On the usage of e-journals in French universities

The article investigates the use of electronic journals in 68 French universities, all members of the Couperin consortium[1]. The study is based on statistical data on usage, libraries and journal packages collected by the EPEF project (Evaluation des Périodiques Electroniques dans le Réseau Universitaire Français) from different sources (publishers, Ministry, survey results), from 2001 to 2005. It provides information about local metrics for eight packages (serials), e.g. usage trends; requests per package and user; costs per request and user; as related to the scientific specialty of the university (STM, SS&H, multidisciplinary, etc.). Whenever possible and reasonable, the data is compared with results from UK and US research. The paradoxical situation of SS&H is discussed. This is the first study on usage statistics in French universities and probably one of the first based on a nationwide survey. Methodological problems are discussed.

CHERIFA BOUKACEM-ZEGHMOURI
Senior Lecturer
GERIICO, Charles de Gaulle University, France

JOACHIM SCHÖPFEL
Head of Department
INIST-CNRS and Document Numérique & Usages, France

Introduction

The availability of electronic journals in the French academic community improved in a spectacular way since the creation of the national university consortium Couperin[1] in 1999, expanding from 53,403 (2000) to 477,391 (2005) online subscriptions (factor 8), while the overall expenses for scientific and technical information (STI) in university libraries increased only at a rate of 17%, from €70m (2001) to €82m (2005). But except for some local data, so far little was known on the overall uptake and real usage of these resources; more generally, very few statistical surveys on usage data of online digital resources have been published in France up to now.[2]

Funded by the French National Agency of Research (ANR) from 2006 to 2009, the Evaluation des Périodiques Electroniques dans le Réseau Universitaire Français (EPEF) project[3] collects nationwide usage data from university libraries and analyses general and domain-specific trends. Additionally, it introduces international standards on usage statistics and metrics in France[4] and conducts qualitative surveys on usage behaviours.

In particular, EPEF is based on the COUNTER project (http://www.projectcounter.org) and derived studies[5,6] and the deep log analysis of the London CIBER team[7,8,9,10]. The specifics of the EPEF approach are a successful collaboration between academics, scientists, LIS professionals (librarians) and end-users, interdisciplinary work at the crossroads of social, economical and information sciences, and the emphasis laid on economical aspects ('value for money').

The following article reports on preliminary EPEF results, in particular on local metrics, and addresses some methodological shortfalls and issues.

Methodology

Descriptive data was collected from two sources:
Usage statistics: data was collected from eight academic publishers (ACS, Blackwell, CNRC

Press, Elsevier, IOPP, Kluwer, Springer and Wiley) for 68 universities for the period 2001-2005. Only 25-30% of the data was COUNTER compliant (JR1, 3 or 4).

Library characteristics: data on university libraries, e.g. budget, number of users[11], disciplines[12], portals etc., was collected from the Ministry of Education ('ESGBU statistics'[13]). This data was completed by results of a national survey on digital resources in the academic establishments ('ERE statistics'[14]). The period covered was 2003– 2005, except for the budget which had data only for 2005.

Other data collected by EPEF, e.g. impact factor and citations[15], usage statistics for bibliographic databases, web logs from local systems (Lyon, Grenoble, Paris, etc.), survey data from a sample of scholars and academics, was not used for this first report on results.

The data was integrated into a single database through Access then migrated to SAS; the statistical analyses were done with Excel and Access[16].

Results

The presentation of local metrics in this section generally follows the COUNTER definition if not otherwise indicated.

Trend in number of full text article requests
From 2001 to 2005, the requests for full text articles offered by the eight publishers mentioned above increased rapidly from around three million to more than 13 million (see Figure 1).

In the first years the annual increase rate was 50-80% but then slowed down to 10-30%, an annual rate comparable to that observed by Bevan *et al* for NESLi2[5]. We have no usage statistics for 1999 but based on the (low) number of e-subscriptions in

1999 we can estimate that the number had been noticeably below one million requests.

Of these requests, 60% came from sciences, technology and medicine (STM), 30% from greater and 6% from smaller or medium-sized multi-disciplinary universities. Faculties of law and business (LB), including economics and management, accounted for 3%; only 1% came from social sciences and humanities (SS&H), including arts, universities.

