

Wired Research: Transaction Log Analysis of E-Journal Databases to Assess the Research Activities and Trends in UAE University

By Ahmed Taha

Library Specialist for Research and E-information Services,
UAE University Libraries Deanship, United Arab Emirates.

Abstract

E-journals come bundled with the powers of the Internet particularly the Web to revolutionise the entire scholarly communication process and changed the way that the researchers seek and exchange scientific information. The UAEU library is currently going through a paradigm shift; switching from paper-based to digital formats. This work aims at better understanding the pathways taken when the researchers connect to favourite e-journals by analysing the generated transaction logs. Examples on electronic uses were derived from ScienceDirect™ database transaction logs of the articles that appeared in 2003. The data analysis revealed that the sheer predominance of electronic access to e-journals was in the domains of environmental, materials, medical, and business studies since these areas are gaining a significant priority to UAEU research funding and graduate programmes. The implication of the choice and heavily use of favourite e-journals was discussed with particular reference to the future of e-information services, and subscription to core e-journals. This study argues that the emergence and fit of e-journals in the library e-services has significantly improved efficiency of information practices as well supported a range of research activities and trends.

Introduction

As the needs and expectations of library users changed in the networked environment, libraries are diligently endeavouring to find the best ways to define their user communities, understand what they value, and evolve networked library electronic collections and services to meet their growing demands for cultural and scholarly information. Web-based e-resources are emerging as a dominant paradigm for constructing and composing distributed scholarly information applications and enabling library-

wide interoperability. As one of the most important forms of scholarly communication, scholarly journals are also undergoing a process of change to be constructed as digital communication forums. In the academic arena, proliferation of online bibliographic databases and e-journals as media for scholarly publications has revolutionised the ways that the researchers conduct information searching and communicate their research findings.

The web-based e-resources however facilitate the exploration of user behaviour far beyond what is

possible in a print environment as the web-based services can be configured to log all relevant user transactions that occurs on the e-journal database systems. By analysing these transaction logs, it is made possible to obtain a detailed picture of what users are doing and how they use a service. Most of the studies have focused on online public access catalogues (OPAC) and traditional IR systems in order to identify the extent to which specific features are used; moreover, to identify general pattern of system used (Banks 1999). However, many further studies have recently been conducted on analyses of web log files of e-journals (e.g. Sullenger 1997).

The analysis of user logs from networked libraries offers new opportunities to assess research trends and practices in an institution's user communities. Journal websites could play a critical role in scholarly communication for three reasons:

- i. increasing use of the web as an information source both inside and outside academia;
- ii. centrality of journals in disseminating scientific research; and
- iii. astonishing increase in the number of the scholarly journals available on the Web in the last two years, including both the new only e-journals and traditional print journals having online versions (Vaughan and Thelwell 2003)

The work analysed the transaction logs generated from the e-journal databases Elsevier ScienceDirect, which covers the period of January to December 2003. The goal is to get better understanding of the pathways taken when the researchers connect to their interesting e-journals to predict the research trends and activities conducted in UAE university for proposing an acquisition strategy focusing on development of relevant digital collection; in particular, e-journals.

Transaction Log Analysis: *pros and cons*

Transaction log analysis (TLA) was developed about twenty-five years ago to evaluate the database performance. It has been further evolved as a

method used to study unobtrusively interactions between the online-accessed database systems and the end-users for obtaining some insights into users' information seeking behaviour, and evaluation of query interfaces (Peters 1993; Convey 2002). The information recorded is stored in an electronic file referred to as transaction log. The contents of transaction logs are usually formatted in fields to facilitate quantitative analysis. Analysis of transaction logs data enhances had better understand of how the users conduct search on online information systems or web sites with the intention of improving their design and functionality to meet user needs and expectations.

TLA is thus an effective method to study such activities as the frequency and sequence of feature use; system response times; hit rates; error rates; user actions to recover from errors; the number of simultaneous users; and session lengths (Convey 2002). The networked library, however, has gained the most benefits from TLA for i- effective evaluation and improvement of e-services, ii- Identify user communities, iii- Identify patterns of e-resources use, iv- inform the redesign and development of the library Web portal, and v- assess end-user satisfaction. On the other hand, the primary disadvantages of TLA are that extracting data can be time-consuming and the data can be difficult to interpret.

