

JOSEPH CHEE CHANG

7 Allegheny Center #1012 Pittsburgh PA 15212 | 412-980-8551 | josephcc.cmu@gmail.com | josephcc@cmu.edu | http://joe.cat

OVERVIEW

- Ph.D. research focused on 1) combining crowds and ML to help requesters synthesize and learn from data; and 2) designing novel information systems and interactions to better support sensemaking using both qualitative and quantitative approaches.
- Ph.D. work led to full-paper publications at CHI (5), UIST (2), CSCW (1) and IUI (1), including 3 Best Paper Honorable Mentions Awards (<5%) from the top HCI Conference and was invited to give an encore talk at the AAAI HCOMP conference.
- Knowledgeable in machine learning and natural language processing.
- Strong development skills with an emphasis on WebUIs.

EDUCATION

Ph.D. in Language Technologies (2020)

Focus: HCI, Crowdsourcing, Information Systems, Sensemaking. CHI best paper honorable mentions x3

Carnegie Mellon University, United States

Advisor: Aniket Kittur

M.S. in Computer Science, Information Systems (2012)

Focus: Natural Language Processing, Information Extraction. Thesis work published at ACL 2012.

National Tsing Hua University, Taiwan

Advisor: Roger Jang

B.S. in Computer Science (2010)

Focus: IC Electronic Design Automation. Won in a national algorithm design and implementation competition (3000+ lines of C++)

Yuan Ze University, Taiwan

SELECTED RESEARCH PROJECTS

Alloy: Crowdsourced open-domain non-factual question-answering [1,2] 2 Best Paper Honorable Mentions in ACM CHI'16

- **Problem:** Answering complex questions such as "How do I grow better tomatoes?" often requires individuals to conduct extensive online research. Can we crowdsource this process using ML and 100 crowdworkers conducting microtasks distributedly?
- **Approach:** A crowd-powered QA system that extracts and synthesizes web clips across webpages into coherent articles.
- **Results 1:** Combining crowds and ML techniques, the system's short text clustering algorithm outperformed existing computation- and crowd-based approaches, and reduced monetary cost by 70% when compared to an existing crowd-based system.
- **Results 2:** The final output articles compared favorably to highly curated and reputable sources from top Google search results.
- Collaborated with an industry partner (Bosch) to support their real-world deployment.

Revolt: Structuring unknown unknowns while creating training labels with the crowd [3]

published in ACM CHI'17

- **Problem:** It can be costly for requesters to create labeling guidelines that covers the long-tail of edge cases in complex datasets, leading to inconsistent labels from the crowdworkers. How can we use crowdsourcing to also scale the complexity of data?
- **Approach:** Crowdworkers find and categorize items not covered by guidelines as they label for post-hoc decisions by requesters.
- **Results:** Accurate and consistent labels without the needs for creating comprehensive guidelines beforehand.

Mesh: Scaffolding comparison tables for online product research [4]

cond. accepted, UIST'20 (scores: 4.5, 4.5, 4.5, 4.0, 4.0/5.0)

- **Problem:** Consumers can choose from many options based on thousands of reviews, but making confident decisions can be hard because online information can be subjective, biased, and scattered across sources incurring high cognitive & interaction costs.
- **Approach:** A comparison table builder that scaffolds users' research by capturing their decisions and interpretation of data.
- **Results:** Lab and field deployment studies showed the system can lower interaction costs and that participants generated learning summaries rated as more insightful and confident than participants who used Google Sheets as a baseline condition.

SELECTED PUBLICATIONS

1. **Joseph Chee Chang**, Aniket Kittur, and Nathan Hahn. 2016. Alloy: Clustering with Crowds and Computation. ACM CHI Conference (Best Paper Honorable Mention). 12 pages. doi.org/10.1145/2858036.2858411
2. Nathan Hahn, **Joseph Chee Chang**, Ji Eun Kim, and Aniket Kittur. 2016. Knowledge Accelerator: Big Picture Thinking in Small Pieces. ACM CHI Conference (Best Paper Honorable Mention). 12 pages. dl.acm.org/doi/10.1145/2858036.2858364
3. **Joseph Chee Chang**, Saleema Amershi, and Ece Kamar. 2017. Revolt: Collaborative Crowdsourcing for Labeling Machine Learning Datasets. ACM CHI Conference. 13 pages. doi.org/10.1145/3025453.3026044
4. **Joseph Chee Chang**, Nathan Hahn, and Aniket Kittur. 2020. Mesh: Scaffolding Comparison Tables for Online Decision Making. (to appear) ACM UIST Conference. 14 pages. joe.cat/mesh.pdf (preprint)
5. Joel Chan, **Joseph Chee Chang**, Tom Hope, Dafna Shahaf, and Aniket Kittur. 2018. Solvent: A Mixed Initiative System for Finding Analogies between Research Papers. ACM CSCW Conference. 21 pages. doi.org/10.1145/3274300

WORK EXPERIENCE

Ph.D. Research Intern @ Machine Teaching Group

Worked on crowdsourcing and machine learning [3]. Mentors: Saleema Amershi and Ece Kamar.

Microsoft Research

May-August 2016

Full-time Search Engineer

Worked on the Knowledge Graph and search log analysis using Hadoop.

Yahoo!

April-July 2013

Skill Keywords: [Web Development] Modern JavaScript, **TypeScript***, **ReactJS***, **D3***, **Firebase***, SQL [Stats/ML/Data] **Python***/Scipy, R, Hadoop Map-Reduce [Others]

Browser Extensions*, iOS, Android, Java, Obj-C, C. *proficient/strong expertise

Status: Permanent resident, United States.