

Summary of TR Evaluation

- Extremely important!
 - TR is an empirically defined problem
 - Inappropriate experiment design misguides research and applications
 - Make sure to get it right for your research or application
- Cranfield evaluation methodology is the main paradigm
 - MAP and nDCG: appropriate for comparing ranking algorithms
 - Precision@10docs is easier to interpret from a user's perspective
- Not covered
 - A-B Test [Sanderson 10]
 - User studies [Kelly 09]