

Formulate Research Hypotheses

- Typical hypotheses in IR:
 - Hypothesis about user characteristics (tested with user studies or user-log analysis, e.g., clickthrough bias)
 - Hypothesis about data characteristics (tested with fitting actual data, e.g., Zipf's law)
 - Hypothesis about methods (tested with experiments):
 - Method A works (or doesn't work) for task B under condition C by measure D (feasibility)
 - Method A performs better than method A' for task B under condition C by measure D (comparative)
 - Introduce baselines naturally lead to hypotheses
- Carefully study existing literature to figure out where exactly you can make a new contribution (what do you want others to cite your work as?)
- The more specialized a hypothesis is, the more likely it's new, but a narrow hypothesis has lower impact than a general one, so try to generalize as much as you can to increase impact
- But avoid over-generalize (must be supported by your experiments)
- Tuning hypotheses