Chapter 38
Web-Scale Discovery:
A Library of Babel?

William Breitbach
California State University, USA

ABSTRACT

With significant developments in library discovery systems, many libraries are exploring options for improving access to content. Although these new systems appear to be better than what has come before (federated searching), many problems remain. Libraries therefore should consider a number of interrelated issues/challenges before investing. These challenges include: an information glut, devaluation of metadata, disconnection of content from discourse communities, the notion that libraries compete with Google, and the creation of false expectations for users searching for scholarly content.

INTRODUCTION

When it was first proclaimed that the Library contained all books, the first impression was one of extravagant happiness (Borges, 1962, p. 55).

In Jorge Luis Borges' famous short story *The Library of Babel*, a universal library exists where every possible book is located. However, its vastness and lack of organization make the library essentially useless. Because of the continued failure of users to understand the library's organization, many go mad, commit suicide or become paranoid. Although the author of this chapter does not anticipate such calamity as libraries adopt Web-scale discovery systems (WSDS), there are a wide range of potential problems libraries need to consider before we reach the stage of "extravagant happiness."

There have been significant developments in WSDS over the past few years. Many libraries have adopted or are considering the adoption of such services. WSDS have great potential, but

DOI: 10.4018/978-1-4666-1821-3.ch038

readers should pursue their adoption with a critical eye. Before investing in these expensive systems, libraries should consider the problems they are trying to solve and whether these systems will solve them. To assist in the evaluation of these systems, a discussion of the following interrelated issues/challenges govern the body of this essay: an information glut, devaluation of metadata, disconnection of content from discourse communities, the notion that libraries compete with Google, and the creation of false expectations for users searching for scholarly content.

BACKGROUND

The goal of a single search for all library resources has been around for some time. The first glimpse of the presumed future came with federated searching. However, the poor search performance and poorly designed user interfaces left much to be desired. Moreover, most of the studies on federated searching focused on user satisfaction and perception (for examples, see Armstrong, 2009; Williams, Bonnell, & Stoffel, 2009) rather than student outcomes or performance measures of the systems. A few studies of federated searching did indeed attempt to tackle the more essential issue of task performance. These studies used research scenarios to assess user outcomes by counting the number of documents saved, time saved with federated search, and quality of results (Haya, Nygren, & Widmark, 2006; Belliston, Howland, & Roberts, 2007). However, these studies did not show better task performance with federated searching, begging the question of why libraries would invest in systems that make little to no difference in learning/discovery outcomes?

A new wave of discovery services has arrived. WSDS offer many of the same promises of federated search but differ in some very important ways. Federated search engines search and retrieve records from multiple databases from multiple hosts simultaneously. WSDS are different in that they use a single centralized and pre-aggregated index. This offers opportunity for improving performance (speed), relevance, and normalization of data. Jason Vaughan's much discussed issue of *Library Technology Reports* (2011) gives a substantive description of the promises of Web-scale discovery, but does not adequately address the associated challenges. Some of the early studies of WSDS show increases in the use of content by measuring SFX®[1] click-throughs (Garrison, Boston, & Bair, 2011; Way, 2010). Although these studies are certainly interesting, welcome, and even promising, additional information on how users interact with the information beyond simply clicking is warranted. That said, a recent task-oriented usability study shows some evidence that EBSCO Discovery Service™[2] performs better than federated searching in terms of finding relevant content and in ease of use (Williams & Foster, 2011). Williams & Foster (2011) created task oriented scenarios to test different parts of EBSCO Discovery™. These tasks were based on research scenarios such as finding an article on a complex topic, finding and emailing records, accessing full text, and discovering books in a local collection. In the author's view, the type of task-oriented research in the Williams and Foster (2011) study offers a good start, but comparative research where students solve common task-oriented problems in both WSDS and traditional databases is warranted.

Issues, Controversies, Problems

The notion of a single discovery interface for all available resources has great surface appeal. A beautiful and elegant search box that alleviates the need to navigate complex Web sites, learn multiple information systems, and which streamlines access to content has been a dream for at least a decade. Moreover, the vision of delivering high quality content within an interface that is somewhat familiar to novice users sounds compelling. This is essentially the promise of WSDS.