

## ENTOMOLOGICAL WEBSITE USAGE PATTERNS

R. W. MANKIN

United States Department of Agriculture, Agricultural Research Service  
Center for Medical, Agricultural, and Veterinary Entomology, P.O. Box 14565, Gainesville, FL 32604

## ABSTRACT

Usage patterns of entomological research websites were examined to assess their current roles as information resources. A 5-year review of logfiles at three Florida entomological research websites indicated that usage has increased since 1999 and that visitors increasingly have taken advantage of Internet search engines to find pages with high information content. Websites provide opportunities for dissemination of information (for example, in sound files or databases) that is difficult to include in traditional refereed publications. Given the rapid growth of website usage, research organizations may wish to consider formal procedures for vetting such information.

Key Words: Internet, world wide web, search engine

## RESUMEN

Los patrones de uso de páginas electrónicas para investigación entomológica fueron examinados para determinar sus papeles actuales como recursos de información. Una revisión de un periodo de 5-años de los registros de archivos de tres páginas electrónicas de investigación entomologica indica que su uso ha aumentado desde 1999 y los visitantes han tomado una ventaja creciente de sistemas de búsqueda de la Red-electrónica para encontrar páginas con un alto contenido de información. Las páginas electrónicas proveen oportunidades para la diseminación de información (por ejemplo, en un archivo de sonido o una base de datos) que son difíciles para incluir en publicaciones tradicionales reguladas. Tomando en cuenta el crecimiento rápido del uso de páginas electrónicas, las organizaciones de investigación tal vez quieren considerar los procedimientos formales para evaluar esta información.

Entomological institutions began establishing websites in the early 1990s to disseminate research and provide membership services (VanDyk 2000). The Florida Entomological Society (FES) was one of the pioneers in this effort, beginning online publication of new issues of the *Florida Entomologist* at the Florida Center for Library Automation (FCLA) in 1994 (Zenger & Walker 2000). Online journals at society and institutional websites have since become important resources for researchers, but few studies have been published on scientific website usage patterns (Davis 2004) or the roles of institutional websites as information resources (Treise et al. 2003; Lederbogen & Trebbe 2003).

In addition to the original online issues of *Florida Entomologist* in .pdf format, FES launched a membership website in 1998 which hosted a search engine for online issues, *FESsite* (Table 1), and began offering an .html version of new articles at BioOne ([www.bioone.org](http://www.bioone.org)) in 2002. A search engine for all *Florida Entomologist* issues was initiated at FCLA in 2004. To evaluate usage of current FES website resources in the context of activity at similar websites, a review was conducted of activity between 2000 and 2004 at *FESsite*, the United States Department of Agriculture, Agricultural Research Service, Center for

Medical, Agricultural, and Veterinary Entomology website, *CMAVEsite*, and a personal research site at CMAVE, *Perssite*. The usage was compared with activity at the *Florida Entomologist* journal homepage, the American Chemical Society journal server (Davis 2004), and two entomological websites at the University of Florida (UF). The two UF sites were the Entomology Department Newsletter and the highly visited, Featured Creatures site (Table 1), hosted jointly with the Florida Division of Agriculture and Consumer Services (FDACS).

## MATERIALS AND METHODS

## Website Logfiles

Daily logs from *FESsite* were analyzed from 5-Sep-99 to 31-Dec-04, using Mediahouse or LiveSTATS (Deepmetrix Corp., Gatineau, Quebec, CA). Daily logs from *CMAVEsite* were analyzed from 14-Jan-00 to 24-Dec-04, with Analog ([www.analog.cx](http://www.analog.cx), Cambridge, UK). Additional analyses were conducted on a subset of the CMAVE daily activity in a personal research site, *Perssite* (Table 1). File-editing records were examined and directory searches were conducted to count the numbers of pages on the sites at the be-

TABLE 1. DESCRIPTIONS AND INTERNET ADDRESSES (URLS) OF FREQUENTLY ACCESSED CMAVE, FES AND UF ENTOMOLOGY WEB PAGES, WITH ASSOCIATED SITE/PAGE NAMES.