The usage trends varied between the categories of universities; from 2003 to 2005 the increase in usage was 45% for STM and 37-38% for multi-disciplinary universities. Nevertheless, the most important growth rates were observed in LB (+111%) and SS&H universities (+666%).

Full text article requests per title
The average number of full text article requests per title increased from 23.0 (2001) to 27.6 (2005), an increase of 20%. Yet, the usage data does not show a steady growth (see Figure 2).

After an important increase between 2001 and 2003 (>50%), the average number of requests per title slowed down again. This may be an effect of the big deals, with a growing number of low-use and formerly unsubscribed titles ('long tail').

Globally, this data seems to be consistent with other findings (cf. Bevan *et al*[5] who distinguished access to formerly subscribed and unsubscribed titles, a distinction EPEF did not make).

Full text article requests per publisher package
Over the observed period (2001-2005), the Elsevier Freedom collection received the most requests (78%), followed by ACS (10%) and Wiley (6%). The other publishers accounted only for 6%. The strong position of Elsevier was relatively stable, with annual variations between 75% and 80%.
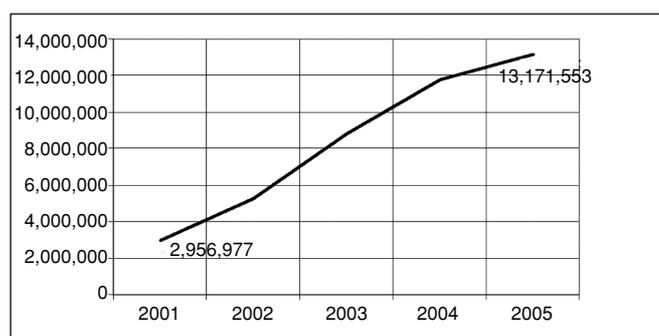

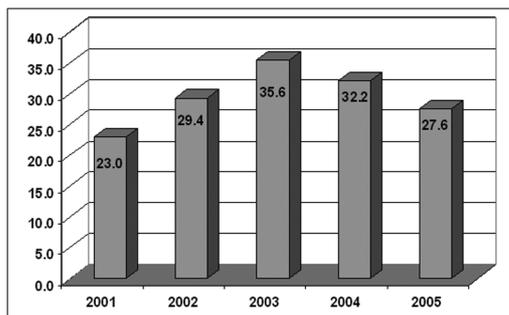*Figure 1. Full text article requests (2001–2005, 68 universities)*

Figure 2. Full text article requests per title 2001-2005

Full text article requests per user

The full text article requests per user increased between 2003 and 2005 from 19 to 41 (+116%). We can observe a significant difference between universities (see Figure 3).
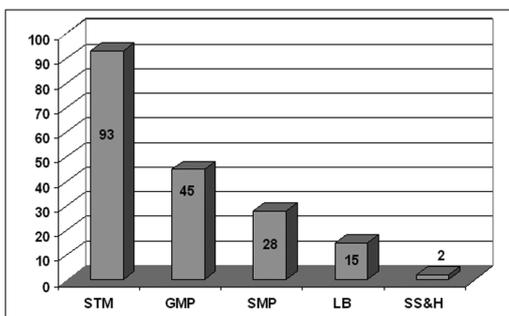

Figure 3. Full text article requests per user and university type 2005

> Abbreviations of university types: STM – sciences, technology and medicine; GMP – large multidisciplinary; SMP – small or medium-sized multidisciplinary; LB – law and business, including economics and management; SS&H – social sciences and humanities, including arts

Academics from STM faculties use electronic resources negotiated by the academic consortium six times more than their colleagues from law and business and nearly 50 times more than those from SS&H disciplines.

Most requested titles

Based on a representative sample of big STM universities in 2005 (17% of overall downloads), the 10 most requested titles ('high range titles', cf. Conyers[6]) received 12% of the overall traffic. The same analysis for SS&H universities indicates a more 'centralized' usage, with the ten most requested titles obtaining nearly 19%. Like Conyers, we observed that STM titles predominated in the 'high usages ranges'.