E-Journal TLA Related Literature

With the widespread use of the Web-based e-resources, there has been some research related to the analysis of Web log files of e-journals. Zhang (1999) has conducted a study on the evaluation of electronic journals and monitoring their use patterns as revealed in the web log files of the accessed e-journals. Eason *et al.* (2000) have examined the general value to users of a range of electronic journal functions and their usefulness in the specific context of the SuperJournal Project in light of generated logged data of usage. Jones *et al.* (2000) has conducted a transaction log analysis on user activity gained insights into query and browsing patterns.

Transaction log analysis has been used to obtain insight into user behaviour for decades. Ke et al. (2002) have analysed transaction logs of the Taiwan ScienceDirect OnSite (SDOS) E-journal system, and interpreted the results to track the use pattern and information seeking behaviour of the end-users. Banks (1999) has discussed the usefulness of using transaction logs data for resource scheduling of the reverence services. Spink et al. (1998) have investigated the nature, manifestations, and behaviour of successive searching conducted by end-users in digital environments by means of TLA to derive criteria for use in the design of information retrieval interfaces and systems to support convenient access and search tactics.

UAEU Library Networked Information Environment

The UAEU Library is operating in a dynamic multicultural society, which is of positive perception to

information and communication technologies. Library mission as stated in its strategic plan is “*To build, organise, and provide access to an enduring collection of recorded knowledge and information; to support the needs of the teaching, learning, and scholarly research; to serve in a leadership role with other academic libraries*”. To achieve this, it has placed great emphasis on the acquisition of innovative library technology of its day since 1992. Currently, the library is wireless-networked throughout the five university campuses; concomitantly developed interactive web portal (<http://www.libs.uaeu.ac.ae>), which provides vastly access and several options to search and reach a wide variety of networked electronic resources and services. To pursue advanced e-information service, the library launched in 2003 EZProxy™ portal (<http://ezproxy.uaeu.ac.ae>) providing the virtual users efficient functionality and ease-of-access to its databases menus.

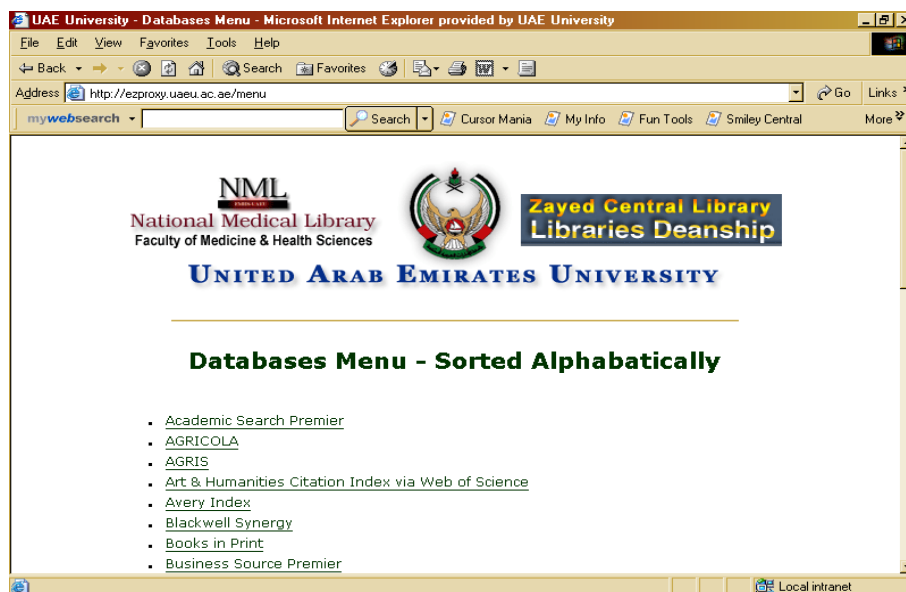


Figure 1: Databases Menu (15 textual and 20 bibliographic)

Migration to E-Journal Holdings

During the last four years, the UAEU library has made further improvements toward achieving its goal of empowering users through advanced library and information technologies. The UAEU Library possesses hybrid journal holdings to offer an increasing range of electronic journals as a result of the advent of web-based electronic resources into the library setting. This moving further towards the provision of e-journal services was attributed to many factors converged to make such rapid and peaceful transition possible and right, among these were:

1. Strong support from higher university administration
2. Adequate networking and IT infrastructures in the place
3. Ready-to-access large number of scholarly e-journals from remote areas
4. The urgent need to support growing the course websites as an effective learning tool, and scholarly research activities

The major transition period initiated in 2002 to decline p-journal subscription and compensated by increasing subscriptions to online-accessed journals. It was through three renewal cycles when the major cancellation of p-journal subscriptions took place as seen in Fig.2

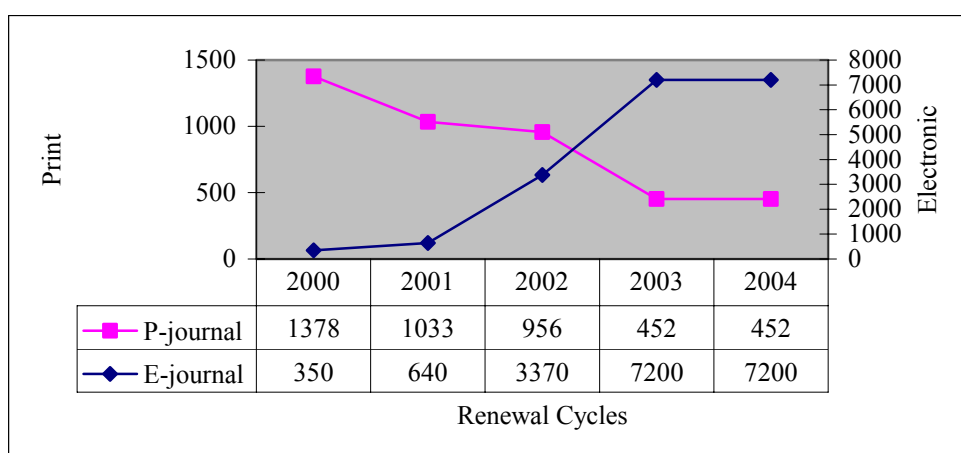


Figure 2: Migration to E-journal subscriptions

Scholarly Research Activities in UAE University: an overview

In order to develop research activities of the UAE University to international standards, research must be an integral part of the academic life of both Faculty and students. To achieve the goals established for accountable and competitive research, research infrastructures have been developed with the guidance of the Scientific Research Sector (SRS). Research laboratories have been equipped with the most up-to-date facilities and equipment. Concurrent with these activities, the SRS has developed mechanisms to evaluate and fund research proposals

and to recommend allocation of research personnel, equipment, etc. All faculty members are eligible to apply for financial support for research projects. These applications are peer-reviewed according to international standards; the evaluation includes additional criteria such as the relevance of the proposed projects to UAEU research priorities; the research record of accomplishment of the applicant; resources needed for pursuing the stated objectives; and the potential involvement of graduate students or collaboration with overseas partners.

In addition to generous SRS support, the research faculty are actively pursuing alternative sources to

support their research including international organisations, universities, companies, etc. along with these research activities, there are four MSc degree programmes; these are on i- environmental studies (interdisciplinary), ii- materials science and engineering, iii- water resources, and iv- MBA. These programmes were modelled on the American system of graduate studies. Some of the dissertation proposals submitted to graduate programmes often generated from SRS projects. At present, 86 research projects funded by SRS are being developed principally in the following areas:

1. *Engineering Sciences*: projects largely focusing on the petrochemical and related studies; environmental and civil engineering; materials and polymer technologies; biotechnology; and water resources.
2. *Environmental Studies*: these are interdisciplinary projects mostly involved graduate students. The fields of interests are arid land ecology; wild life and conservation; marine biology; medical zoology; and environmental microbiology.
3. *Economic and Business Studies*: the project mainly dedicated to investigate financial and economic conditions of the UAE with focusing on e-commerce and e-business.
4. *Education*: the projects mostly oriented to investigate the potential impacts of educational technologies on schooling performance of the students; curricular development also takes its share.
5. *Social studies*: the projects give special concern to UAE social problems related to expatriated workforce; social pathology; population structure, ethnic conflicts, family affairs, cultural heritage, etc.