Description	URL	Site/Page name
CMAVE website	cmave.usda.ufl.edu	<i>CMAVEsite</i>
CMAVE home	cmave.usda.ufl.edu/index.html	<i>CMAVEhpage</i>
IFAHI Unit home <sup>1</sup>	cmave.usda.ufl.edu/ifahi/index.html	
Formis database <sup>2</sup>	cmave.usda.ufl.edu/formis/	
CMAVE publ. list	cmave.usda.ufl.edu/publications.html	
Personal website	cmave.usda.ufl.edu/~rmankin	<i>Perssite</i>
Personal home	cmave.usda.ufl.edu/~rmankin/index.html	<i>Pershpage</i>
Sound library	cmave.usda.ufl.edu/~rmankin/soundlibrary.html	<i>Perssound</i>
FES website	flaentsoc.org	<i>FESsite</i>
FES home	flaentsoc.org/index.html	<i>FEShpage</i>
FlaEnt search <sup>3</sup>	flaentsoc.org/FEASearch.cfm	
FlaEnt home link <sup>4</sup>	flaentsoc.org/fe.html	
FES pests <sup>5</sup>	flaentsoc.org/fespestweblinks2.html	
FES Fla. insects <sup>6</sup>	flaentsoc.org/arthropodiversity/	
UF/FDAC featured creatures	creatures.ifas.ufl.edu/	
UF Entomology newsletter	entnews.ifas.ufl.edu/	

<sup>1</sup>CMAVE Imported Fire Ant and Household Insects Research Unit home.

<sup>2</sup>Bibliography of ant literature.

<sup>3</sup>Link to original search engine for 1993-2004 *Florida Entomologist* online issues.

<sup>4</sup>FES page with information and links to *Florida Entomologist* journal homepage (Fig. 1B).

<sup>5</sup>Links to information about Florida pest insects.

<sup>6</sup>Florida Arthropod Conservation homepage.

ginnings and ends of the analysis periods. The Florida Entomologist home page (<http://www.fcla.edu/FlaEnt/index.htm>) was monitored with Net.Data (IBM, Armonk, NY).

Requests for FES, CMAVE, and personal research website pages were also compared with records of monthly page views of the UF, IFAS Entomology and Nematology Newsletter and the UF, IFAS/FDACS Featured Creatures websites analyzed by LiveSTATS from 1-Jan-02 to 31-Dec-04.

#### Logfile Analysis Procedures

Logfile software typically provides information about the originators of requests, and the numbers of successful and unsuccessful requests for pages (files with .htm or .html extensions), as well as graphics, .pdf, .wav, or other files embedded in the pages (Srivastava et al. 2000). Unless otherwise specified below, the counts listed in this report refer only to successful requests for (.htm and .html) web pages and not directly to requests for embedded files.

CMAVE and FES website page monthly totals were estimated by summing the counts from all successful views of pages at the website in the 30 days preceding the count. To evaluate the contribution of specific pages to website totals, monthly page views also were counted individually for several frequently accessed FES, CMAVE, and personal research pages (Table 1). The Featured Creatures and Newsletter monthly totals were determined from daily counts.

Monthly totals were counted for page queries initiated by easily identifiable search engines, including Google, MSN, Yahoo, Ask Jeeves, Alta Vista, HotBot, and Lycos. Potential relationships between website page-view rates and search engine query rates were tested with Proc GLM (SAS Institute 1988).

## RESULTS

### Website and Home Page View Rates

Between 2000 and 2005, the rates of page views of the FES (*FESsite*), CMAVE (*CMAVEsite*), and personal research (*Perssite*) websites increased 2.8-, 6.7-, and 9.5-fold, respectively (Fig. 1A), or 0.6-, 1.3-, and 1.9-fold per year, respectively. The greatest rate of increase was seen at the Featured Creatures website which increased 6.3-fold between January 2002 and December 2004, or 3.1-fold per year. The activity at FES and personal research website pages in 2003-2004 was similar in magnitude to the 3316 monthly queries of the American Chemical Society Server by chemists at Cornell University in 2002-03 (Davis 2004). Apart from the homepage, the *Florida Entomologist* search engine was the most frequently viewed page at *FESsite*. The insect sound library (Leslie 2002) was the most frequently viewed page at *CMAVEsite*. These two pages are considered in greater detail below.