Taking all categories together, the three international titles 'in the charts' in 2005 were *Tetrahedron Letters*, *Tetrahedron* and *Cell.* The top three French journals were medical titles: *Revue de Médecine Interne, Annales Françaises d'Anesthésie et de Réanimation* and *Archives de Pédiatrie.*

Cost per full text article request

The average cost per full text article request in 2005 is €1.60, that is about 20% above the observed average cost in the UK or USA (around US$2.00 or £1.50).[5,17,18] But again, the subject-related differences are significant (see Figure 4).
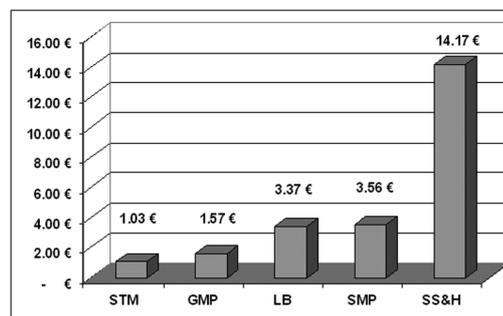

Figure 4. Cost per full text article request and university type 2005

For key to abbreviations, see Fig. 3

The difference between large and smaller or medium-sized institutions (€1.03/1.57 to €3.56) corresponds approximately to that described by Bevan *et al*[5] between 'high use, high cost subscription' and 'lower use but lower cost' institutions (£1.48 to £5.00). The only exception again is the SS&H category with average cost metrics roughly four times higher than comparable international figures.

Cost per user

According to the official statistics for 2005 (ESGBU), the university libraries spent 6.3m for e-journals (86% through big deals), with 53% from STM and 33% from large multidisciplinary (GMD) universities.

Related to the number of users, the differences in investment are significant (see Figure 5).

Expenditure on e-journals per user in STM universities was about two times higher than in multidisciplinary universities and up to ten times higher than in SS&H and LB universities. Of course, this must be seen in relation to the real usage (see above, Figures 3 and 4).
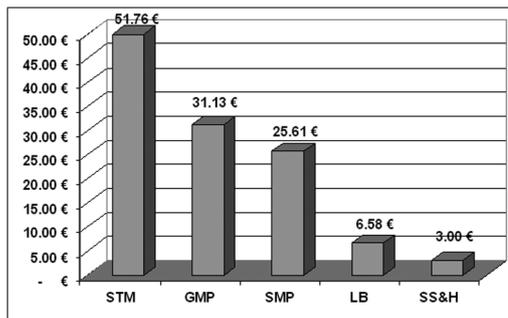
Figure 5. Investment in e-journals per user and university type 2005

For key to abbreviations, see Fig. 3

Discussion

The EPEF project faced some methodological problems that tend to diminish the reliability of the collected data. First, as mentioned above, only 25–30% of the usage data collected from publishers was COUNTER compliant, a lack of homogeneity that induces the risk of flawed metrics and interpretation.

Second, compared to the ESGBU data the ERE survey-based statistics were often revealed to be of lesser value, in particular concerning data on investment and staffing. This again induces risk of error (e.g. a latent overestimation of the two metrics 'full text article requests per user' and 'costs per user') even if large-scale studies on a consortium-level like EPEF necessarily have to deal with these methodological shortfalls.

Another potential bias the EPEF project has not addressed so far is the possible overestimation of usage by double downloads (HTML and PDF, double clicks etc.).

Even so, taking into consideration the potential impact of these aspects, the collected data of the first phase of the EPEF project appears sufficiently significant to provide some major results that can be resumed as follows:

■ After a period (before 1999) of relatively slow uptake compared to other countries with important public research funding, French universities globally caught up with the international level of digital resource usage and costs.
■ We can observe significant differences between the French universities with regards to licensed contents, investments and usage. Most investment is done in STM faculties where the usage is highest.

■ For the SS&H, the situation is paradoxical: even if the investment is quite low, we observe a high cost per usage and user related to the low usage statistics – too high compared to other disciplines and countries.

