Evaluating the online activity of users of the e-Bug web site

Ed de Quincey1, Patty Kostkova1*, Gawesh Jawaheer1, David Farrell1, Cliodna A. M. McNulty2 and Julius Weinberg1 on behalf of the e-Bug Working Group†

1City eHealth Research Centre, City University, Northampton Square, London EC1V 0HB, UK; 2Health Protection Agency Primary Care Unit, Microbiology Department, Gloucester Royal Hospital, Great Western Road, Gloucester, Gloucestershire, GL1 3NN, UK

*Corresponding author. Tel: +44-20-7040-4084; Fax: +44-20-7040-8795; E-mail: patty@soi.city.ac.uk
†Members are listed in the Acknowledgements.

Web server log analysis is being increasingly used to evaluate the user behaviour on healthcare resource web sites due to the detailed record of activity that they contain. This study aimed to use this information to evaluate the e-Bug web site, a healthcare resource that provides a range of educational resources about microbes, hand and respiratory hygiene, and antibiotics. This evaluation was conducted by analysing the web server logs of the e-Bug web site for the period January 2008 to November 2009, using a proprietary application named Sawmill. The e-Bug web site has had >900000 page views generated from >88000 users, with an increase in May 2009 during the swine flu epidemic and a further increase in September 2009 following the official launch of e-Bug. The majority of visitors were from the UK, but visits were recorded from 190 different countries. Word® document resources were downloaded >169000 times, with the most popular being a swine flu fact-sheet. PowerPoint® document resources were downloaded >36000 times, with the most popular relating to the ‘chain of infection’. The majority of visitor referrals originated from search engines, with the most popular referral keywords being variations on the e-Bug name. The most common non-search engine referrals were from other healthcare resources and agencies. Use of the site has increased markedly since the official launch of e-Bug, with average page views of >200000 per month, from a range of countries, illustrating the international demand for a teaching resource for microbes, hygiene and antibiotics.

Keywords: education, web server logs, evaluation

Introduction

The importance of understanding how users interact with healthcare resources provided on web sites has been demonstrated in a number of studies.1,2 User-centred evaluation, such as task analysis and think aloud protocol,3 where quantitative and qualitative feedback is elicited from an observed user, reduces the number of potential respondents recruited.

Implicit evaluation methods such as web server log analysis are therefore being increasingly seen as a principal source of information for determining online user behaviour.1 Web server logs contain information about which pages have been viewed, when they were viewed, the referral page (if present) and the Internet Protocol (IP) address of the device from which the user has viewed the page, along with the browser type (e.g. Internet Explorer or Firefox). The advantage, therefore, is that web server logs contain a complete and accurate record of all browsing activity on a web site.

e-Bug is a European Commission-funded educational project aimed at young people. It combines traditional methods of classroom delivery complemented with online, browser-based (Flash) games to teach pupils in junior and senior schools about microbes, hand and respiratory hygiene, and antibiotics. The project includes 10 associate European countries that translated and disseminated the resources. Lesson plans with extension activities and videos with media are available on the e-Bug web site, along with educational games that can be used alone or with the pack activities.4

The main aim of analysing the web logs was to determine the general use of the e-Bug web site, the geographic distribution of users, the number of times different pages were viewed and downloaded, and from where visitors were referred. Evaluation of the games will be published separately.

Methods

This evaluation was conducted by analysing web server logs of the e-Bug web site for the period January 2008 to November 2009. The web site was launched in three languages (English, Czech and French) in piloting countries in January 2008 and translated into other e-Bug associate countries’ languages after modifications in December 2009. The evaluation was conducted in line with the methodologies presented in previous studies by Madle et al.,1,2 where server log analysis was used to determine the browsing behaviour on the National Electronic Library of Infection (NeLi).

The server logs provided data about users in a number of fields, including: the IP or hostname of the origin of the request; the date and time of the request; the type of request; the page requested; the returned...
status of the page; the number of bytes transferred; and the referring
Universal Resource Locator, if available. These raw data were then
cleaned and analysed using a proprietary application named Sawmill. This
removed search engine crawlers (which are automated and index
content) to produce a set of standard reports that covered information
such as visits, hits, content viewed, visitor demographics, systems and
referrers.