The general trend of increasing rates of page views was modulated by a yearly cycle. As expected with smaller audiences at holidays, the

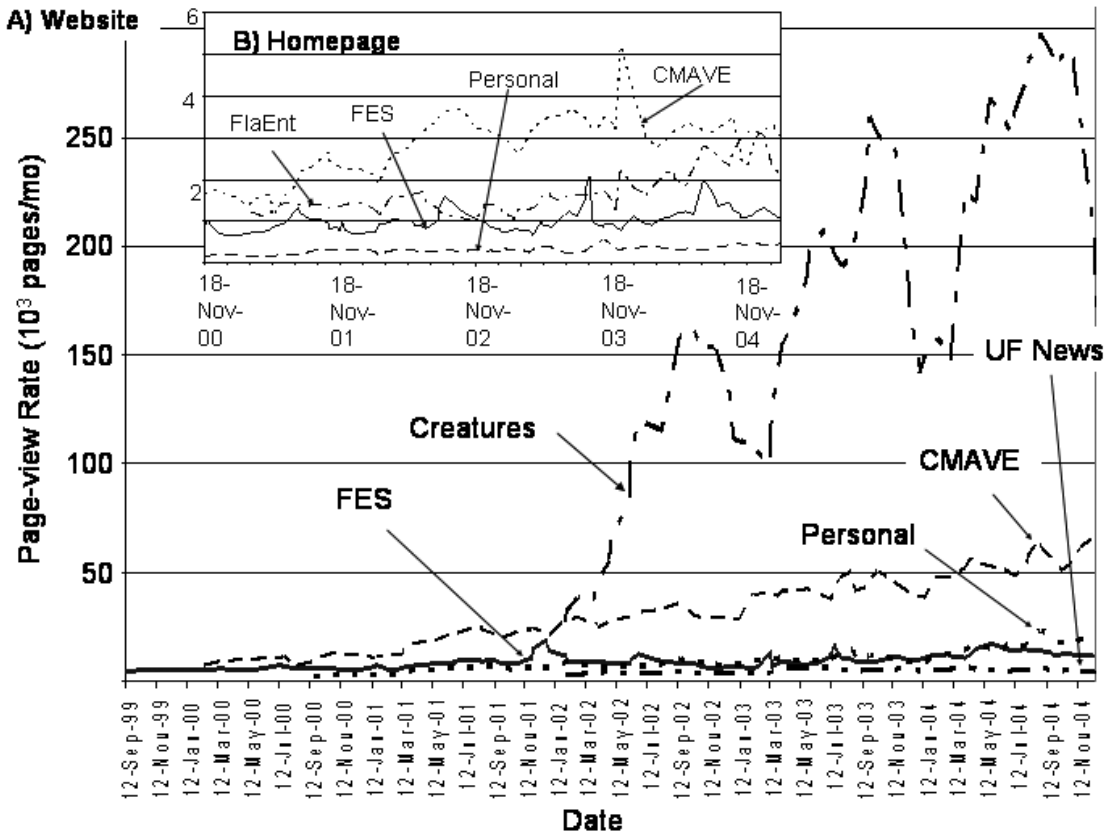


Fig. 1. A) Numbers of Featured Creatures (dot-dashed line), CMAVE (dotted line), FES (solid line), UF Entomology newsletter (dot-dot-dashed line), and Personal website (dashed line) pages viewed monthly between September 1999 and December 2004 (page addresses in Table 1); B) Numbers of CMAVE, FES, Personal, and FlaEnt journal (dot-dashed line) homepages viewed per month between September 1999 and December 2004 (page addresses in Table 1).

view rates at all of the entomological web sites decreased briefly each year in December and January. Increased rates of page views were observed at *FES* site in each year in June and July, probably due to activities involving the annual meeting. Increased rates of viewing at the Featured Creatures website occurred in August and September,

near the beginning of the elementary and secondary school year.

