Lecturer in Data Science at the iSchool, U. of Sheffield

www.gianlucademartini.net
Research Interests

- **Entity-centric Information Access** (2005-now)
  - Structured/Unstructured data (SIGIR 12), TRank (ISWC 13, WSemJ 16)
  - NER in Scientific Docs (WWW 14), Prepositions (CIKM 14)
  - IR Evaluation (ECIR 16 Best Paper Award, IRJ 2015)

- **Hybrid Human-Machine Systems** (2012-now)
  - ZenCrowd (WWW 12, VLDBJ), CrowdQ (CIDR 13)
  - Human Memory based Systems (WWW 14, PVLDB)
  - Hybrid systems overview (COMNET, 2015)

- **Better Crowdsourcing Platforms** (2013-now)
  - Platform Dynamics (WWW 15)
  - Pick-a-Crowd (WWW 13), Malicious Workers (CHI 15)
  - Scale-up Crowdsourcing (HCOMP 14), Scheduling (WWW 16)

Sponsored by: H2020, EPSRC, U Sheffield, Amazon
Hybrid Human-Machine Info Systems

• Use of **paid micro-task crowdsourcing** (i.e., Amazon MTurk)

• To build information systems that leverage
  – the power of machines to **scale over large amounts of data**
  – the ability of **human intelligence** to keep the quality of data processing high
5-year Analysis of MTurk workload

- Mturk-tracker.com
  - Collects metadata about each visible batch (Title, description, rewards, required qualifications, HITs available, requester, etc), that is, set of similar tasks or HITs
  - Records batch progress (every ~20 minutes)
  - Covers 130M tasks
  - 2009-2014

HIT Pricing

5-cents is the new 1-cent
#mturk dynamics

- HIT reward has increased over time
- **Audio transcription** is the most popular task
- Demand for Indian workers has decreased
- **Surveys** are most popular for US workers
- 1000 new requesters per month join
- 10K new HITs arrive and 7.5K HITs get completed every hour

- Check #mturk dynamics for the main findings
Facebook is not waiting for its initial public offering to make its first big purchase.

In its largest acquisition to date, the social network has purchased Instagram, the popular photo-sharing application, for about $1 billion in cash and stock, the company said Monday.
ZenCrowd

• Combine both algorithmic and manual linking
• Automate manual linking via crowdsourcing
• Dynamically assess human workers with a probabilistic reasoning framework
ZenCrowd Architecture

Lessons Learnt

• Crowdsourcing + Prob reasoning works!
• But
  – Different worker communities perform differently
  – Many low quality workers
  – Completion time may vary (based on reward)
• Need to find the right workers for your task (see WWW2013 and CHI2015 papers)
• Need to make sure high priority tasks are completed fast (see WWW2016 paper)
Behavioral Patterns of Malicious Workers

**Instruction**: Please attempt this microtask ONLY IF you have successfully completed 5 microtasks previously.

Response: ‘this is my first task’

eg: Copy-pasting same text in response to multiple questions, entering gibberish, etc.

Response: ‘What’s your task?’ , ‘adasd’, ‘fggfg gsd lilkj’

**Instruction**: Identify 5 keywords that represent this task (separated by commas).

Response: ‘survey, tasks, history’, ‘previous task yellow’

**Instruction**: Identify 5 keywords that represent this task (separated by commas).

Response: ‘one, two, three, four, five’

These workers abide by the instructions and provide valid responses, but stumble at the gold-standard questions!

Scheduling HITs

- Fair Scheduling
  - Priority of HITs but avoid starvation
  - Assign HITs of the same type (no context switch)

Happy to help

• Take away: “Crowdsourcing works but…”

• I have used crowdsourcing (i.e., Amazon MTurk) over the last 6 years

• And worked on improving its effectiveness/efficiency over the last 4 years

• If you have tried / want to try it but
  – Find that you get very bad quality data back
  – Don’t know how to best design the task / reward / instructions

• Let me know!

  gianlucaademartini.net