Current Awareness Service of the RePEc Digital Library:
Progress, Performance and Potentials

Heting Chu
Palmer School of Library & Information Science
Long Island University/C.W. Post Campus
720 Northern Blvd., Brookville, NY 11548, USA
Phone: 516-299-2177 Fax: 516-299-4168
hchu@liu.edu

Thomas Krichel
Palmer School of Library & Information Science
Long Island University/C.W. Post Campus
720 Northern Blvd., Brookville, NY 11548, USA
Phone: 516-299-2843 Fax: 516-299-4168
krichel@openlib.org

Abstract

NEP (New Economics Papers) is a current awareness service for the RePEc (Research Papers in Economics) digital library. Since its initiation in April 1998, over 50 individual lists have been created to loosely represent subfields within economics. Those lists in total made more than 37,000 announcements of about 28,000 working papers that were added to RePEc in the past five years. This article examines the growth and development of the NEP service. The performance of NEP is measured in terms of timeliness, coverage ratio, and usage. In exploring the various NEP parameters and their relationships, we discuss the potentials and other perspectives of the NEP service. Although it can be further improved, NEP could become an innovative model for current awareness services of digital libraries technically as well as organizationally.

Introduction

Digital libraries have been established in the past decade for various constituencies and communities. RePEc (Research Papers in Economics) at http://www.repec.org is one of them, specially created to facilitate the dissemination of working papers, journal articles and software objects in the field of economics. RePEc, a collaborative effort of over 100 volunteers in 30 countries, provides over 208,000 items of interest, over 107,000 of which are freely available online. The founders of RePEc have provided detailed descriptions about the digital library in their writings (e.g., Cruz & Krichel 2000; Krichel 2000).

Digital libraries have been researched and written about ever since their creations (Fox & Urs 2002; SchwartZz 2000). For instance, Davis and Lagoze (2000) illustrated the design and development of a digital library of computer science research reports

(NCSTRL) although it was later discontinued. Cole, Allen and Schmitz (2000) particularly explained the building of a digital collection of educational resources in agriculture at a university. Many more are published with regard to standards and related topics such as interoperability for digital libraries (e.g., Suleman & Fox 2001). Evaluation of digital libraries, the subject that the information science community is always keen on, received proper attention as well. Saracevic (2000), for example, discussed the challenges facing digital library evaluation and suggested a conceptual framework for evaluation derived from the systems approach. Choudhury et al (2002) described an evaluation framework for digital library services by focusing on an existing digital library project. More specific issues in evaluating digital libraries (e.g., usability) are also examined in the literature (e.g., Xie & Wolfram 2002).

Services offered by digital libraries vary from institution to institution. They generally fall into two broad categories: traditional services and services unique in the digital environment. Personalization, plagiarism detection, analysis and processing of digital information are some example of unconventional digital library services (Fox & Urs 2002). Reference services of digital libraries, on the other hand, belong to the traditional category. However, such services have evolved so rapidly and extensively that they become an emerging domain called digital reference services (e.g., Chowdhury 2002; Sloan 2001; Stemper & Butler 2001). Current awareness seems to be a service that borders the boundary between conventional and unconventional digital library services. It is not a major item on the traditional library service menu. Yet, it would achieve its full potential in digital libraries thanks to the Internet technology.

NEP (New Economics Papers) is a current awareness service of the RePEc digital library. This paper is intended to describe the development, performance and potentials of the NEP service.

The NEP Service

NEP (http://netec.wustl.edu/NEP/), started in 1998, provides current awareness service to all those who have subscribed to the individual NEP lists (see Appendix) for distributing papers newly entered into the RePEc digital library. The history and operations of the NEP service are comprehensively reported in Cruz, Krichel and Trinidad (2003).

Unlike RePEc which holds both working paper (i.e., recent research reports prior to formal publication) and article (i.e., peer reviewed writings) data, NEP covers only working paper data. Papers recently arrived in the distributed RePEc databases are selected and compiled by a general editor of nep-all (All New Papers), the overall list that functions as the source provider for individual NEP lists. Each year, nep-all on the average distributes about 40 issues of new economics papers to the editors of individual NEP lists (see Figure 2). Those individual NEP list editors then select papers from nep-all, compile the selected papers into an issue of a NEP report, and disseminate via their own lists to respective list subscribers. Figure 1 illustrates the NEP service process.

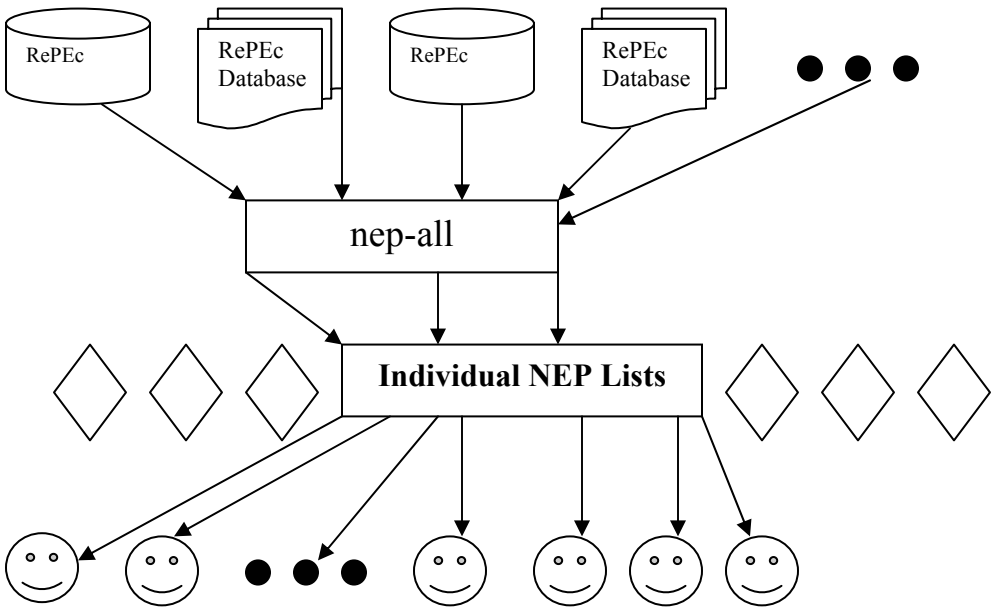

Figure 1. The NEP Service Process

It should be pointed out that there is not necessarily a one-to-one relationship between each issue of nep-all and other individual NEP list reports. One individual NEP list report may announce papers from different issues of nep-all. In addition, one paper from a single nep-all issue can be announced by more than one NEP list or not selected for announcement by any NEP list. For this reason, the term "announcement" is used to refer to the inclusion of an individual paper from net-all in a specific NEP list report.

Current awareness, as indicated before, is not new as a service in disseminating the latest information to the user in brick-and-mortar libraries. Nor is it unheard of in the world of digital libraries. The Contents-to-Go of the TORPEDO Ultra Digital Library Initiative is a case in point (Stackpole & King 1999). That current awareness service emails electronic table of contents for all journals in its collection to its subscribed users. Other kinds of current awareness include the use of lists to share information among catalogers (Condron & Tittemore 2001). PhysDoc (http://physnet.uni-oldenburg.de/PhysNet/physdoc.html) collects and organizes links to physics resources worldwide by continent, country and town (Severiens, et al 2000). It has the component for current awareness but is at best built upon the pull technology. Users have to visit PhysDoc in order to learn what is new in the field.

In comparison, NEP is based on the push technology. New economics papers are sent to subscribers via individual NEP lists. The basic model of NEP was set up by Thomas Krichel in 1998. NEP has been running, although not free of problems (Cruz,

Krichel & Trinidad 2003), without change until now. It perhaps could become an innovative model for current awareness services of digital libraries. Among the many features that distinguish NEP from other current awareness services, NEP operates on a scale larger than Contents-to-Go (Stackpole & King 1999), as will be described below.

NEP's Growth & Development

The first issue of nep-all was distributed on May 4, 1998. Since then, a total number of 215 nep-all reports have been issued by the end of June in 2003. Figure 2 describes the total number of issues nep-all distributes (i.e., # of issues) and the average number of papers each nep-all issue carries (i.e., issue size) every year from 1998 to 2003.

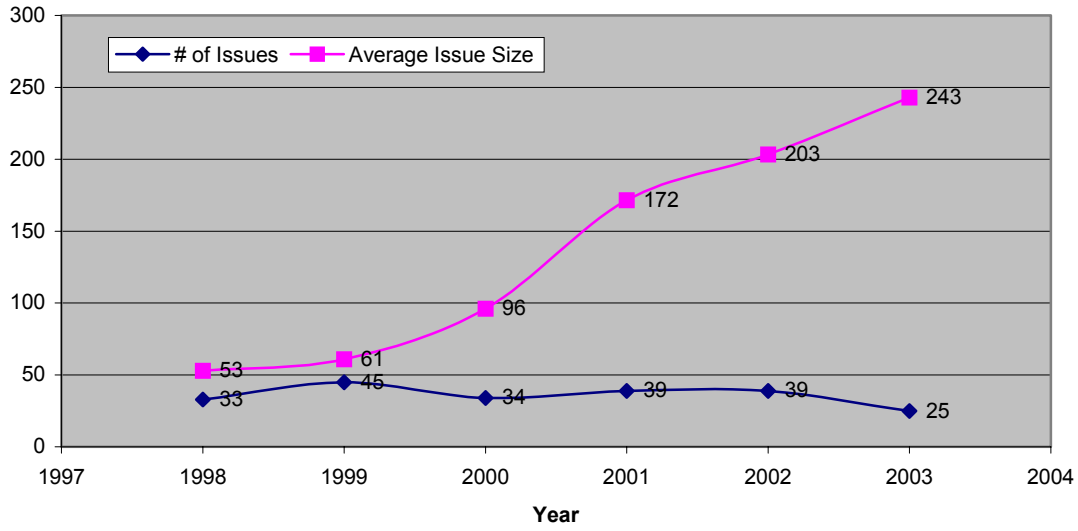

Figure 2. NEP-all Issues and Average Size

As shown in Figure 2, nep-all issue size steadily increases over the years while the number of issues nep-all produces annually remains about the same. This implies that the NEP service grows considerably in terms of the number of papers it distributes to individual NEP lists. Each nep-all issue in 2003 contains on average 243 papers, more than quadruple of the issue size in 1998. The most noticeable change in issue size occurred between 2000 and 2001 when the average number of papers each nep-all issue carries jumped from 96 in 2000 to 172 in 2001. In the time period covered in this article, nep-all in total delivered 215 issues that all together consist of 28,433 papers.

Once nep-all sends out an issue, it is the individual NEP lists that ultimately select and deliver the papers contained within to their respective list subscribers. There are 56 such individual NEP lists altogether at the time of this writing, excluding nep-all (see Appendix). Figure 3 gives the number of new NEP lists created and the accumulated number of such lists available on a yearly basis.

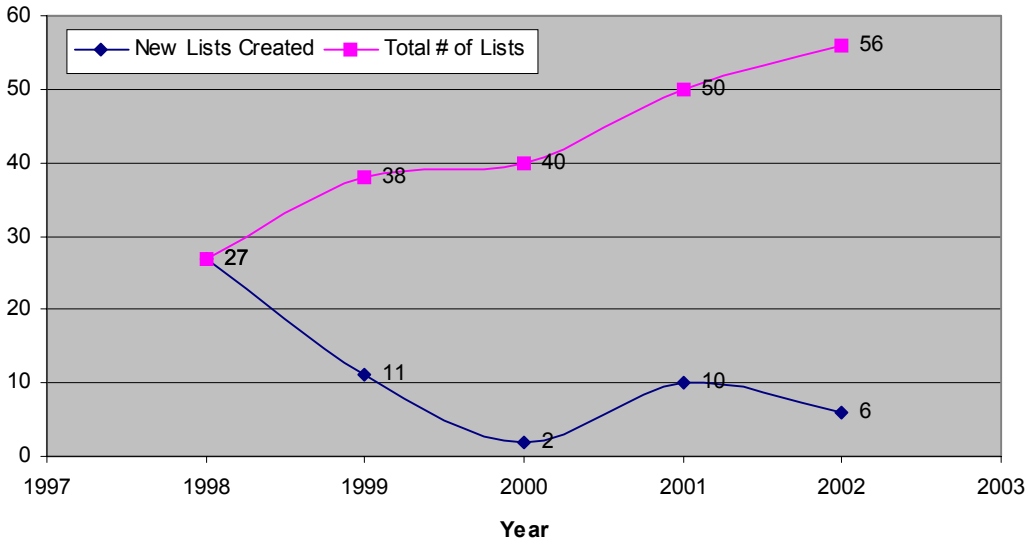

Figure 3. Growth of Individual NEP Lists

The year of 1998 saw the largest number of new NEP lists created, a necessity for launching the NEP service. While 11 and 10 more new lists were formed in 1999 and 2001 respectively, only two were set up in 2000. One possible explanation for this phenomenon is that the NEP service had reached an early level of maturity in 2000. In late 2002, NEP moved to a new list software system. No new list was opened while technical changes took place. The new list information for 2003 is not included in this project. Via the network of 56 individual NEP lists, a total number of 37,588 announcements of nep-all papers are made to 21,911 list subscribers. Further analysis and discussion of the individual NEP lists will be reported in a separate paper.

As explained previously, an announcement in NEP refers to the inclusion of an individual paper from net-all in a NEP list report. Figure 4 displays the total number of announcements all NEP lists made of nep-all papers every year. In 1998, the beginning year of the NEP service, it is understandable that fewer nep-all papers were announced due to several factors (e.g., issue size and number of lists). A leap in the number of announcements was subsequently observed from 2000 to 2001, parallel to the growth of nep-all issue size (see Figure 2). It should be noted that the announcement number in 2003 includes the data for the first half of the year. We will not be surprised if the final figure of announcements for 2003 turns out to be the largest of all since 1998.

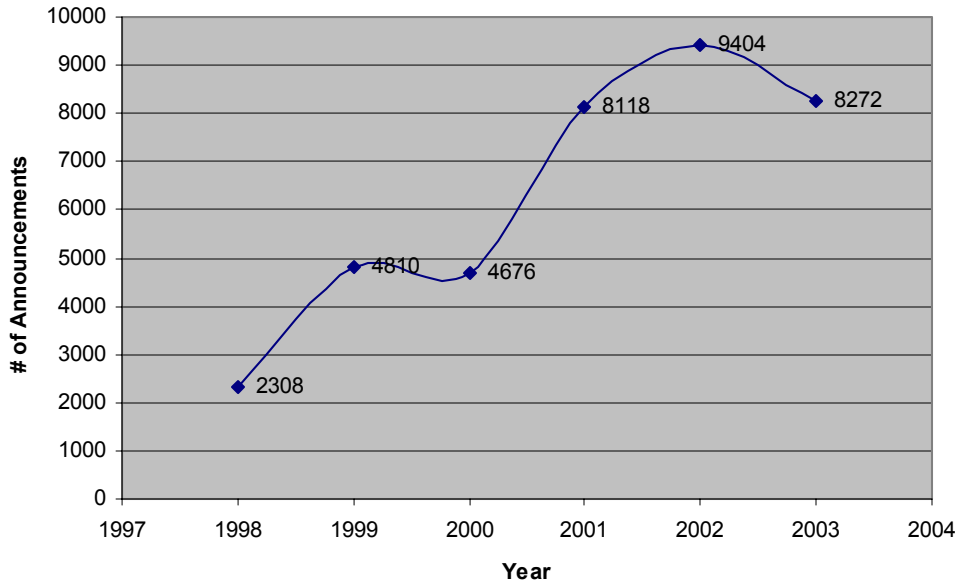

Figure 4. Number of Paper Announcements by Year

In a little bit over five years, NEP has grown and developed into a distinctive and fully functioning current awareness service for the RePEc digital library. We presented on a yearly basis the growth and development of the NEP service in terms of the nep-all publication frequency, average nep-all issue size, evolution of the NEP lists, and total number of announcements the NEP lists made. The NEP service overall has been enjoying healthy growth since its creation in 1998. In addition, substantial progress was made between 2000 and 2001.

NEP's Performance

Timeliness, coverage ratio, and usage are the three parameters we have chosen to measure the performance of the NEP service in this study. As NEP is a current awareness service for digital libraries, we have selected our evaluation criteria accordingly. The evaluation frameworks suggested by, for example, Choudhury et al (2002) and Saracevic (2000), although pertinent, are not exactly for evaluating the current awareness service of digital libraries.

Timeliness

For any current awareness service, timeliness should be one of the first dimensions to be considered. The value of a current awareness service will diminish if the service cannot be provided in a timely fashion. With respect to NEP, we define timeliness as the difference between the time when an issue of nep-all is distributed and the average time when papers from that issue of nep-all are announced in individual NEP list reports. Timeliness in this writing is measured in hours. Figure 5 shows the average timeliness data from 1998 to 2003.

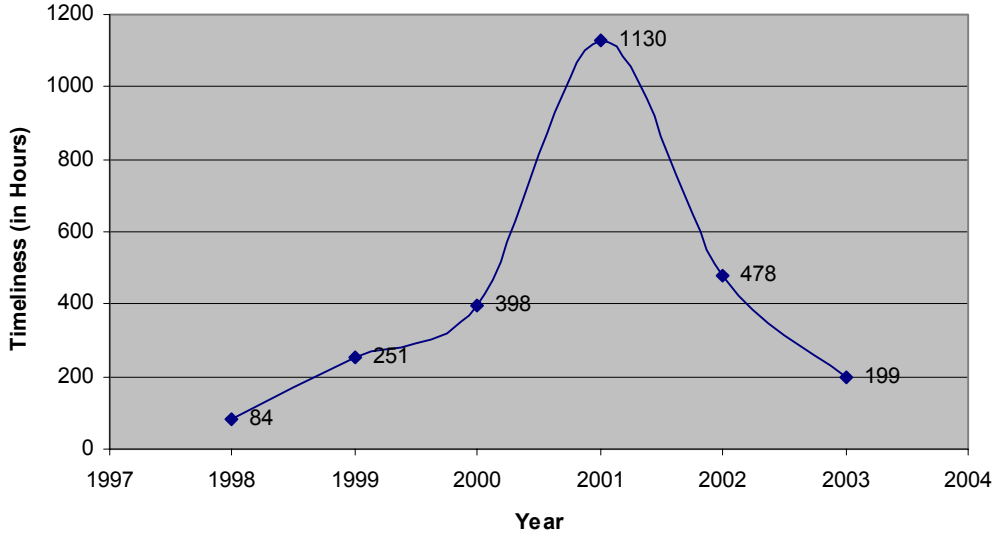

Figure 5. Timeliness of the NEP Service

Figure 5 indicates that 1998 is the year when NEP performed the best judging by timeliness. Several factors contribute to the result. First, NEP was more manageable then as it was young and small in size. Second, the editors of individual NEP lists were more enthusiastic about the NEP service when it is new. Third, the task for individual NEP list editors becomes more demanding as nep-all grows in frequency and size. It requires more time and efforts on the editor's side to wade through the piles of papers issued by nep-all (Cruz, Krichel & Trinidad 2003). As described by Cruz, Krichel and Trinidad (2003), a series of technical changes and problems occurred from 2000 to 2002 have also affected the timeliness of the NEP service.

According to Figure 5, the year of 2001 saw the biggest delay in announcing nep-all papers. The significant increase of nep-all issue size (see Figure 2) partially explains this problem in concern. Like in many other situations, human beings need some time to make adjustment to new changes. The sudden burst of nep-all papers in a sense brought challenges to the NEP list editors. A close examination of nep-all issue size data in 2001 reveals that only eight out of the 39 issues have fewer than 100 papers each. The rest of the nep-all issues may carry as many as 478 papers in a single report.

The average hours needed to announce a paper from nep-all is 443 across the years, which is equivalent to less than 19 days. This figure does not seem very impressive by itself. However, the NEP management has since been taking measures to ensure a more timely service. As a result, the mean timeliness for the year of 2003 reduces to 199 hours from 478 hours in 2002.

Coverage Ratio

Three possibilities exist for papers contained in a nep-all issue regarding the number of announcements they received: none, once, and more than once. Based on this categorization, coverage ratio is defined as the number of papers that received at least one announcement in a NEP list report, divided by the total number of papers in each issue of nep-all. In fact, coverage ratio deals with nep-all papers that are announced once and more, the second and third possibilities we just enumerated. As for papers in nep-all issues that received no announcement, we plan to explore them in a future paper to find out, for instance, the suitability of the general structure that all the 56 NEP lists represent in economics.

Figure 6 exhibits the average coverage ratio of the papers of each nep-all issue by year.

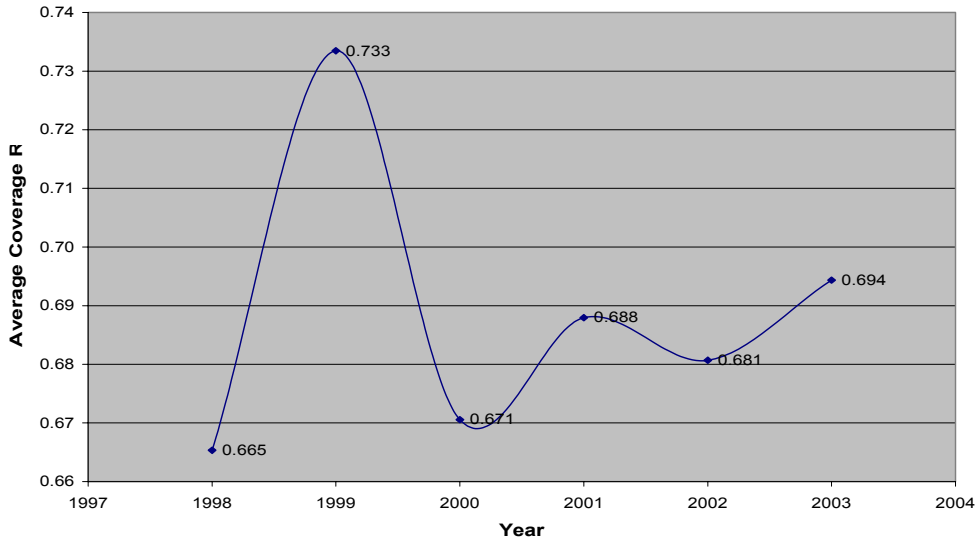

Figure 6. Average Coverage Ratio by Year

The fluctuations in coverage ratio appear small over the years with 0.068 (or less than 7 percent if expressed in percentage) being the largest difference found from one year to another in the data series. The mean coverage ratio throughout the time is 0.7 or 70 percent. In other words, 70 percent of the papers distributed in nep-all issues subsequently get announced via one or more NEP lists. There are on average 30 percent of the nep-all papers that did not receive any announcements.

It was anticipated based on common sense that coverage ratio would change as the nep-all issue size gradually increases over the years (see Figure 2). The reality appears to be rather static ($r = -0.16$, $p = 0.05$), which presumably hides two contrasting trends. First, as more individual NEP lists are created, all else being equal, the coverage ratio should improve. On the other hand, it seems tempting to theorize that as the size of nep-all increases, again all else being equal, there will be a decline in the coverage ratio. Cruz, Krichel and Trinidad (2003) have empirical evidence to corroborate this idea, but have no theoretical explanation. We think that their empirical analysis makes sense. The theoretical underpinning of this observation could be that individual editors have a target report size. That is, for each issue of their report, they will try to have a target size of

papers. If there are few papers in nep-all, their selection will be more lenient. If the nep-all issue is large, they will be more selective.

While timeliness clearly associates with the quality of awareness service, coverage ratio affects both the service quality as well as service substantiality. A current awareness service will not be highly regarded if only a small fraction of new information in a given area is reported. In this sense, NEP performs quite well although no analogous data is available for comparison purposes.

Multiple announcements of a single nep-all paper, mentioned in the beginning of the discussion on coverage ratio, result from the unique distribution network NEP has developed for disseminating new working papers in economics. The distribution network is unique in that 56 individual lists are currently employed in the NEP service with little coordination among them. Each NEP list editor is not aware of how other NEP editor colleagues work and is not required to consider if a particular paper is also announced in other list(s). In addition, individual NEP lists are created without an overseeing mechanism to ensure their mutual exclusiveness. They are only intended to loosely represent the subfields within economics. It is therefore common, rather than unusual, to have a single nep-all paper announced in more than one NEP list. Overlaps in announcement of nep-all papers can be treated as a synonym for multiple announcements.

The average multiple-announcement ratio over the years is 0.4 or 40 percent of all the nep-all papers. Specially speaking, multiple-announcement ratio refers to the number of a single paper that was announced by more than one NEP list, divided by the total number of papers in each issue of nep-all. Of all the papers included in nep-all issues, 40 percent of them are announced by more than one NEP list. We will elaborate on the ramifications of this finding later when we associate multiple announcements with other NEP parameters examined in this study.

Usage

Usage of the current awareness information NEP provides, given the availability of data source, is assessed in the number of papers downloaded from the RePEc collection as a result of NEP list announcements. Even so, data of such downloads is only available from the nep-all issue dated October 27, 2002 and on, or the last 34 out of the 215 nep-all issues considered in this article. We understand the limited nature of the usage data (e.g., downloading does not necessarily lead to actual usage such as reading). The data, nevertheless, should be able to help reflect NEP's performance in usage to a certain extent. Figure 7 lists the average number of NEP-announced paper downloads from each nep-all issue in a month because some months (e.g., March 2003) had five nep-all issues whereas others had three (i.e., June 2003) or four issues (e.g., November 2002). Since there is only one nep-issue with download data in October 2002, it is not included in Figure 7.

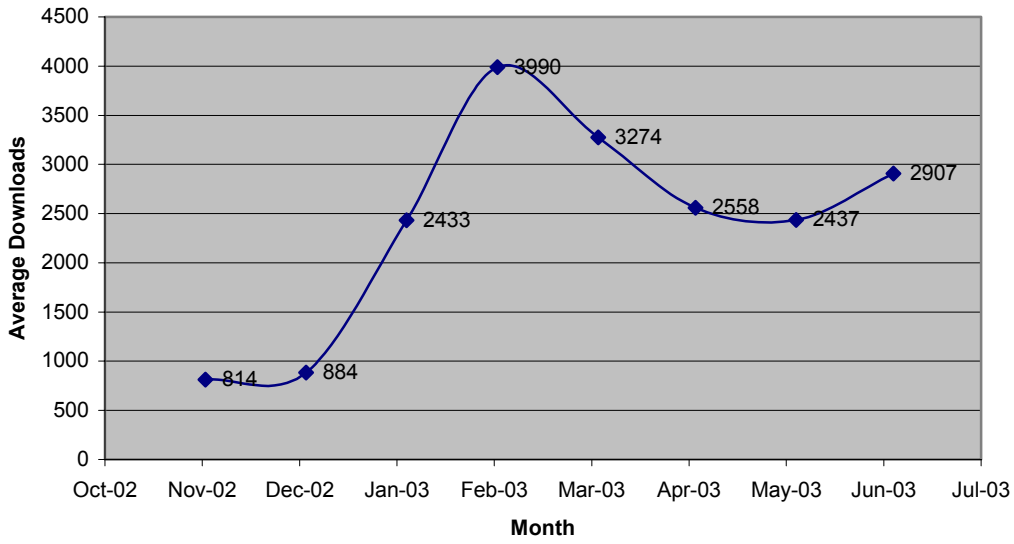

Figure 7. Average Downloads Per NEP-all Issue

As shown in Figure 7, average downloads of NEP-announced papers per nep-all issue peaked in February 2003. A possible explanation for this outcome is that people are usually ready to read more research literature after the holiday season is over. The same perhaps can be used to explain the low downloads from November to December in 2002 during which major holidays fall. With the exception of March 2003, the remaining months in the data set recorded similar download volume. In addition, usage data gathered for this study is cumulative over time, so papers announced earlier in the time sequence automatically get more downloads. Note that we also cannot distinguish downloads as a result of NEP announcements from those done directly from RePEc.

Out of the 10,164 papers announced in NEP list reports during the time period where download data is available, 10,799 downloads were made. This number indicates that every NEP-announced paper on the average is downloaded at least once. The objective of the NEP service, namely, to provide subscribers with up-to-date information to the RePEc digital library, is therefore met because NEP-announced papers are downloaded for possible use.

The user dimension undoubtedly influences the usage of the NEP service. We plan to research in that area and other human related aspects of NEP when we sort out the raw data of such kind.

Potentials and Other Perspectives of the NEP Service

What has been presented so far demonstrates that the NEP service, despite of the difficulties and problems it encountered (Cruz, Krichel & Trinidad 2003), has been thriving since its creation thanks to the joint and dedicated efforts of many volunteers.

Looking at the interrelationships among the parameters we have examined, we find that downloading of NEP-announced papers correlates significantly ($p = 0.01$) with nep-all issue size ($r = 0.58$), number of announcements ($r = 0.75$), and multiple announcements ($r = 0.77$). These findings suggest that nep-all should increase its size to facilitate more downloads. Meanwhile, either single or multiple announcements of nep-all papers should be further encouraged. They have an even greater impact than nep-all issue size in enhancing the value of the NEP service because NEP list subscribers may not download any papers if they are not announced. The immediate implication of these results seems at least two-fold: One is that existing NEP lists ought to try announcing more papers from nep-all. The other suggestion would be to create more lists to cover the 30 percent of papers that are reported in nep-all but receive no announcements from any existing NEP lists.

Timeliness, as a measurement of NEP performance, has little impact on the downloading of NEP-announced papers. Peer review of one paper in economics on average takes four years. So the NEP delays appear really tiny in comparison. Besides, a paper is downloaded in a current awareness service perhaps mainly for its relevance and less so for its currency. Nevertheless, timeliness should be upheld as a standard for having quality current awareness services as we discussed previously.

The issue size of nep-all shows a weak association with timeliness ($r = 0.28$, $p = 0.01$) and a strong impact on both number of announcements ($r = 0.88$, $p = 0.01$) and multiple announcements ($r = 0.82$, $p = 0.01$). In other words, a large nep-all issue size would not automatically delay the announcement of papers in that issue if NEP list editors do the job decently. But an increase in nep-all issue size would help augment the announcement of papers and consequently downloads.

With the recent implementation of the mailman list software (Cruz, Krichel & Trinidad 2003) plus other positive factors (e.g., devoted NEP volunteers), NEP shows great potentials in its future operation and development as a current awareness service.

References

Choudhury, Sayeed, et al. (2002). A framework for evaluating digital library services. *D-Lib Magazine, 8(7-8)*, http://www.dlib.org/dlib/july02/choudhury/07choudhury.html.

Chowdhury, Gorbinda G. (2002). Digital libraries and reference services: Present and future. *Journal of Documentation,* 58(3), 258-283.

Cole, Timothy W., Allen, Robert S., and Schmitz, John G. (2000). Building an outreach digital library collection. *Illinois Libraries, 82(4),* 239-250.

Condron, Lyn, and Tittemore, Cecilia Piccolo. (2001). Listservs and listprocs for catalogers. *Cataloging and Classification Quarterly, 32(3),* 25-29.

Cruz, José Manuel Barrueco, Krichel, Thomas, and Trinidad, Jerimiah C. (2003). Organizing current awareness in a large digital library. Presented at 2003 Conference on Users in the Electronic Information Environments in Espoo, Finland, September 8-9, 2003. http://openlib.org/home/krichel/papers/espoo.pdf.

Cruz, José Manuel Barrueco, and Krichel, Thomas. (2000). Cataloging economics preprints: An introduction to the RePEc project. *Journal of Internet Cataloging, 3(3),* 227 -241. http://openlib.org/home/krichel/papers/shankari.pdf.

Davis, James, and Lagoze, Carl. (2000). NCSTRL: Design and development of a globally distributed digital library. *Journal of the American Society for Information Science, 51(3),* 273-280.

Fox, Edward, and Urs, Shalini. (2002). Digital libraries. *Annual Review of Information Science and Technology, 36,* 503-589.