Part of the increased activity between 2000 and 2005 at the *FES* and *CMAVE* websites could have been due to an increase in the numbers of web pages at each site, but the percentage increase in page-view rates exceeded the percent-

TABLE 2. REGRESSION EQUATION STATISTICS FOR RELATIONSHIPS BETWEEN PAGE-VIEW RATES AND SEARCH ENGINE QUERY RATES, LISTED BY DECREASING  $r^2$ .

Name	$F^1$	$r^2$	Root MSE <sup>2</sup>
<i>Perssite</i>	161.06	0.797	2586.48
<i>FESsite</i>	143.13	0.777	1722.06
<i>CMAVEsite</i>	131.37	0.762	8032.98
<i>Perssound</i>	60.89	0.596	672.64
<i>Pershpage</i>	48.24	0.541	68.14
<i>FEShpage</i>	23.10	0.360	275.28
<i>CMAVEhpage</i>	13.42	0.247	686.24

<sup>1</sup> $df = 1,41$ ,  $P < 0.001$  for all variables except for *CMAVEhpage*, with  $P = 0.007$ .

<sup>2</sup>Mean Square Error.

TABLE 3. REGRESSION COEFFICIENT ESTIMATES AND STANDARD ERRORS (SE) FOR RELATIONSHIPS BETWEEN PAGE-VIEW RATES AND SEARCH ENGINE QUERY RATES.

Name (unit)	Estimate	SE	P
$A_{CMAVEsite}$ (pages)	13550.130	2299.41	<0.0001
$B_{CMAVEsite}$ (pages/query)	9.750	0.85	<0.0001
$A_{CMAVEhpage}$ (pages)	2202.080	196.43	<0.001
$B_{CMAVEhpage}$ (pages/query)	0.267	0.073	0.007
$A_{FESsite}$ (pages)	1497.300	626.68	0.0216
$B_{FESsite}$ (pages/query)	6.700	0.56	<0.001
$A_{FEShpage}$ (pages)	589.100	100.10	<0.0001
$B_{FEShpage}$ (pages/query)	0.430	0.09	<0.0001
$A_{Perssite}$ (pages)	2141.320	722.57	0.0050
$B_{Perssite}$ (pages/query)	12.830	1.01	<0.0001
$A_{Perssound}$ (pages)	590.070	187.92	0.0031
$B_{Perssound}$ (pages/query)	2.050	0.26	<0.0001
$A_{Pershpage}$ (pages)	203.070	19.04	<0.0001
$B_{Pershpage}$ (pages/query)	0.180	0.03	<0.0001

age increase in web pages. The numbers of web pages increased only from 136 to 619 (455%) at *CMAVEsite*, 85 to 188 (221%) at *FESsite*, and 32 to 84 (262%) at *Perssite*.

In contrast with the website activity, there was less evidence of a trend in the rates of views of FES, CMAVE, and personal research homepages (Fig. 1B). However, there was an increased rate of

