Annotative Indexing

Charles L.A. Clarke

University of Waterloo, Canada

Abstract

This talk presents and explores annotative indexing, a novel framework that unifies and generalizes traditional inverted indices, column stores, object stores, and graph databases. As a result, annotative indexing can provide the underlying indexing framework for retrieval systems that integrate sparse retrieval, dense retrieval, entity retrieval, knowledge graphs, and semi-structured data. While our reference implementation primarily supports human language data in the form of text, annotative indexing is sufficiently general to support a wide range of other data types. The talk will include examples of SQL-like queries over a JSON store built on our reference implementation that include numbers and dates. Taking advantage of the flexibility of annotative indexing, the talk will also demonstrate a fully dynamic inverted index incorporating support for ACID properties of transactions with hundreds of multiple concurrent readers and writers.