Krichel, Thomas. (2000). RePEc: An open library for economics: To appear as a book chapter. http://openlib.org/home/krichel/papers/salisbury.html.

Saracevic, Tefko. (2000). Digital library evaluation: Toward an evolution of concepts. *Library Trends, 49(2),* 350-369.

Schwarts, Candy. (2000). Digital libraries: An overview. *Journal of Academic Librarianship, 26(6),* 385-393.

Severiens Thomas, et al. (2000). PhysDoc - A distributed network of physics institutions documents: Collecting, indexing, and searching high quality documents by using Harvest. *D-Lib Magazine, 6(12),* http://www.dlib.org/dlib/december00/severiens/12severiens.html.

Sloan, Bernie. (2001). Reference service in the digital library: A report on the ready for reference project. *Library Hi Tech News, 18(10),* 14-19.

Stackpole, Laurie E., and King, Richard James. (1999). Electronic journals as a component of the digital library. *Issues in Science and Technology Librarianship, (22),* http://www.library.ucsb.edu/istl/99-spring/article1.html.

Stemper, James A., and Butler, John T. (2001). Developing a model to provide digital reference services. *Reference Services Review, 29(3),* 172-188.

Suleman, Hussein, and Fox, Edward. (2001). The Open Archives Initiative: Realizing simple and effective digital library interoperability. *Journal of Library Administration, 35(1-2),* 125-145.

Xie, Hong, and Wolfram, Dietmar. (2002). State digital library usability: Contributing organizational factors. *Journal of the American Society for Information Science & Technology, 53(13),* 1085-1097.

Appendix: Individual Lists of the NEP Service

Abbreviation	Full Name	Date of Creation
nep-all	*All New Papers*	*5/4/1998*
nep-acc	Accounting	8/11/2001
nep-afr	Africa	10/22/2001
nep-agr	Agricultural Economics	4/27/1999
nep-cba	Central Banking	10/23/2000
nep-cbe	Cognitive and Behavioural Economics	8/16/2002
nep-cdm	Collective Decision-Making	5/25/1998
nep-cfn	Corporate Finance	10/22/1998
nep-cmp	Computational Economics	10/9/1998
nep-com	Entrepreneurship	8/16/2001
nep-cul	Cultural Economics	10/18/2002
nep-cwa	Central and Western Asia	12/6/2001
nep-dcm	Discrete Choice Models	7/28/1998
nep-dev	Development	4/28/1999
nep-dge	Dynamic General Equilibrium	6/24/1998
nep-ecm	Econometrics	4/27/1998
nep-edu	Education	4/27/1999
nep-eec	European Economics	7/20/1998
nep-eff	Efficiency and Productivity	6/1/1998
nep-ene	Energy Economics	4/27/1999
nep-ent	Urban and Real Estate Economics	10/24/2002
nep-env	Environmental Economics	8/10/1998
nep-ets	Econometric Time Series	4/27/1998
nep-evo	Evolutionary Economics	5/21/1998
nep-exp	Experimental Economics	4/27/1998
nep-fin	Finance	4/22/1999
nep-fmk	Financial Markets	6/10/1998
nep-geo	Economic Geography	3/20/2002
nep-gth	Game Theory	5/18/1998
nep-hea	Health Economics	4/27/1998
nep-his	Economic History	4/28/1999
nep-hpe	History and Philosophy of Economics	9/1/1999
nep-ias	Insurance Economics	11/5/1998
nep-ifn	International Finance	6/29/1998
nep-ind	Industrial Organization	4/26/1999
nep-ino	Innovation	9/28/1999
nep-lab	Labour Economics	4/22/1999
nep-lam	Sports and Economics	7/20/1998
nep-law	Law and Economics	4/28/1999
nep-ltv	Unemployment, Inequality and Poverty	9/4/1998
nep-mac	Macroeconomics	11/15/2001
nep-mfd	Microfinance and Financial Development	7/25/2001
nep-mic	Microeconomics	4/27/1998
nep-mon	Monetary Economics	10/19/1998
nep-net	Network Economics	9/7/1998
nep-pbe	Public Economics	4/28/1998
nep-pke	Post Keynesian Economics	6/21/1998

nep-pol	Positive Political Economy	4/28/1998
nep-pub	Public Finance	5/20/1998
nep-reg	Regulation	5/13/2000
nep-res	Resource Economics	11/6/2001
nep-rmg	Risk Management	11/26/2002
nep-sea	South East Asia	8/22/2001
nep-spo	Central and South America	8/16/2001
nep-tid	Technology and Industry Dynamics	5/21/1998
nep-tra	Transition Economics	11/28/2001
nep-ure	Industrial Competition	10/23/2002