Chapter 9
A Data Mining Algorithm for Accessing Research Literature in Electronic Databases: Boolean Operators

Valentine Joseph Owan
https://orcid.org/0000-0001-5715-3428
University of Calabar, Nigeria & Ultimate Research Network (URN), Nigeria

ABSTRACT

Searching and retrieving relevant research materials from electronic databases are difficult for many students and early career researchers. Many researchers have abandoned beneficial research projects because they believe that related literature is unavailable to ground their work. This chapter serves as a guide to students, professionals, and internet users on how to pull information from electronic databases easily. The chapter begins by clarifying the concept of electronic databases, the evolution of electronic databases, and the processes involved in indexing scholarly works in an electronic database. The advantages and disadvantages associated with the use of electronic databases are also discussed. The chapter describes how electronic database search works, with insights into some poor practices. The concept of Boolean operators and how they can be used to easily mine desired contents from electronic databases are discussed. The knowledge and use of Boolean operators might become unavoidable in enabling researchers to locate relevant materials for their projects.

INTRODUCTION

The research process includes problem ideation, conceptualisation, literature review, study design, data collection and analysis, and ultimately, disseminating results (in different formats) to publishing outlets. In this chapter, two of these processes are important – literature review and result dissemination. The literature review process entails searching for previously published content from different databases (traditional or electronic). Literature review presupposes that the material to be reviewed is already

DOI: 10.4018/978-1-6684-3364-5.ch009

published or disseminated to the public domain. Therefore, a symbiotic relationship exists between literature review and research dissemination. This is because the review of research literature depends on previously disseminated works. Besides, a review of existing literature is required to offer future research directions and sustain dissemination practices.

The published research literature is often managed using physical or electronic databases, with each requiring different methods of assessing indexed contents. To locate research literature from traditional/physical databases, one must visit the library. After that, a catalogue is used to search for such information depending on the form of library classification. In other cases, individuals must go through different files and folders, searching for such information. However, information search and retrieval patterns work differently from the traditional system in electronic databases. This chapter is not aimed at comparing these systems but to discuss electronic databases, their evolution, advantages and disadvantages, the indexing processes and how indexed information can be assessed using a state-of-the-art approach. The emphasis of this chapter is on the Boolean operators due to their growing importance and utility in mining records from large databases or data warehouses. These operators allow for reducing or expanding the number of records returned. Boolean operators may save time by restricting the search scope and eliminating irrelevant results to users' needs.

Therefore, after reading this chapter, readers are expected to be able to: explain the meaning of electronic databases; discuss the evolution of electronic databases; define the term data mining; mention at least four advantages and disadvantages of electronic databases; describe electronic database search works; discuss the indexing processes of scholarly materials in an electronic database; make practical use of Boolean operators in retrieving literature materials from electronic databases.

ELECTRONIC DATABASES

Meaning

There is no universally accepted meaning for electronic databases due to variations in usage across disciplines, authors, time, and place. However, an electronic database can be defined as an organised collection of information that has been saved on one or more digital media. These materials can be textual or non-textual but are in digital forms to aid easy retrieval and consumption. Stored materials in electronic databases are accessible to users with authorisation and privileges. Electronic databases are used for storing, retrieving, and managing information on a broad range of subjects. Textual materials include books, journal articles, theses, dissertations, patents, conference proceedings, monographs, and other resources stored in digital forms. These resources can be accessed using electronic devices such as computers, smartphones and tablets. On this note, Koedinger et al. (2008) state that e-learning tools, instructional technology, the Internet, and institutional archives of student records have resulted in large data warehouses that educators may access. Therefore, electronic databases are treasures of knowledge discovery that can be investigated and used to help us understand how students learn (Mostow & Beck, 2006). It has also been stated that a large amount of data is available in electronic educational databases that may be used to make better management choices (Bala & Ojha, 2012).

Because of the drive toward open science or open access publications, electronic databases offer a rich variety of scholarly content beyond the offerings of traditional libraries. Besides, in contemporary times, big data has occasioned diverse forms of non-textual materials that can meaningfully be stored