ONLINE VISITS TO LARGE PUBLIC LIBRARIES IN AUSTRALIA – A COMPARISON

Searchers Working Paper 6

Vivienne Waller
December 2008

The Searchers

Institute for Social Research
Swinburne University of Technology
Acknowledgements

This Working Paper is part of the Australian Research Council project The Searchers. We would like to acknowledge the generous support of Hitwise and the State Library of Victoria. Particular thanks go to Sandra Hanchard and Ben Carroll of Hitwise and to Sue Hamilton, Ian Patterson and Janice van de Velde of the State Library of Victoria. This research was funded by Australian Research Council grant LP0775215, conducted under the guidance of Chief Investigators Professor Julian Thomas and Professor Denise Meredyth and with the assistance of Dr Ian McShane. The Institute for Social Research takes all responsibility for the content. The views of the authors are not necessarily those of the State Library of Victoria.
**Background to the analysis**

The Searchers project examines the strategic challenges for major public libraries presented by the on-line information environment. The project has three components: research on current information-seeking and the role of public libraries in online information provision, analysis of wider developments in online environments and search models, and evaluation of the policy and strategic implications for public libraries highlighted by the research. The project will provide guidance for the State Library of Victoria in developing appropriate models of information provision and contribute new thinking on the role and position of major libraries in the digital age.

**Scope of the analysis**

This report contributes to the research on current information-seeking. It uses data from Hitwise to compare the type and number of visitors to the websites of the State Library of Victoria (SLV) with the type and number of visitors to the websites of the National Library of Australia (NLA) and the State Library of Queensland (SLQ).

Specifically, this report compares the following three aspects of visits to the websites. These are:

- the physical location of online visitors,
- the lifestyle characteristics of online visitors,
- the frequency of online visits.

The characteristics of online visitors are compared using Mosaic lifestyle profiles linked to Hitwise data on visits to the websites. The frequency of online visits is compared by looking at the websites' ranking and market share in the reference category.

The data in this report does not distinguish between online visitors who came to the library website via search engine results, those who came via a hyperlink and those who typed in the URL or had it bookmarked.

Further analysis of online visits to the website of the State Library of Victoria can be found in the Searchers Working Paper 3: Virtual Visitors to the SLV website, August 2008.

**Hitwise methodology**

Each different method for measuring web activity has its own particular benefits and limitations. Hitwise measures internet use using a network centric methodology
and a proprietary method to match internet use with lifestyle profiles. The nature of this method is indicated in Figure 1.

![Figure 1: Schematic depiction of Hitwise methodology](image)

As figure 1 shows, each household is assigned a lifestyle grouping on the basis of characteristics of adjacent households. This grouping is attached to the web activity data originating from a household that provides data to Hitwise.

Mosaic 2006 lifestyle groupings are constructed indirectly from more than 1,000 variables and directly from 207 variables deriving from a range of sources including 2001 ABS census data, 1999/2000 Australian Taxation Office data, and marketing surveys. Similarly, Mosaic 2008 lifestyle groupings are constructed from 238 variables including 2006 ABS census data. (For more information on how these profiles are developed see http://www.mosaicaustralia.com.au/).

Hitwise data covers more than one third of Australian internet subscriptions, including homes, businesses, schools, universities, and libraries. It is the resulting aggregate, summarized and weighted statistical information that is used in this paper. As a network centric methodology, the main strength of Hitwise lies in the ability to make comparisons between different websites. It is also particularly useful in having some demographic data attached. Another advantage of Hitwise data is that, as it is based on activity through ISP subscriptions, it excludes the web activity of ‘spiders’, automated computer scripts which ‘crawl’ the World Wide Web indexing pages.
**Caveats to this study**

The main disadvantage of using Hitwise data for this study is that it reports on a sample, rather than total web activity at a website. Although the Internet Service Providers that provide data to Hitwise include a representative cross-section of sizes, there may be some sample bias, the direction of which is impossible to detect. Because it is reporting on a sample, Hitwise data is presented in percentage terms rather than as absolute numbers.

The biggest caveat to this comparison of visitors to the websites of SLV, SLQ and NLA is the fact that the websites are not directly comparable. The architecture of each website is complex, made up of a number of websites, and a variety of subdomains. For the purposes of reporting, this analysis glosses over these complexities.

In addition, there are limitations to what Hitwise reports on separately. For example, it was not possible to separately identify visits to the SLQ catalogue, and it was not possible to include visits to the SLV and NLA catalogue in the analysis of demographic data. It also was not possible to separate Libraries Australia from the catalogue of NLA.

The data on lifestyle and physical location of visitors compares the following websites:

SLQ – slq.qld.gov.au (including SLQ catalogue)

SLV – slv.vic.gov.au (excluding SLV catalogue)

NLA – nla.gov.au (excluding NLA catalogue, Libraries Australia and Pandora)

The data on ranking and market share compares the following websites:

SLQ – slq.qld.gov.au (including SLQ catalogue)

SLV – (website and catalogue identified separately)

NLA – (website and catalogue identified separately, excludes Pandora, catalogue includes Libraries Australia).
As Figure 2 indicates, in 2007, 81% of visitors to the SLQ website were located in Queensland and 62% of visitors to the SLV website were located in Victoria. The distribution of the Australian population is included for purposes of comparison (ABS 2008).