The paradoxical situation of SS&H can probably be explained by four factors:

1) With explicit support from the government, the national academic consortium in the first years clearly gave preference to licences for STM and other interdisciplinary content from international publishers (big deals, mostly English-language material) in order to close the gap with the most urgent needs for digital resources.
2) There are no big publishing companies for French academic journals, especially in SS&H where the digital offerings of French material have developed only recently, as in the digital collections from CAIRN, Revues.org and Persée.
3) Compared to other areas (medicine, pharmacology, physics, life sciences etc.), the SS&H communities are less accustomed to English-speaking documents and often prefer French resources that during the observed period were not part of the licensed content from Couperin.
4) So far usage statistics of SS&H portals (Revues.org, Persée, CAIRN) have not been available. Nevertheless, as the statistical data shows a significant growth in use of digital resources, there seems to be evidence for the potential of SS&H content.

Conclusions

A more detailed review of these first EPEF results is published in French[19]. Other research is ongoing and will be presented in 2009 and 2010:

■ a qualitative study on user behaviour profiles (semi-directive interviews) with a subset of 17 libraries, with regards to subject- and size-specific variation and use of technologies (tools) that may impact usage reporting, e.g. portals with federated search.
■ a comparative analysis between usage statistics and impact factor meant to contribute to the development of web-based global metrics, based on the 2006 and 2007 data and taking into account the forthcoming recommendations of the UKSG-COUNTER project on usage factors[20].

■ a study on usage of open repositories, in particular the emergent institutional archives, and other freely available material such as content on serials platforms in the social sciences and humanities.

Derived from the experiences of the first project phase and its methodological problems, EPEF suggests that in the future, usage statistics in France should be collected and evaluated on two levels: a federated, nationwide, consortium-wide gathering, storage and processing of publisher and vendor statistics in order to produce metrics on global usage trends; and a campus-based collection of local usage statistics through electronic resource management systems in order to produce, together with data on budget and users (staffing), decision aids for negotiations, acquisition policies and resource sharing.

Following the EPEF results, the global, nation-wide data is not sufficient for an accurate analysis of return on investment (scholarly outcomes related to access and information-seeking behaviour) and need to be completed by local information, even if the latter may be biased, too, for example through underestimation of off campus counts.

Monitoring the usage of digital resources gains is becoming increasingly important for acquisition policy and budget decisions. The COUNTER (and UKSG) recommendations provide the necessary framework both for publishers/vendors and libraries. The EPEF project supplies empirical evidence of usage trends and develops a robust 'methodological toolkit' for the evaluation of usage-related metrics.

One future challenge will be the improvement and implementation of this methodology in the French academic consortium, beyond the project timelines. Another challenge is the ongoing discussion with publishers and vendors for better statistical data, especially for big consortia. Related to the specific French research environment, a third challenge will be the connection and comparison of usage statistics between universities and scientific organizations (CNRS, INSERM, etc.) in order to provide a realistic understanding of 'who uses what and how, and who pays for it'.

The French academic landscape is changing now, offering more autonomy and importance, visibility and funding to the universities. Together with the transition of traditional economic models of scientific publishing to more open,

community-based models, this will provide a fascinating 'laboratory' (and a lot of problems) for studies on usage statistics and behaviours. The first phase of the EPEF project laid the methodological foundations for future improvement and implementation. So, the fourth and last challenge will be the adaptation of the empirical EPEF 'toolkit' to the emerging scientific and information environment in France.

References

1. Couperin: Consortium Universitaire de Publications Numériques, with more than 200 members (universities, HE and research institutions). At: http://www.Couperin.org (Accessed 23 May 2008)

   See also the article by C Forestier and P Carbone in this issue of *Serials.*

2. Assadi, H, Beauvisage, T, Lupovici, C, Cloarec, T, User and uses of online digital libraries in France, *Lecture Notes in Computer Science*, 2004, 2769, 1–12. http://thomas.beauvisage.free.fr/pubs/ecdl2003.pdf (Accessed 23 May 2008)

3. The project team for EPEF was made up of: 10 academics, scientists and professionals in library and information sciences and economics from Lille, Paris, Lyon and Nancy. The EPEF website is under preparation.