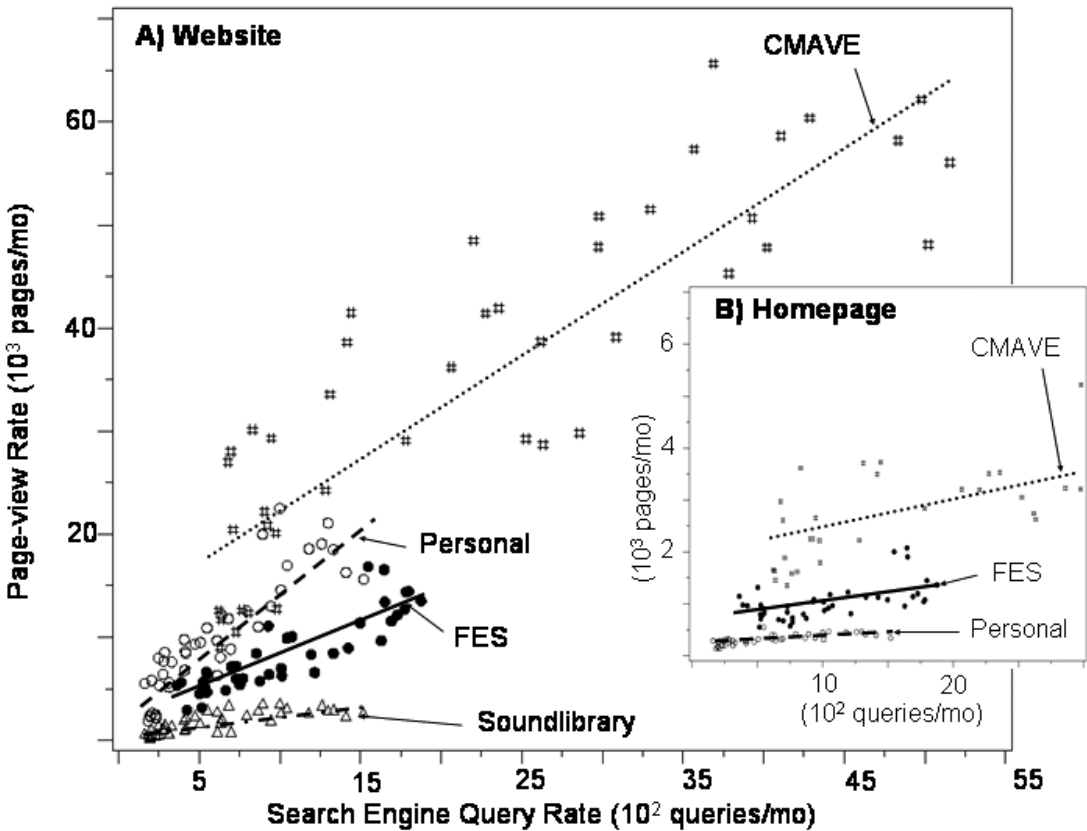


Fig. 2. Comparison of A) website and B) homepage view rates with search engine query rates for CMAVE (hashes, dotted lines), Personal (open circles, dashed lines), FES (dots, solid lines), and Insect sound library (triangles, dot-dashed line), websites.

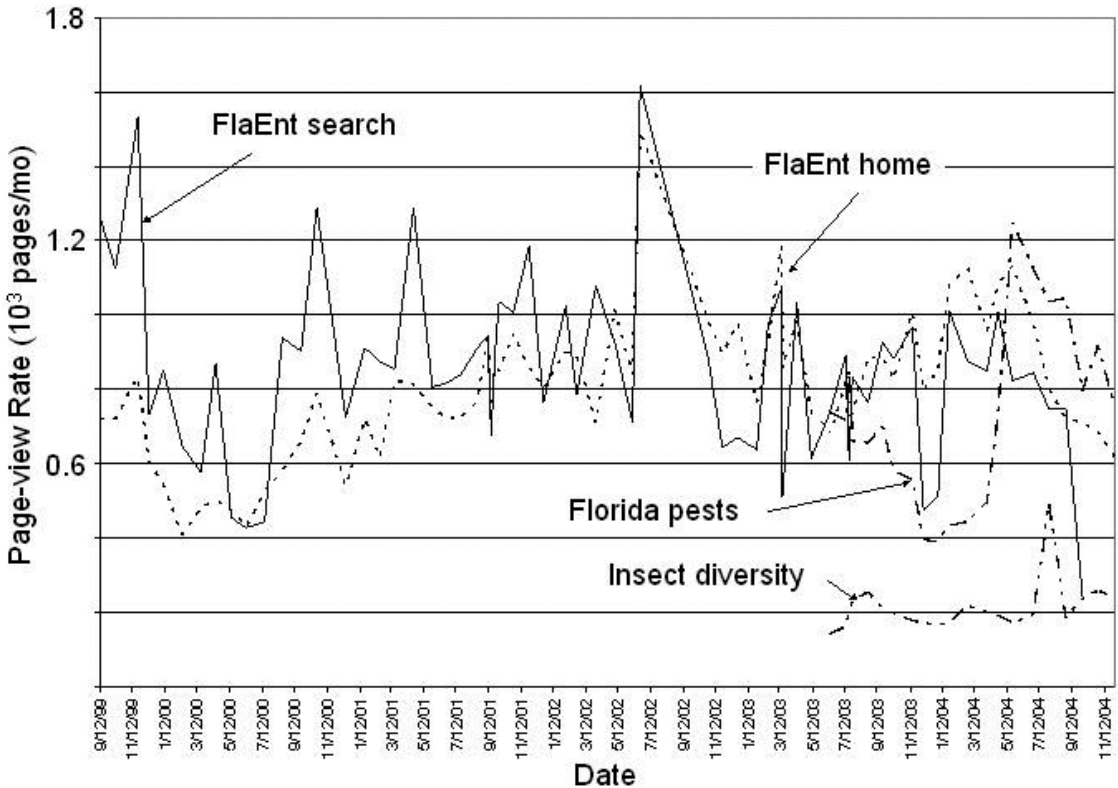


Fig. 3. Rates of viewing of frequently accessed pages at *FES*site: *Florida Entomologist* text-search engine (solid line), *Florida Entomologist* information page (dotted line), Florida pest control methods (dot-dashed line), insect diversity (dot-dot-dashed line).

viewing of the *Florida Entomologist* journal homepage during 2004 (Fig. 1B) similar to the website trend in Fig. 1A. The difference between the usage patterns for *CMAVE*site, *FES*site, *Pers*site, and the corresponding homepages suggests that the behavior of visitors at those three sites (but not at the journal) may have changed over time. If a typical user enters a website through the homepage and then visits five or fewer pages altogether (Cooper 2001), about 15% or more of the pages viewed would be expected to be homepages. Instead, the percentage of homepage views at the *CMAVE*, *FES*, and personal websites had declined below 8% by August 2004. Perhaps, experienced visitors were bookmarking pages of interest in one session and then returning to the bookmarked pages directly rather than through the homepage. Alternatively, visitors may have accessed multiple pages directly from a search engine, in which case the rates of website page views might be reflected in search-engine query rates.

#### Comparisons of Page-Views and Search Engine Queries

The hypothesis that the rates of viewing of pages at *CMAVE*, *FES*, and personal research

websites between 2000 and 2005 were proportional to search engine query rates was tested under the model:

$$\text{PageViewRate} = A + B \text{ QueryRate},$$

where *PageViewRate* is the number of page views per month at a specified website and *QueryRate* is the number of search engine requests per month for pages at that site. The regressions for the *CMAVE*, *FES*, and personal websites, the corresponding home pages, and the insect sound library page are compared in Table 2. The regression coefficients are listed in Table 3 and the relationships are graphed in Fig. 2. The coefficients of determination for website page-view rates and search engine query rates were larger than for the home pages, but all of the regressions in Table 2 were statistically significant.

The slopes of the regression equations for *CMAVE*site, *FES*site and *Pers*site (Table 3, Fig. 2A) were all greater than 6 views per query, while the slopes of the regression equations for the corresponding home pages (Fig. 2B) were all less than 0.5 views per query. The large difference in the slopes of the website and homepage page-

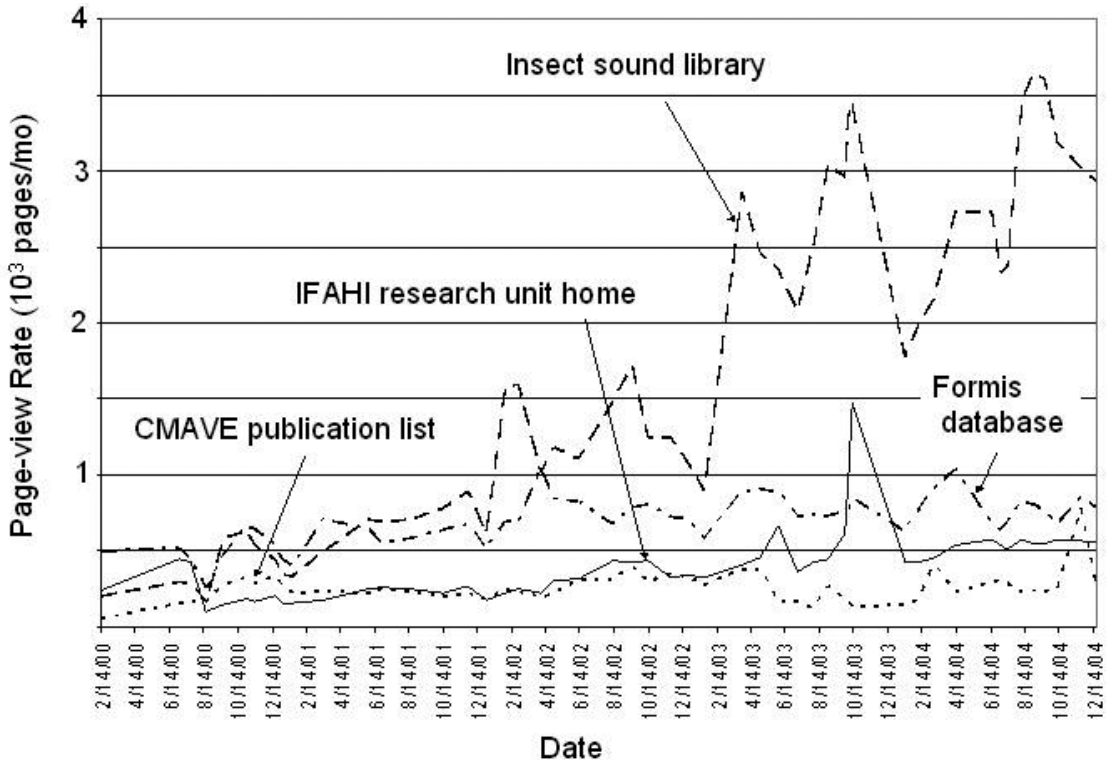


Fig. 4. Rates of viewing of frequently accessed pages at *CMAVEsite*: Insect sound library (dashed line), IFAHI research unit home (solid line), Formis database of ant literature (dot-dashed line), CMAVE publications list (dotted line).

view-rate regressions suggests that many of the viewers had queries in search for specific information present at the site rather than for general information about the hosting institution. Such information most likely would be present in high-content pages rather than homepages. An analysis of frequently viewed pages was conducted to determine whether they were characterized by high levels of information content.

#### Analysis of Frequently Viewed Pages

Page-view rates for the four most frequently viewed pages other than the website homepages in 2004 are shown for FES in Fig. 3 and CMAVE in Fig. 4. It should be noted that the *Florida Entomologist* search engine page was replaced with an improved version at the FCLA website, which reduced the rate of website page views (Fig. 1B) after September 2004. The introductory page to the *Florida Entomologist* has been viewed frequently since the FES website began, and two pages initiated in 2003 that contain information about Florida pests and Florida arthropod ecology

have been viewed at increasing rates in 2004. These, together with the annual meeting program and abstract pages, contain a large fraction of the scientific information available at *FESsite*.

The most frequently accessed pages at *CMAVEsite* other than the homepage in 2004 were the insect sound library, the Formis database, the IFAHI Research Unit home page, and the CMAVE publication list (Fig. 4). The sound library contained information that is not easy to include in a traditional refereed publication, although increasing numbers of journals allow for online posting as supporting online material. The Formis database is a popular source of bibliographic information on ant literature. The CMAVE website has a larger number of high-content pages than *FESsite*, which possibly contributed to its higher rate of page-views in 2004 (Fig. 1A).

