Associating Searching on Search Engines to Subsequent Searching on Sites

Adan Ortiz-Cordova, College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA, USA

Bernard J. Jansen, Social Computing Group, Qatar Computing Research Institute, Doha, Qatar

ABSTRACT

In this research study, the authors investigate the association between external searching, which is searching on a web search engine, and internal searching, which is searching on a website. They classify 295,571 external – internal searches where each search is composed of a search engine query that is submitted to a web search engine and then one or more subsequent queries submitted to a commercial website by the same user. The authors examine 891,453 queries from all searches, of which 295,571 were external search queries and 595,882 were internal search queries. They algorithmically classify all queries into states, and then clustered the searching episodes into major searching configurations and identify the most commonly occurring search patterns for both external, internal, and external-to-internal searching episodes. The research implications of this study are that external sessions and internal sessions must be considered as part of a continuous search episode and that online businesses can leverage external search information to more effectively target potential consumers.

KEYWORDS

External Search, Internal Search, Site Search, Web Search Engines, Website Searching

INTRODUCTION

Searchers often use multiple information platforms such as search engines or websites in order to address a task that involves searching. A substantial percentage of Web users commonly employ major search engines (e.g., Google, Baidu, NAVER, Yandex) as entry points to the Web. These searchers may traverse from the major search engine to a particular commercial website (e.g., Amazon, Gmarket, TMall, Ulmart) where they may continue to refine their search queries. However, given the difficulty of collecting data from multiple sites, searching on each of these platforms have been analyzed in isolation (i.e., treated as separate searching sessions). In reality, given the single underlying search task, the searches on these separate platforms are more likely to comprise a single searching episode and should be examined holistically in order to better understand the user’s underlying task.

In our research, we investigate one aspect of this multi-platform searching context, namely the search that brings the user from the major search engine to the particular website. Understanding this searching situation is important, as the majority of websites rely on search engines to direct substantial percentages of their traffic, with more than 80% of Web users employing a search engine as their starting point (Bughin et al., 2011) for locating other websites. This search engine traffic is critical for many online businesses given that traffic from search engines is continual, and provides a direct funnel of potential customers, clients, users, etc.

One method of engaging this inbound traffic is to provide a searching capability using a site search application so that users can find the information they are looking for, thereby remaining on
the website. As many websites are structurally complex and information-rich, visitors can leverage an internal site search capability to locate what information they want or need. Therefore, site search is an important informational architecture and navigational feature for a website.

We define both the capability and the act of searching a website as referred as *internal search*, and we define queries submitted during internal search as *internal search queries*. We define internal search as *one or more queries submitted to a site’s specific search service in order to find information that is contained on that site*. The capability and the act of searching using a general purpose search engine is referred to as an *external search*, and queries submitted during external search are referred to as *external search queries*.

In this research, we investigate the transition between *external search* (ExS), which is searching on a major web search engine, and *internal search* (InS), which is the subsequent searching on a particular website.

Our research is focused on those searches where a searcher conducts an ExS and then a subsequent InS in continuation of the same search task. In these *external to internal search* (Ex2InS) episodes, the searcher conducts the ExS by submitting a query to a major search engine and then submitting subsequent queries, one or more, to a website search appliance.

So, though ExS and InS can each occur in isolation (i.e., ExS with no InS, or InS with no ExS), this research focuses on the combined occurrences, specifically the transition from ExS to the subsequent InS usage. Figure 1 illustrates the ExS, InS, and Ex2InS concepts.

Previous research has not thoroughly examined the relationship between ExS and InS. In fact, prior research has mainly treated these engagements as separate and distinct searching sessions; though, obviously, they are at least connected by the last ExS query because it is the referral query on the web search engine that brings the visitor to a particular website.

Therefore, our premise is that the InS is a continuation of the last web search engine session query. In these situations, there is a connection between the ExS and the InS, and they are each part of the same underlying searching task. Therefore, it is essential to investigate the InS behaviors that are manifested following ExS in order to develop better InS capabilities, design searching personalization, identify sponsored search keyword generation, discover user intent, and document missing content that the online business may not have in its data collections.

---

**Figure 1. Contrast of External to Internal Search (Ex2InS), External Search (ExS), and Internal Search (InS)**
Privacy Management Architecture Privacy Technologies
www.igi-global.com/chapter/privacy-management-architecture-privacy-technologies/28143?camid=4v1a

Online Services Delivered by NTO Portals: A Cross-Country Examination
www.igi-global.com/chapter/online-services-delivered nto-portals/50240?camid=4v1a