The key points investigated were:

1. The number and geographical location of e-Bug web site users,
   sessions and pages accessed.
2. The most commonly accessed resources on the e-Bug site. The
   resources, in the form of PowerPoint® presentations and Word®
   documents on the e-Bug web site are categorized into two main
   areas, ‘Junior’ and ‘Senior’, relating to the educational level at
   which the resources are aimed. These files are also provided in
   a single compressed ‘Zip’ file, which contains all of the resources in
   the section. This is referred to as a ‘Pack’ download.
3. Referrers—the sites that visitors were being referred from and the
   search keywords that were being used to find the site.

Results

General use of the e-Bug web site

The e-Bug web site had 934,532 page views from 88,427 visitors
in the 23 months analysed. There were 3,208 complete pack
downloads. Downloads of junior and senior individual activities,
and PowerPoint presentations were similar over the time
period. Figure 1 shows the number of visitors accessing the
site, the number of individual pages of the site accessed and
the total resource downloads (PowerPoint and Word documents)
during each month from January 2008 to November 2009.

The activity on the site has increased steadily since 2008, with
the rise in May 2009 coinciding with the swine flu epidemic when
a link to the e-Bug site was placed on the UK HPA web site to
direct visitors to the various swine flu resources that have been
posted on the site. The increase in September 2009 coincided
with the official e-Bug launch, which received national news
coverage. Complete pack downloads mainly occurred after
September 2009.

Geographical distribution of visitors

Over one-third of the visitors (36%) were from the UK, followed
by France (17%), Portugal (13%) and Belgium (8%). In total,
visits were recorded from 190 different countries. The Czech
Republic, the other pilot web site country, accounted for only
2% of visitors. Interestingly, the USA and Australia, both English-
speaking countries, accounted for 4% and 2% of the total
visitors, respectively.

Most commonly accessed resources

Most commonly accessed Word documents

Visitors downloaded 2,796 individual Word documents 16,9607
times. Junior page Word documents accounted for 74,627 down-
loads and 55,756 were senior page Word documents. The most
popular documents were the junior and senior swine flu fact-
sheets (54,52 downloads) and a handout document related to
‘Hand Hygiene’.

---

Figure 1. e-Bug web site: number of visitors, page views and resource downloads per month (January 2008 to November 2009).
Download of documents in French accounted for 42310 downloads, compared with 49520 in English. This figure does not include several other English documents, such as the various swine flu factsheets that were posted on the site as they were produced in response to the epidemic and were not an original part of the e-Bug web site. The figures have been included in the totals and Table 1.

Table 1 shows the top 20 Word document resources that were downloaded from the e-Bug web site. The most frequently downloaded document was a resource on swine flu, with 80.1% of downloads being from the UK, 4.2% from the USA, 1.4% from Ireland, 1.2% from Austria, 1% from France and the remaining 67 countries each representing downloads of <1%.

The referrals from the HPA web site were mainly directed to the e-Bug homepage but, interestingly, a number of referrals were recorded that directed visitors to information about hand hygiene and respiratory hygiene. This behaviour was predominantly demonstrated in April 2009 (3190 referrals) and May 2009 (3443 referrals), during the initial phases of the swine flu pandemic.

Most commonly accessed PowerPoint documents

Over 300 (339) individual PowerPoint documents were downloaded from the e-Bug web site (Table 2) 36235 times. Of the PowerPoint documents that were in the junior and senior sections of the site, 9106 were junior document downloads and 22375 were senior document downloads. The most popular document, with 6012 downloads, was the ‘Chain of Infection’ PowerPoint presentation followed by a presentation containing information about ‘Good Microbes’ (1743 downloads).

External site referrers

Referrals from 1821 individual external web sites generated 90473 page views. The majority of these referrals originated from search engines such as Google and Bing (16 of the top 20 referrers). Table 3 shows the top 10 non-search engine referral sites since January 2008, along with the number of page views and visitors they produced.