4. Boukacem, C, Schöpfel, J, Statistiques d'utilisation des ressources électroniques en ligne: le projet COUNTER, *Bulletin des Bibliothèques de France*, 2005, 50(4), 62–66. Available at: http://archivesic.ccsd.cnrs.fr/sic_00001543/fr/ (Accessed 23 May 2008)

5. Bevan, S, Dalton, P, Conyers, A, How usage statistics can inform national negotiations and strategies, *Serials* 2005, 18(2), 116-123. Available at: http://serials.uksg.org (Accessed 27 May 2008); doi: 10.1629/18116 at: http://dx.doi.org/10.1629/18116. (Accessed 27 May 2008)

6. Conyers, A, Building on sand? Using statistical measures to assess the impact of electronic services, *Performance Measurement and Metrics* 2006, 7(1), 37–44.

7. Nicholas, D, Huntington, P, Watkinson, A, Scholarly journal usage: the results of deep log analysis, *Journal of Documentation*, 2005, 61(2), 248–280.

8. Nicholas, D, Huntington, P, Evaluating the Use and Users of Digital Journal Libraries. In: *Digital Libraries*, Ed. Papy, F, 2008, London: ISTE Wiley.

9. Tenopir, C, Read, E, Manoff, M, Baker, G, Nicholas, D, King, DW, What does usage data tell us about our users? *Online Information 2007 Conference*. London, 4–6 December 2007. Proceedings. 2007, London, Incisive Media.

10. Nicholas, D, Huntington, P, Jamali, H R, User diversity: as demonstrated by deep log analysis, *The Electronic Library*, 2008, 26(1), 21–38.

11. The French official definition of 'users' globally corresponds to the common 'academic staff FTE' in the UK or US academic environment and includes academics, scientists, PhD (graduate) students, postdocs.

12. Boukacem, C, Inter-library loan services and access to electronic resources in French university libraries: a marriage of reason, *Interlending & Document Supply* 2003, 31(4), 218-227. Available at: http://archivesic.ccsd.cnrs.fr/sic_00000946/fr/
    The French university libraries are classified in five groups following their dominant scientific domains: STM (sciences, technology, medicine), GMD (grand multidisciplinary) and SMD (small or medium-sized multidisciplinary), SS&H (social sciences and humanities, including arts), LB (law and business, including economics and management).

13. See: http://www.sup.adc.education.fr/asibu/ (Accessed 18 June 2008)

14. See: http://www.sup.adc.education.fr/bib/Acti/Electro/accueil.htm#ere (Accessed 23 May 2008)

15. *Journal Citation Reports* (2005 edition) from the Institute for Scientific Information's *Web of Knowledge* (Thomson Scientific) at http://isiwebofknowledge.com/ (Accessed 27 May 2008)

16. Acknowledgment to Rachel Kamga (PhD student at Paris 8 University) for data preparation and analysis.

17. Franklin, B, Managing the Electronic Collection with Cost per Use Data, *IFLA Journal* 2005, 31(3), 241–248. Available at: http://ifl.sagepub.com (Accessed 28 May 2008)

18. Shepherd, P, How can we track what we do? COUNTER now and future? *P-D-R Special Meeting. Beyond the Model License: Pharma Libraries on the Move*. La Grande Motte, February 28–29, 2008.

19. Boukacem-Zeghmouri, C, Kamga, R, Les périodiques numériques dans le réseau universitaire français: premier état des lieux de la consultation, *Bulletin des Bibliothèques de France* 2008, 53(3) (forthcoming). Available at: http://bbf.enssib.fr (Accessed 27 May 2008)

20. See: http://www.uksg.org/usagefactors (Accessed 27 May 2008)

■ **Chérifa Boukacem-Zeghmouri**
Université Charles de Gaulle
Domaine Universitaire du 'Pont de Bois'
UFR IDIST
BP 149
59653 Villeneuve d'Ascq
France
E-mail: boukacemc@yahoo.fr

■ **Joachim Schöpfel**
INIST-CNRS
2 Allée du Parc de Brabois
54519 Vandoeuvre-lès-Nancy
France
Tel : ++33 (0) 383 504764
Fax : ++33 (0) 383 504732
E-mail: schopfel@inist.fr