

# Matt Lease

Associate Professor  
Information & Computer Science

<http://ir.ischool.utexas.edu>

[ml@utexas.edu](mailto:ml@utexas.edu)



**Early Career Awards:** DARPA Young Faculty, NSF CAREER, IMLS Early Career

**Additional Past Support:** ONR, Qatar QNRF, Portugal FCT, Yahoo!, Google, Amazon



UT Austin Information Retrieval & Crowdsourcing Lab, Spring 2018

# Research Areas



- **Information Retrieval & Search**
  - Neural Information Retrieval: At the End of the Early Years. *IRJ* 2018
  - Efficient Test Collection Construction via Active Learning. arXiv 2018.
  - ArabicWeb16: A New Crawl for Today's Arabic Web. ACM SIGIR 2016.
  - Active Learning to Maximize Accuracy vs. Effort in Interactive Information Retrieval. ACM SIGIR 2011.
- **Human Computation & Crowdsourcing** (e.g., Human-in-the-loop)
  - Why Is That Relevant? Collecting Annotator Rationales for Relevance Judgments. *AAAI HCOMP* 2016.
    - *Best Paper Award*
  - Probabilistic Modeling for Crowdsourcing Partially-Subjective Ratings. *AAAI HCOMP* 2016.
  - Crowdsourcing Information Extraction for Biomedical Systematic Reviews. *AAAI HCOMP* 2016.
  - Combining Crowd and Expert Labels using Decision Theoretic Active Learning. *AAAI HCOMP* 2015.
- **Natural Language Processing**
  - An Interpretable Joint Graphical Model for Fact-Checking from Crowds. *AAAI* 2018.
  - Aggregating and Predicting Sequence Labels from Crowd Annotations. *ACL* 2017.
  - Exploiting Domain Knowledge via Grouped Weight Sharing [for] Text Categorization. *ACL* 2017.
  - Active Discriminative Text Representation Learning. *AAAI* 2017.

# Goals & Skills



- **Research Goals: Information Retrieval**
  - i) improve back-end, core search algorithms
  - ii) enhance human-centered, interactive search
  - iii) enable new forms of search
- **Research Goals: Human Computation**
  - i) optimize human-in-the-loop interactive computing;
  - ii) expand crowdsourcing's reach to tackle more ambitious problems
  - iii) optimize crowdsourced data collection (e.g., quality, cost, and speed)
- **My graduate training spanned IR & NLP** across 6 universities
  - IR: U. Mass. Amherst, with James Allan & Bruce Croft
  - NLP: Brown University, with Eugene Charniak & Mark Johnson
  - Lab Rotations: U. Washington, Johns Hopkins U., Charles U. (Prague), Saarland U. (Germany)
- **My expertise is diverse, spanning 1) IR, HCOMP, & NLP areas; and 2) AI (back-end) & HCI (front-end)**
- The Texas Advanced Computing Center (TACC) provides **unparalleled distributed computing and storage capabilities** to UT Austin faculty & students

# Combining AI & HCI



- IR and HCOMP both *place people at the center of computing*
  - information seekers in IR and online contributors in HCOMP
- I approach both areas via *hybrid, social computing*
  - I orchestrate *man-machine partnerships* which creatively **blend front-end human-computer interaction (HCI) design with back-end artificial intelligence (AI)** modeling of people and their tasks.
- By capitalizing on the strengths of each party - man and machine - I seek to compensate for the other's limitations to create a whole greater than its parts.
  - IR systems can utilize front-end HCI design to empower searcher intuition and creativity, while back-end AI algorithms interpret ambiguous human queries, sift through vast information, and suggest potentially relevant results.
  - In HCOMP, front-end HCI design can enable workers to more easily understand and complete tasks, while back-end AI modeling of workers and tasks enables principled optimization of data collection.
- As with Amdahl's Law, optimizing only part of a joint system is limiting. By tackling both front-end HCI and back-end AI, we maximize overall effectiveness.

# Contact Information



- Matt Lease
- Associate Professor
- Information and CS, U. of Texas at Austin
- **Email:** [ml@utexas.edu](mailto:ml@utexas.edu)
- **Phone:** (512) 471-9350
- **Lab webpage:** <http://ir.ischool.utexas.edu>
- **Linked In:** <https://www.linkedin.com/in/mattlease/>
- **Twitter:** [@mattlease](https://twitter.com/mattlease)
- **Slides online:** <https://www.slideshare.net/mattlease>