#### DISCUSSION

The patterns described above suggest that increased usage of Internet search engines has se-

lectively increased the rates of viewing of specific FES and CMAVE web pages. Overall, page-view rates have increased, but the greatest increase has occurred for sites and pages with high information content. Search engines typically assign high visibility to research pages from scientific institutions (Jepsen et al. 2004). This has enabled entomological databases, associated software (Byers 2002), and sound files to become important, easily accessible research tools along with the refereed literature. The popularity of the Featured Creatures website may also have benefited from its high ranking in search engines.

The trend of increasingly data-rich scientific websites is not confined to entomology. Research in bioinformatics (Eiden 2004), biodiversity (Maddison & Schulz 2004) and astrophysics (Kurtz et al. 2005), for example, has become highly dependent on digital libraries and online databases. The rapid growth of information indexed by Internet search engines has enabled researchers to modify their search strategies and accelerate their rate of gathering information (Davis 2004). It has also enabled them to disseminate information faster to wider audiences. The improved visibility of Internet-accessible research has resulted already in increased relative effect of open-access research articles published online (Antelman 2004). If such trends continue unabated, researchers could benefit if institutions developed new procedures to vet website content, perhaps similar to the peer review process of scholarly journals. One of the challenges to such development is that web content is frequently updated and, unlike a journal article, would need to be reverted after a major change.

#### ACKNOWLEDGMENTS

Tom Fasulo (University of Florida) provided LiveSTATS records for monthly page views of the Featured Creatures and Entomology Department Newsletter websites. Tom Walker (University of Florida), and Pat Greany, Paul Shirk, Eric Daniels, and Dianne Un-

derwood (CMAVE) have provided assistance and advice in the development of this manuscript.

#### REFERENCES CITED

- ANTELMAN, K. 2004. Do open-access articles have a greater research impact? *College and Research Libraries* 65: 372-382.
- BYERS, J. A. 2002. Internet programs for drawing moth pheromone analogs and searching literature database. *J. Chem. Ecol.* 28: 807-817.
- COOPER, M. D. 2001. Usage patterns of a web-based library catalog. *J. Am. Soc. Information Sci. Tech.* 52: 137-148.
- DAVIS, P. M. 2004. Information seeking behavior of chemists: A transaction log analysis of referral URLs. *J. Am. Soc. Information Sci. Tech.* 55: 326-332.
- EIDEN, L. E. 2004. A two-way bioinformatic street. *Science* 306: 1437.
- JEPSEN, E. T., P. SELDEN, P. INGWERSEN, AND L. BJORNEBÖRN. 2004. Characteristics of scientific web publications: preliminary data gathering and analysis. *J. Am. Soc. Information Sci. Tech.* 55: 1239-1249.
- KURTZ, J. J., G. EICHORN, A. ACCOMAZZI, C. GRANT, M. DEMEITNER, AND S. S. MURRAY. 2005. Worldwide use and impact of the NASA Astrophysics data system digital library. *J. Am. Soc. Information Sci. Tech.* 56: 36-45.
- LEDERBOGEN, U., AND J. TREBBE. 2003. Promoting science on the web. *Science Comm.* 24: 333-352.
- LESLIE, M. 2002. Catch a buzz. *Science* 297: 743.
- MADDISON, D. R., AND K.-S. SCHULZ (Ed.). 2004. The tree of life web project. <http://tolweb.org>.
- SAS INSTITUTE. 1988. SAS/STAT user's guide, release 6.03 edition, Cary, NC.
- SRIVASTAVA, J., R. COOLEY, M. DESHPANDE, AND P.-N. TAN. 2000. Web usage mining: discovery and applications of usage patterns from web data. *SIGKDD Explorations* 1: 12-23.
- TREISE, D., K. WALSH-CHILDERS, M. F. WEIGOLD, AND M. FRIEDMAN. 2003. Cultivating the science internet audience. *Science Comm.* 24: 309-332.
- VANDYK, J. K. 2000. Impact of the internet on extension entomology. *Annu. Rev. Entomol.* 45: 795-802.
- ZENGER, J. T., AND T. J. WALKER. 2000. Impact of the internet on entomology teaching and research. *Annu. Rev. Entomol.* 45: 747-767.