Text REtrieval Conference (TREC)
https://trec.nist.gov
Community evaluations focused on search and related technologies

- Pioneered use of “pooling” for building large collections
- Built ~200 test collections for dozens of search tasks
- Hundreds of participant teams world-wide
- Premiere venue for determining research methodology
- Model for other efforts in IR and related fields
Answers

- **Live QA (2015-2017)**
  - systems produced answers for Yahoo Answers q’s in real time
  - NIST assessors annotated answer quality

- **Conversational Assistance (2019-)**
  - systems answer series of q’s that model info-seeking dialogues with digital assistants
  - progressively more realistic series in amount of anaphora, recovery from earlier errors, etc.

Machine Learning

- **Deep Learning (2019-)**
  - focus on methods’ effectiveness when large amounts of training data available
  - ties to MS-MARCO benchmark

- **TREC-COVID (2020)**
  - off-cycle effort to build pandemic collection
  - highlighted differences between traditional & neural methods when varying amounts of relevance information available

Not Just Accuracy

- **Fair Ranking (2019-)**
  - recognize multiple stakeholders affected by document rankings
  - explicitly optimize fairness for different parties across set of rankings

- **Health Misinformation (2019-)**
  - penalize wrong/deceptive on-topic information more severely than off-topic
  - consumer health domain has authoritative Truth and ample misinformation for testing