It is clear that the physical spread of visitors to the NLA website most closely matched the distribution of the Australian population with the ACT only slightly overrepresented. SLQ was most likely to have online visitors located in the same State/Territory as the physical library. However, the proportion of online visitors to SLQ located in other States appears to be increasing. Figure 3 indicates the physical location of visitors to library websites between April-June 2008. In 2008, only 70% of visitors to the SLQ website were located in QLD. The proportion of SLQ visitors from Victoria had increased from 5% in 2007 to 7% in 2008. The proportion of SLQ visitors from NSW had more than doubled from 8% in 2007 to 17% in 2008.

The distribution of visitors to the SLV website was similar in 2007 and 2008 with an overrepresentation of visitors from Victoria and Tasmania.
Characteristics of online visitors - Lifestyle groupings

Figure 4 compares the lifestyle profile of visitors to the SLV, SLQ and NLA websites for the period April-June 2007. An index of 100 shows that the proportion of visitors to a website within that lifestyle profile matches the proportion of that lifestyle profile in the Australian online population. Any index over or under 100 shows that the website is over or under-represented within that lifestyle profile. The legend to the graph indicates the prevalence of each lifestyle group in the Australian population. For example Figure 4 indicates that two of the largest groups in the Australian population (F - better off families in core suburban and rural fringe locations and I – disadvantaged peripheral urban and country neighbourhoods) are underrepresented as visitors to the SLV, the former only slightly, the latter quite markedly. The latter group is also underrepresented in visitors to the NLA and SLQ.
To some extent the composition of visitors to the library website reflect the composition of the catchment area. Visitors to SLQ website are predominantly from Queensland while visitors to SLV are predominantly from Victoria. Although data was not available on lifestyle profile at a State level, it is to be expected that SLQ would attract more Anglo-Australian blue collar families from provincial towns, simply because there are likely to be a greater proportion of households like this in Queensland than in Victoria.

Figure 5 compares the lifestyle profile of visitors to the SLV, SLQ and NLA websites for the period April-June 2008. In 2008, the profiles were updated to reflect changing demographic composition. Hence, Figure 4 and Figure 5 are not comparable. Even though the groups from the two segmentation systems share codes and may have similar names, the 2008 segmentation is not a continuation of the 2006 segmentation.
In general, online visitors to SLV are wealthier and better educated than the general online population. It could be that these people were more likely to know about the SLV site or it could be that their online activities were more likely to bring them to the SLV site.

Although the 2008 Mosaic groupings are not comparable to the 2006 groupings, it can be seen that in 2008, the SLV still attracted online visitors who were wealthier and better educated while online visitors on a lower income from the outer suburbs were underrepresented.

The SLV did have an overrepresentation of culturally diverse families, which may reflect the composition of the Victorian online population.

**Frequency of visits – rankings and market share**

The Hitwise Reference category includes all library websites as well as any site that might be used as a reference source, or any site which advertises a reference source. At the end of 2008, this category contained 3,561 websites and received 2.3% of all web traffic.
Figure 6 shows the relative rankings of the catalogues for SLV and NLA and the websites for the three libraries. The top ranking is “1”; in other words, the lower the ranking, the more visits to the website.

Note that the catalogue for SLQ cannot be reported on separately. Even so, the SLV website, excluding the catalogue, has consistently outranked the SLQ website (which includes visits to their catalogue). Both have held fairly stable rankings throughout 2007 and 2008. The NLA website has consistently had the highest rank of all, ranking number 34 in the reference category in November 2008.

The catalogue of the SLV has moved up and down in the rankings quite dramatically, and the catalogue of the NLA has similarly shown a fair degree of up and down movement. Much of this is likely to be an artefact of looking at rankings. The lower the ranking, the less difference in numbers of visits is needed to move above or below closely ranked sites. Hence, small changes in the number of visits can have a large effect on the ranking. However, as discussed next, there are identifiable reasons for the dramatic increase in the ranking of the NLA catalogue from 286 in November 2006 to 21 in November 2008.
Another, more intuitively appealing way of looking at the relative performance of the websites is to look at the market share of each website. Figure 7 compares the market share of each website and the catalogues of the NLA and SLV. It shows a similar pattern to the ranking, but is easier to understand as increases in market share correspond with increases in visits to the site. It can be seen that the SLV website experienced an increase in market share between January-May 2008. This may have been because of the very popular exhibition The Medieval Imagination held during that time. The NLA website reached a peak in early 2007 and its market share has been decreasing in the second half of 2008.

What is immediately striking in Figure 7, however, is the dramatic increase in visits to the NLA catalogue (which includes Libraries Australia) since July 2008. This is quite possibly a direct effect of the agreement made in May 2008 between OCLC and Google. Implementation of the agreement means that an internet user who finds an item through Google book search can also link directly to the catalogues of the closest OCLC member libraries which have that item in their collection.
Figure 8: Comparison of market share – excluding catalogue for NLA and SLV, November 2006-November 2008

Figure 8 compares the market share of the three library websites, excluding the catalogue for NLA and SLV. It depicts a selection of data from Figure 8, making the comparison between these websites clearer. It can be seen that the market share of SLV reached a peak in mid 2007 and has decreased slightly over the last two years. The market share of SLQ has been remarkably stable. NLA has roughly double the market share of SLV, but this has been decreasing since mid 2008.

References: