Chapter 126

Search Engines: Past, Present and Future

Patrick Reid
AstraZeneca, UK

Des Laffey
University of Kent, UK

ABSTRACT

This chapter covers key issues in the area of search engines. It shows the importance of search by explaining what search engines are and their significance to business and society. The mechanics of search are outlined which includes developments to the current age. The search market is then covered which stresses Google’s dominance of most national markets. Search engine optimization is then analysed looking at the key factors which determine position. The chapter also looks at the key funding mechanism for search, paid search advertisements. Finally, the article looks at emerging issues in search, including rich media and mobile, and privacy issues.

INTRODUCTION

Search engines are fundamental to modern life, with Figure 1 illustrating the number of global searches in August 2007.

The use of the term Google as a verb is perhaps the strongest evidence of the impact of search and of Google’s status as the dominant provider. Search has a key role in modern society and as Rangaswamy et al. (2009, p49) write “search results can influence important decisions about someone’s life, health, or a major purchase, or an entrepreneur’s quest for an acquisition target.”

This chapter makes a contribution to the Encyclopedia by outlining the key issues regarding search. It does this by integrating ideas from academic and practitioner audiences to offer an integrated perspective on this important topic. The article firstly covers the key definitions, explains how search engines work and discusses the challenges of Web search. The competitive environment of search is then outlined which stresses Google’s dominance but notes markets where national champions are dominant. The essential topic of search engine
Search Engines

optimization (SEO) is then analysed. The article then covers the evolution of paid search, the use of text based advertisements which are triggered by the terms in a search. Finally before it concludes emerging issues in the search engine field are considered in the areas of the semantic Web, rich media, location based search, mobile search and the challenges of privacy.

BACKGROUND

What is Search? Search capabilities are required for any information system, and the need grows in importance as the volume of data grows (Frana, 2004). There are key differences in the methods of how the search challenge is addressed:

- A directory is a human generated index (database) of websites, with the most well known examples being Yahoo and the Open Directory.
- Organic search refers to computer generated search results which appear based on how relevant the pages are to the user’s search, an early example being Excite with Google the obvious current example.
- Meta-search in turn describes search engines which present aggregated results they have taken from search engines.
- Paid search, also known as sponsored search, refers to payment on a per click basis for text advertisements which are triggered by a search term (Laffey, 2007). Paid search provides the revenue model for organic search and Jansen (2007) estimates that 30% of organic searches lead to one or more clicks on a paid search link.

The Mechanics of Organic Search: The key reference for how search engines work is the seminal paper by Brin and Page (1998). Other authors such as Tassabehji (2003) or Schneider (2007) also cover the mechanics of organic search. A search engine has three essential aspects: the crawler or spider, which retrieves information from webpages, and other documents, it finds as it follows the link structure of the Web; the index, which stores relevant information about the document; and finally, the search engine software which contains algorithms, or rules, which decide how relevant a document is to a user search.

The Surface Web and the Deep Web: The Surface Web refers to pages that can be indexed by search engines. Google stated in 2008 that their systems were aware of 1 trillion unique pages (Google 2008). To place this in context when Google’s prototype search engine was available in 1998, it indexed 24 million pages (Brin and Page, 1998). Estimating the size of the Web is very difficult as crawlers only become aware of new URLs either when they are informed by a webmaster or when they come across a link.