Harnessing the Power of Machine Learning and Artificial Intelligence (AI) to Change the Way We Approach Large, Complex Research Questions

Station 2: AI-Enabled DeDuplication

Part of literature review

Search Development  Deduplication  Title/Abstract Screening  Full Text Screening  Data Extraction/Quality Assessment

Key Concepts

• **Two-Phase approach** to identify duplicates in a set of search results.
• Includes **AI-enabled** step to identify likely sets of duplicates that require manual confirmation.

How it Works

• Phase 1: Maximum Precision – New fields added to .RIS record for starting page and first author last name. Provides enhanced ability to identify exact duplicates that do not require manual confirmation.
• Phase 2: Maximum Recall – Machine learning predicts potential duplicates that are tagged for manual review in reference manager of choice.

Use Case(s)

• Remove duplicates from large set of citations from a multi-database search after search completed before search results screened for relevance.

Similar Tools or Approaches

Reference Managers | Some SR Screening Tools | Bramer method