| College | Research Projects | Fund in US\$ |
|---------------------------|-------------------|------------------|
| Engineering Sciences | 28 | 373,650 |
| Science | 30 | 398,200 |
| Food Systems | 10 | 142,252 |
| Economics/Business | 8 | 132,100 |
| Education | 5 | 50,900 |
| Social Studies/Psychology | 5 | 41,700 |
| Total | 86 | 1,138,802 |

(US\$1.00 = 3.67 AED)

Table 3: UAEU financial support to scientific research during academic year 2002/2003 (adapted from Al-Faris, 2003)

Data Collection and Processing

The analysis of user logs from large-scale networked library offers new opportunities to assess research trends in an institution's user communities. Transaction logs on electronic use were derived from *ScienceDirect* database. The entries for full-text journal viewings were filtered to include only the articles requested between January to December 2003, and then sorted and subtotaled by title. The logs were carefully examined to ensure that there was little or no double counting of articles. The selection of *ScienceDirect* e-journal databases was attributed to the fact that:

- i. it hosts the high ranked and most heavily used journals frequently reviewed by both the students and the professional researchers.
- ii. *ScienceDirect* is usually being taken as an example in many training sessions devoted to scholarly use of e-journals for research purposes.
- iii. special emphasis was paid to analysis of download behaviour and frequency of full-text which enhanced significantly in better understanding of research patterns, as well as in the detection of research trends and activities.

Elsevier ScienceDirect™

- Search capabilities: *easy way to retrieve*

ScienceDirect has evolved successfully to be one of the largest and most heavily used web-based full-text science, technology, and medicine e-journal database worldwide. It hosts the bibliographic information and full-text articles of more than 1,861 journals published by Elsevier Science and co-

partners to access more than five million downloadable articles. Recently it empowered its user with powerful scientific search engine *Scirus™*. Its search portal provides the users with a variety of search options as quick, basic, and advanced; search can be conducted on all sources, on subject groups, or on favourite journals. The UAEU library introduced *ScienceDirect* in January 2003, and is eligible for accessing the full-text articles, which published online in 1998 and onward.

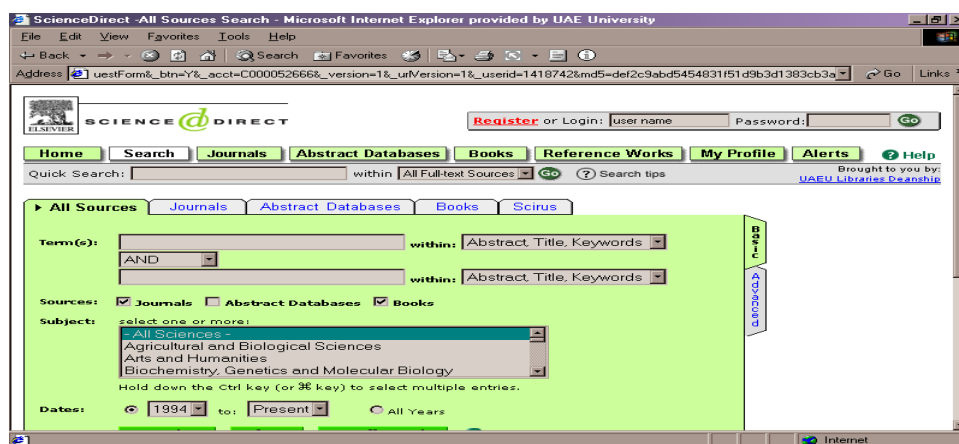


Figure 3: ScienceDirect search portal

- Query Submission: *options for all levels*

The full-text retrievals were performed by the three search strategies, which to the large extent reflect the search literacy of an end-user. A total of 26,028 queries have been submitted to the database in

2003, about 17,784 (68.33%) were executed by means of quick search option; basic search was opted for 8,006 queries (30.76%), whereas the expert advanced search processed the least number of queries to as 238 (0.91%) as seen in Fig.4.

