Dq1 week 1

Technology, Information Technology

According to BrightPlanet the deep Web comprises of two data categories. The first is data that is difficult to obtain using standard searching techniques while the second is data that is not accessible to standard search engines. Examples of the former include information presented in new formats that search engines are yet to recognize and new and/or real time content that the search engines are yet to index. However, BrightPlanet (2012) finds the second category to be of greater importance because it tends to have deeper content that is intentionally set up to be hidden. This essentially means that an ordinary Internet search will miss out on the vast resources hidden within the deep Web.

To appreciate the importance of the deep Web one has only to look at Bergman (2001) findings. To begin with, the deep Web is not only 400 to 550 times larger than the surface Web but is also the largest growing category of new information on the Internet. Also, Bergman (2001) argued that the deep Web's content quality surpasses that of the surface Web by a factor ranging from 1, 000 to 2, 000 times. Most of all this content is highly relevant to most information needs. This element of relevance to most information needs implies that both academia and business should find access to the deep Web's resources important.

The world's biggest search engine, Google was conceived by Sergey Brin and Larry Page when they faced difficulty in finding relevant academic data in the then early Web. Access to relevant, current and quality information is the crux of academia. With conventional search engines accessing only about 16% of the information available on the World Wide Web (LaGuardia Community College 2012) academic researchers risk losing out on a vast

resource of information-rich literature content. This is especially true given Lewandowski and Mayr (2006) findings that: firstly, more and more scholarly content are being provided on the Internet and secondly, that researchers are increasingly consulting general-purpose Internet search engines to retrieve these scholarly documents. In addition to this on-going digitization projects are contributing to the continuous growth of the Invisible Web. In light of this, Academia the importance of accessing the deep Web to academia cannot be gainsaid. Researchers have no option but to devise better tools and techniques to enable better and greater access into this information-rich " portal" of knowledge referred to as the deep Web.

On the other hand, business too has a lot to gain or lose from embracing or spurning the deep Web. Similar to academia, researchers state that the Internet is not only the fastest growing new medium of all time but has also become the information medium of first resort for many businesses (Lyman & Varian 2003). Furthermore, it is difficult to picture any medium-sized or large business without e-Business in this age (Kotadia 2012). This has made the Web a popular platform for conducting competitive intelligence (CI), where CI is defined as the selection, collection, interpretation and distribution of publicly held information that is strategically important to a firm (Boncella 2003). With the deep Web providing better quality and more relevant information, it is only rational for businesses that seek more insightful competitive intelligence to create competitive advantages to access the deep Web. Thus accessing and exploiting the resource in the deep Web is as important to business as it is to academia.

References

Bergman, M. K., 2001. The Deep Web: Surfacing Hidden Value. Journal of Electronic Publishing, 7(1). Available at: http://quod. lib. umich. edu/cgi/t/text/text-idx? c= jep; view= text; rgn= main; idno= 3336451. 0007. 104.

Boncella, R. J., 2003. Competitive Intelligence and the Web. Communications of the Association for Information Systems, 12, pp. 327–340.

BrightPlanet, 2012. Deep Web: A Primer. BrightPlanet. Available at: http://www. brightplanet. com/deep-web-university/deep-web-a-primer/ [Accessed May 5, 2012].

Kotadia, H., 2012. Why Big Data Mining / Analytics is the New Gold Rush | Dr. Harish Kotadia. Blog on Big Data Analytics, Social CRM and Business. Available at: http://hkotadia. com/archives/4887 [Accessed May 5, 2012].

LaGuardia Community College, 2012. Beyond Google: The Invisible Web - Some Fast Facts About the Invisible Web. LaGuardia Community College Library Media Resources Center. Available at: http://library. laguardia. edu/invisibleweb/facts [Accessed May 5, 2012].

Lewandowski, D. & Mayr, P., 2006. Exploring the Academic Invisible Web. Library Hi Tech, 24(4), pp. 529–539.

Lyman, P. & Varian, H., 2003. How Much Information? 2003, University of California at Berkeley.