The referrals from the HPA web site were mainly directed to the e-Bug homepage but, interestingly, a number of referrals were recorded that directed visitors to information about hand hygiene and respiratory hygiene. This behaviour was predominantly demonstrated in April 2009 (3190 referrals) and May 2009 (3443 referrals), during the initial phases of the swine flu pandemic.
Table 3. Page views and visitors generated from referrals from external web sites

<table>
<thead>
<tr>
<th>Site</th>
<th>Number of page views</th>
<th>Number of visitors</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.hpa.org.uk/">http://www.hpa.org.uk/</a></td>
<td>8818</td>
<td>4688</td>
</tr>
<tr>
<td><a href="http://www.nric.org.uk/">http://www.nric.org.uk/</a></td>
<td>1008</td>
<td>870</td>
</tr>
<tr>
<td><a href="http://www.healthyshools.gov.uk/">http://www.healthyshools.gov.uk/</a></td>
<td>948</td>
<td>713</td>
</tr>
<tr>
<td><a href="http://www.neli.org.uk/">http://www.neli.org.uk/</a></td>
<td>560</td>
<td>487</td>
</tr>
<tr>
<td><a href="http://www.dh.gov.uk/">http://www.dh.gov.uk/</a></td>
<td>498</td>
<td>442</td>
</tr>
<tr>
<td><a href="http://www.traininginfection.org.uk/">http://www.traininginfection.org.uk/</a></td>
<td>486</td>
<td>451</td>
</tr>
<tr>
<td><a href="http://www.inpes.sante.fr/">http://www.inpes.sante.fr/</a></td>
<td>469</td>
<td>408</td>
</tr>
<tr>
<td><a href="http://www.netmums.com/">http://www.netmums.com/</a></td>
<td>377</td>
<td>357</td>
</tr>
<tr>
<td><a href="http://www.nhs.uk/">http://www.nhs.uk/</a></td>
<td>348</td>
<td>270</td>
</tr>
</tbody>
</table>

Table 4. Most common keyword referrals from search engines since January 2008

<table>
<thead>
<tr>
<th>Search phrase</th>
<th>Number of page views</th>
<th>Number of visitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>e-bug</td>
<td>4880</td>
<td>3349</td>
</tr>
<tr>
<td>Chain of infection</td>
<td>2573</td>
<td>1822</td>
</tr>
<tr>
<td>Ebug</td>
<td>2522</td>
<td>1707</td>
</tr>
<tr>
<td>e bug</td>
<td>1678</td>
<td>1209</td>
</tr>
<tr>
<td><a href="http://www.e-bug.eu">www.e-bug.eu</a></td>
<td>1385</td>
<td>882</td>
</tr>
<tr>
<td>Useful microbes</td>
<td>1082</td>
<td>770</td>
</tr>
<tr>
<td>e-bug.eu</td>
<td>1017</td>
<td>772</td>
</tr>
<tr>
<td>Harmful microbes</td>
<td>883</td>
<td>656</td>
</tr>
<tr>
<td>Microbios uteis</td>
<td>678</td>
<td>428</td>
</tr>
<tr>
<td>Microbes</td>
<td>644</td>
<td>376</td>
</tr>
<tr>
<td>Good microbes</td>
<td>575</td>
<td>391</td>
</tr>
<tr>
<td>Bad microbes</td>
<td>503</td>
<td>334</td>
</tr>
<tr>
<td>Microbes pathogens</td>
<td>497</td>
<td>406</td>
</tr>
<tr>
<td>Heading</td>
<td>426</td>
<td>283</td>
</tr>
<tr>
<td>Including</td>
<td>425</td>
<td>271</td>
</tr>
<tr>
<td>Senior</td>
<td>423</td>
<td>281</td>
</tr>
<tr>
<td>The chain of infection</td>
<td>401</td>
<td>307</td>
</tr>
<tr>
<td>Higiene das maos</td>
<td>398</td>
<td>364</td>
</tr>
<tr>
<td>Microbes utiles</td>
<td>368</td>
<td>232</td>
</tr>
<tr>
<td>Microbios</td>
<td>355</td>
<td>309</td>
</tr>
</tbody>
</table>