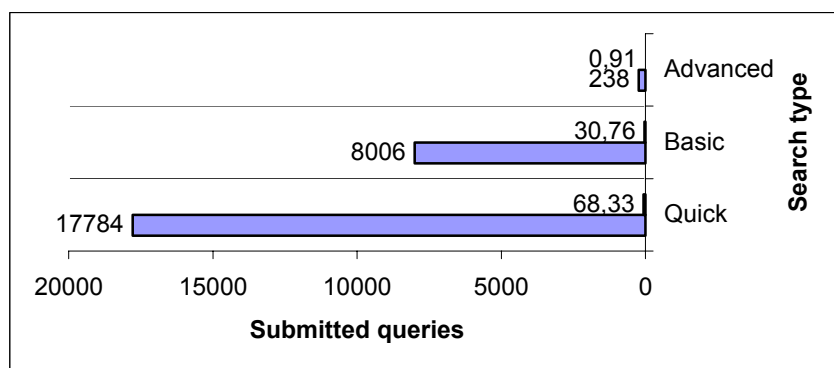


Figure 4: Illustrates queries options

- Full-text Viewing: *format preference*

ScienceDirect supports HTML, PDF formats for viewing, and downloading the articles of interest in full-text, which is the most active usage category because it hosts a large number of scholarly journals that scored high impact factor. The UAEU library's users have viewed in 2003 about 69,571 full-text articles; PDF has been opted to view 54,039 articles, amounting to 77.67% of total access, whereas HTML has been employed in 22.33% of total access to view 15,532 articles. PDF preference might be because an article in PDF is friendly printable, easy to be saved, and uploaded as an e-mail attachment; however, HTML was preferred in copy/paste images and illustrations.

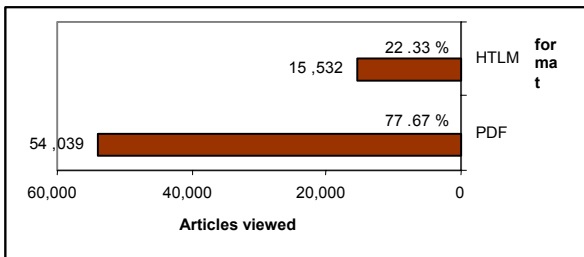


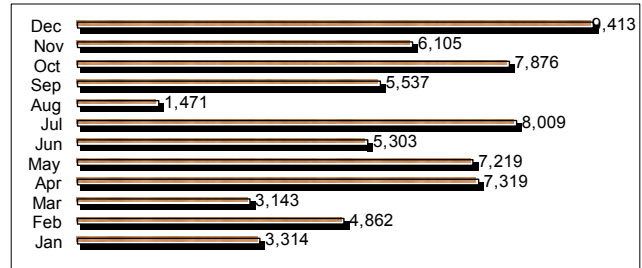
Figure 5: Articles viewed in HTML and PDF formats

- Download Behaviour: *the heavy use months*

The period of analysis (Jan-Dec. 2003) witnessed a great variability in searching intensity. Two periods have recorded distinct heavy uses of searching and retrieving transactions:

- April to July (second semester): when the graduate students and research project principals gathering relevant literature prior to submitting their scholarly works e.g. dissertations, technical reports,

- September to December (first semester): the newly registered students to MSc degree programmes consulted the database to retrieving support information for literature review, whilst the faculty used the retrieved information for preparing curriculum and research proposals.
- August is the summer vacation to record the lowest rate of viewing, as seen in Fig.6.



(n=69,571)

Figure 6: Ddepicts the number of request for full-text articles per month (2003)

- Download Behaviour: *subject-wise*

As *ScienceDirect* adequately supports the researches in the fields of natural science, medicine, engineering, and food science in particular; and business and social studies in general, it was not surprised that a large segment of the reviewed articles belonging to engineering (20.8%), natural sciences (collectively 34.7%), environmental studies (11.5%), food science (12.8), and medical (8.2) as seen in Fig.7. This data gives a practical tool for predicting the trends and intensity of activities in different fields of research i.e. natural sciences monopolised 88% of articles viewed, the rest 12% taken by social sciences.

