

# The Cranfield Evaluation Methodology

- A methodology for laboratory testing of system components developed in 1960s
- Idea: Build reusable test collections & define measures
  - A sample collection of documents (simulate real document collection)
  - A sample set of queries/topics (simulate user queries)
  - Relevance judgments (ideally made by users who formulated the queries) → Ideal ranked list
  - Measures to quantify how well a system's result matches the ideal ranked list
- A test collection can then be reused many times to compare different systems