**Most common search keywords**

The majority of referrals were generated from search engines, with Google alone accounting for 56 520 page views. The referral keywords from all search engines were aggregated, with 18 588 different search phrases being used. Table 4 shows the top 20 search phrases used by visitors to find the e-Bug web site. The most common search phrases were related to e-Bug, with 11 482 page views originating from searches for ‘e-bug’, ‘Ebug’, ‘e-bug’, ‘www.e-bug.eu’ and ‘e-bug.eu’. The prominence of the search phrase ‘chain of infection’ is down to the fact that it represents a very specific combination of query terms and, therefore, e-Bug ranks highly on search engine results; unlike less specific query terms, such as ‘antibiotics’, which are covered by numerous other bigger web sites than e-Bug (e.g. Wikipedia).

Interestingly, swine flu and related terms were not one of the main search phrases, indicating that the increase in page views and downloads seen during April and May 2009 was primarily due to direct links from sites such as the HPA.

**Discussion**

The relative ease of collecting and analysing web server logs illustrates the importance of this technique to evaluate the success of a web site such as e-Bug and how the data can be shared amongst multiple partners in a project of this type.

Use of the site has increased markedly since the official launch of e-Bug in September 2009. The average number of page views per month between January 2009 and August 2009 was 28 553. Since September 2009, the average has increased to 209 625 page views per month. This is primarily due to the promotion that the site received during this period and, also, this coincides with the start of the school year.

The distribution of users from a number of European countries demonstrates the international success the e-Bug Project is having and the importance of providing localized versions of all of the resources. There have been visitors from 190 different countries, indicating that the demand for resources of this type is universal, not solely European.

The almost 200 000 downloads of Word documents and PowerPoint presentations for both junior and senior packs of materials, the majority occurring since September 2009, again illustrate the demand for a teaching resource for microbes, hygiene and antibiotics after the official launch and dissemination in countries. It is perhaps not surprising that two of the top three most popular Word documents were factsheets about swine flu, but this is an indication of the impact that an information resource directed at schools can have, as well as the importance of links from popular external sites, such as the HPA, during such outbreaks. It is also interesting that the top downloads were generally documents that provide information about handwashing and hand hygiene, indicating that these are topics about which people are currently looking for information.

The high number of page views originating from people using search engines demonstrates the importance of sites like Google and Bing in the success of a resource like e-Bug. Five of the top seven search phrases were related to the name of the site, e.g. searches for ‘e-bug’, ‘ebug’ and variations of the actual domain name www.e-bug.eu, showing that the Project is already well known. The other most popular search phrase referral was related to the ‘chain of infection’ and can be explained by the presence on the site of a PowerPoint resource targeting that subject. On the current Google UK web site, this document is ranked fourth for the search phrase ‘chain of infection’. Besides search engines, sites such as the HPA and NeL1 and the National Resource for Infection Control (NRIC) also demonstrate the importance of having links from popular sites and the numbers of visitors that these links can bring.

**Future work**

With the site receiving increased interest in the UK, and also as the resources are translated into more languages, it is expected...
that there will be a related increase in traffic to the site from an increasing number of countries. This will provide further data that can be utilized to study user behaviour on the site, with the potential for personalizing resources for specific groups of users, e.g. for particular countries, students, teachers etc. Future work will also look into the types of resource in which users are more interested and identifying which subjects can be developed further.

Acknowledgements

Members of the e-Bug working group include Herman Goossens, Niels Adriaenssens, Stijn De Corte, Julius Weinberg, Patty Kostkova, David Farrell, Jette Holt, Marianne Noer, Jenny Kremastinou, Koula Merakou, Dimitri Gennimata, Giuseppe Cornaglia, Raffaella Koncan, Pawel Grzesiowski, Anna Olczak-Pienkowska, Antonio Brito Avo and Jose´ Campos.

Funding

This work was supported by DG SANCO of the European Commission (2005211).

Transparency declarations

This article is part of a Supplement sponsored by the European Commission Health and Consumer Protection Directorate-General DG SANCO (grant number 2005211).

References