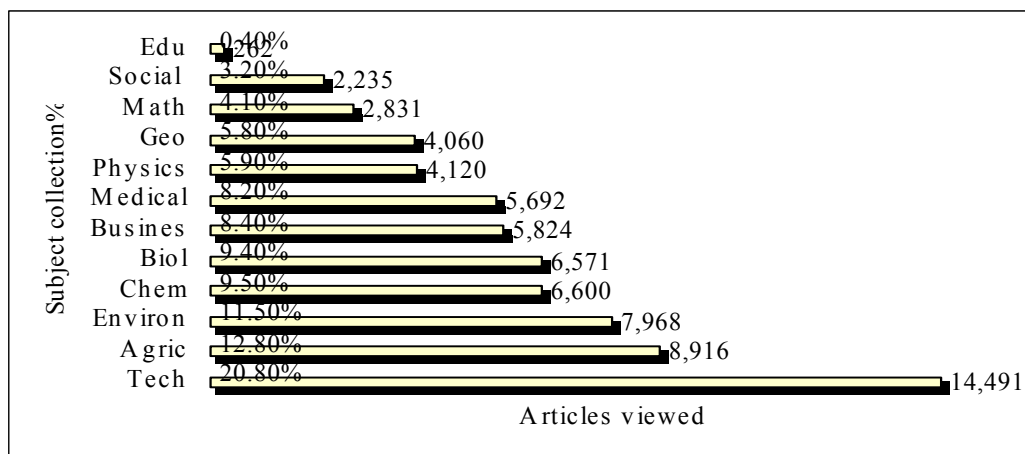


Figure 7: Depicts articles viewed per subject

- Transaction Categories: *user choice*

Two categories of transaction logs defined the user choice while conducting online search and viewing the returns: article abstract and Full-text access. Full-text access was found to be the dominant activity of searching to perform 69,571 views out of 82,861 (84%); whereas abstract viewing prior to assessing the relevance of the article was received very weak attention to view only 13,290 records (16%).

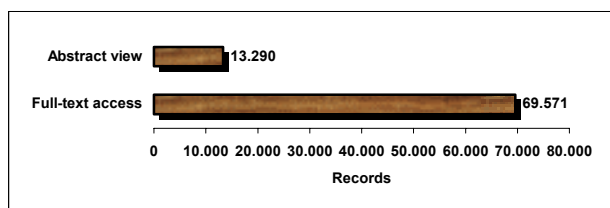


Figure 8: Shows view categories of transaction logs

- Download Tendency: *subscribed vs. unsubscribed holdings*

ScienceDirect™ allows its subscribers to access to the full database e-journals collection, including full-text access to e-journal titles that were not subscribed to its equivalent print version. The derived transaction logs revealed very interesting searching

phenomenon as the database users largely have targeted unsubscribed titles; for instance, more than 500 downloads were retrieved from fifteen titles each; eight of them were unsubscribed in print; approximately, 54,000 downloads were retrieved from 1,233 unsubscribed journals, whereas about 15,000 from subscribed titles. On the other hands, TLA revealed that nothing were downloaded from 440 titles available in ScienceDirect™ as seen in Fig.9. Most of these journals were irrelevant to research interest of many research faculty and graduate students.

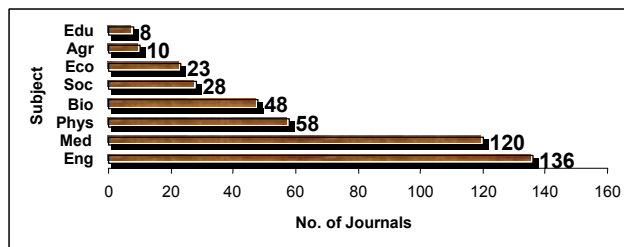


Figure 9: Depicts the zero-used journals per subject

Some peculiar download behaviour was predicted; for instance, some user heavily downloaded from journals of their interest at once in one search session, and not repeated again. Such behaviour will give a deceptive picture for download activities and will not help in assessing the research trends as well

as in proposing a rational policy for e-journals acquisition.

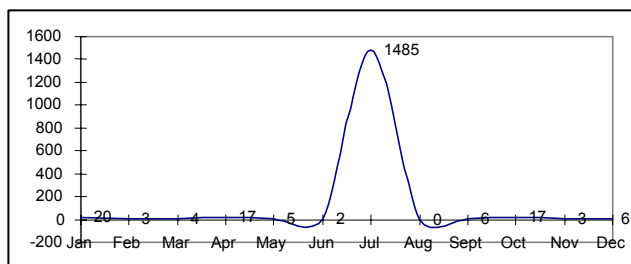


Figure 10: Illustrates download pattern (Cement & Concrete Research ISSN: 00088846)

Discussion and Conclusions

The objectives of the study were threefold:

- i. To describe the emergence role and fit of e-journals in library networked information services for research projects at UAE University,
- ii. To gain some insight into the range and pattern of e-journal usage researchers,
- iii. to predict and understand the impact of e-journal on research activities and trend.

The findings of the transaction log analysis revealed that the researchers at UAEU regularly consulted and use a limited number of e-journals available in *ScienceDirect*TM database, and seem to be unaware of availability of e-journals in their disciplines are hosted in other databases. The most striking observation to emerge from the TLA was that the sheer predominance of electronic access to *ScienceDirect*TM e-journals was found to be in the fields of environmental, materials, medical, engineering, and business studies since these areas are believed to have significant priority to the UAE university research funding graduate studies. In conclusion, the findings of this study can be used efficiently in collection development of core e-journals and identify areas for improvement; which ranging from user interface and functionality to product training. Promotion, awareness, and e-literacy may all be incentive factors to consider maximising usage of library e-resources. Analysis of usage of such online sys-

tems can provide valuable data on information-seeking behaviour of researchers in context of networked information environment.

Acknowledgement

I am immensely grateful to my colleagues for fruitful exchange of intellectual ideas, which enriched the manuscript, and warmly appreciated my wife Muzna and my elegant twins Amr and Omran for keeping long the patience and cheer me on to complete the task.

Cited References

Al-Faris, A-R. F., 2003, Research centres and decision making in the United Arab Emirates, *Journal of Development and Economic Policies* (Kuwait), vol.5, no.3, pp 113-137.

Banks, J. 1999, Can transaction logs be used for resource scheduling? An analysis, *Reference Librarian*, vol.63, no.1, pp 95-108.

Bollen, J., and Luce, R., 2002, Evaluation of digital library impact and user communities by analysis of usage patterns, *D-Lib Magazine*. Available at: <http://www.dlib.org/dlib/june02/bollen/06bollen.html> (viewed 28/2/2004)

Bollen, J.; et al., 2003. Usage analysis for the identification of research trends in digital libraries, *D-Lib Magazine*. Available at: <http://www.dlib.org/dlib/may03/bollen/05bollen.html> (viewed 28/2/2004)

Convey, D. T., 2002, *Usage and usability assessment: library practice and concer*. Washington, DC, Digital Library Federation. Available at: <http://www.clir.org/pubs/reports/pub105/contents.html> (viewed 14/2/2004)

Eason, K.; et al. 2000, Use and usefulness of functions in electronic journals: the experiences of SuperJournal, *Program*, vol.34, pp 1-28.

Jones, S.; et al., 2000, A transaction log analysis of a digital library, *International Journal on Digital Libraries*, vol.3, no.1, pp 152-169.

Ke, H-R.; et al, 2002, Exploring behaviour of e-journal uses in science and technology: transaction log analysis of Elsevier *ScienceDirect* Onsite in Taiwan, *Library and Information Science Research*, vol.24, no.3, pp 265-291.

Peters, T.A., 1993, The history and development of transaction log analysis, *Library Hi-Tech.*,vol.11, no.1, pp 41-58.

Spink, A.; et al., 1998. Modelling user's successive searches in digital environments. Available at: <http://www.dlib.org/dlib/april98/04spink.html> (viewed 18/2/2004)

Sullenger, P., 1997, A serials transaction log analysis, *Serials Review*, vol.23, no.1, pp 21-26.

Vaughan, L., and Thelwell, M., 2003, Scholarly use of the web: what are the key inducers of links to journal web sites?, *Journal of the American Society for Information Science and Technology*, vol.54, no.1, pp 29-38.

Zhang, Z., 1999, Evaluating electronic journals and monitoring their usage by means of WWW server log analysis, *Vine*, vol.111, no.1, pp 37-42.

Biographical note

Ahmed Taha
ALA, ASIST, MLA SLA
P. O. Box 1441, Al-Ain
United Arab Emirates
E-mail: Jazraby@uaeu.ac.ae;
Jazraby@mail2Uae.com