

On the usage of e-journals in French universities

CHERIFA BOUKACEM-ZEGHMOURI

Senior Lecturer
GERIICO, Charles de Gaulle University
Lille, France

JOACHIM SCHÖPFEL

Head of Department
INIST-CNRS and "Document Numérique & Usages"
Nancy, France

Abstract

The article investigates the usage of electronic journals in 68 French universities members of the COUPERIN consortium. The study is based on statistical data on usage, libraries and journal packages collected by the EPEF project from different sources (publishers, Ministry, survey results), from 2001 to 2005. The article 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 are 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 nation-wide survey. Methodological problems are discussed.

Introduction

The availability of electronic journals in the French academic community improved in a spectacular way since the creation of the national university consortium COUPERIN¹ 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 €82 (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 (see Assadi et al. 2004).

Funded by the French National Agency of Research ANR from 2006 to 2009, the EPEF² project 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 (see Boukacem & Schöpfel 2005) and conducts qualitative surveys on usage behaviours.

In particular, EPEF is based on the COUNTER project³ and derived studies (see Bevan et al. 2005 or Conyers 2006) and the deep log analysis of the London CIBER team (see Nicholas et

¹ Consortium Universitaire de Publications Numériques with more than 200 members (universities, HE and research institutions), see at <http://www.couperin.org> and Carbone in this issue.

² *Evaluation des Périodiques Electroniques dans le Réseau Universitaire Français*. The project team: 10 academics, scientists and professionals in library and information sciences and economics from Lille, Paris, Lyon and Nancy. The EPEF web site is under preparation.

³ See <http://www.projectcounter.org>

al. 2005 and 2008, Tenopir et al. 2007, Nicholas & Huntington 2008). The specificities of the EPEF approach are a successful collaboration between academics, scientists, LIS professionals (librarians) and end users, interdisciplinary work at the crossroad 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 were collected from two sources:

Usage statistics: Data were collected from eight academic publishers⁴ for 68 universities, for the period 2001-2005. Only 25-30% of the data were COUNTER compliant (JR1, 3 or 4).

Library characteristics: Data on university libraries, e.g. budget, number of users⁵, disciplines⁶, portals etc., were collected from the Ministry of Education (“ESGBU statistics”⁷). These data were completed by results of a national survey on digital resources in the academic establishments (“ERE statistics”⁸). The covered period was 2003-2005, except for the budget with data only for 2005.

Other data collected by EPEF, e.g. impact factor and citations (JCR 2005), usage statistics for bibliographic databases, web logs from local systems (Lyon, Grenoble, Paris etc.), survey data from a sample of scholars and academics, were not used for this first report on results.

The data were integrated into a single database through Access then migrated to SAS; the statistical analyses were done with Excel and Access⁹.

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 of full-text articles offered by the eight publishers mentioned above increased rapidly from around 3m to more than 13m (see figure 1).

⁴ ACS, Blackwell, CNRC Press, Elsevier, IOPP, Kluwer, Springer, Wiley.

⁵ 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.

⁶ The French university libraries are classified in five groups following their dominant scientific domains (see Boukacem 2003): 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).

⁷ See <http://www.sup.adc.education.fr/asibu>

⁸ See <http://www.sup.adc.education.fr/bib/Acti/Electro/accueil.htm#ere>

⁹ Acknowledgment to Rachel Kamga (PhD student at Paris 8 University) for data preparation and analysis.

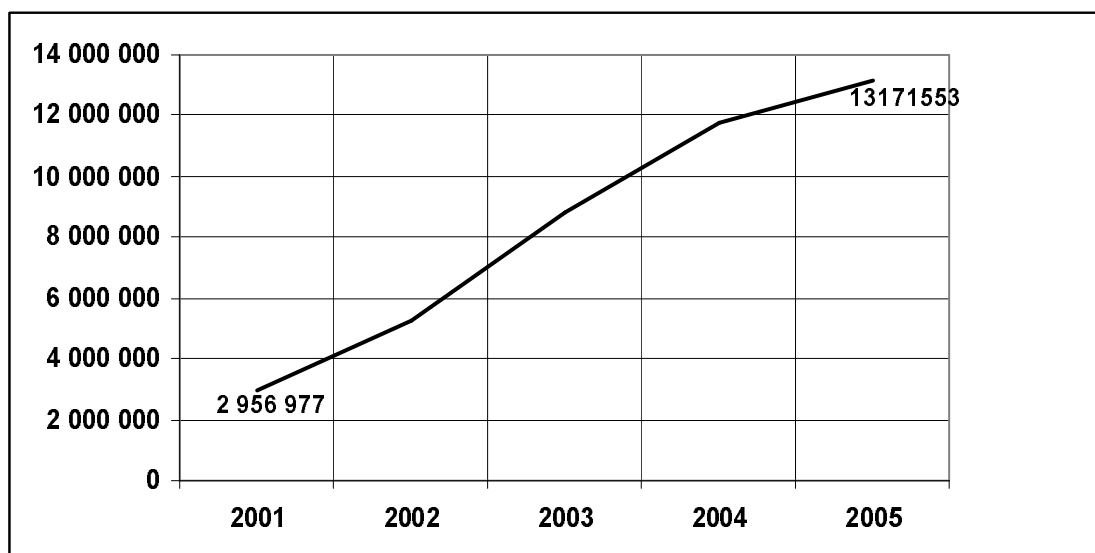


Figure 1: Full-text article requests (2001-2005, 68 universities)

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. (2005) for NESLi2. 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 1m requests.

60% of these requests came from STM, 30% from greater and 6% from smaller or medium-sized multidisciplinary universities. Faculties of law and business (LB) accounted for 3%; only 1% came from SS&H universities.

The usage trends varied between the categories of universities; from 2003 to 2005 the increase of usage was 45% for STM and 37-38% for multidisciplinary 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 don't show a steady growth (see figure 2).

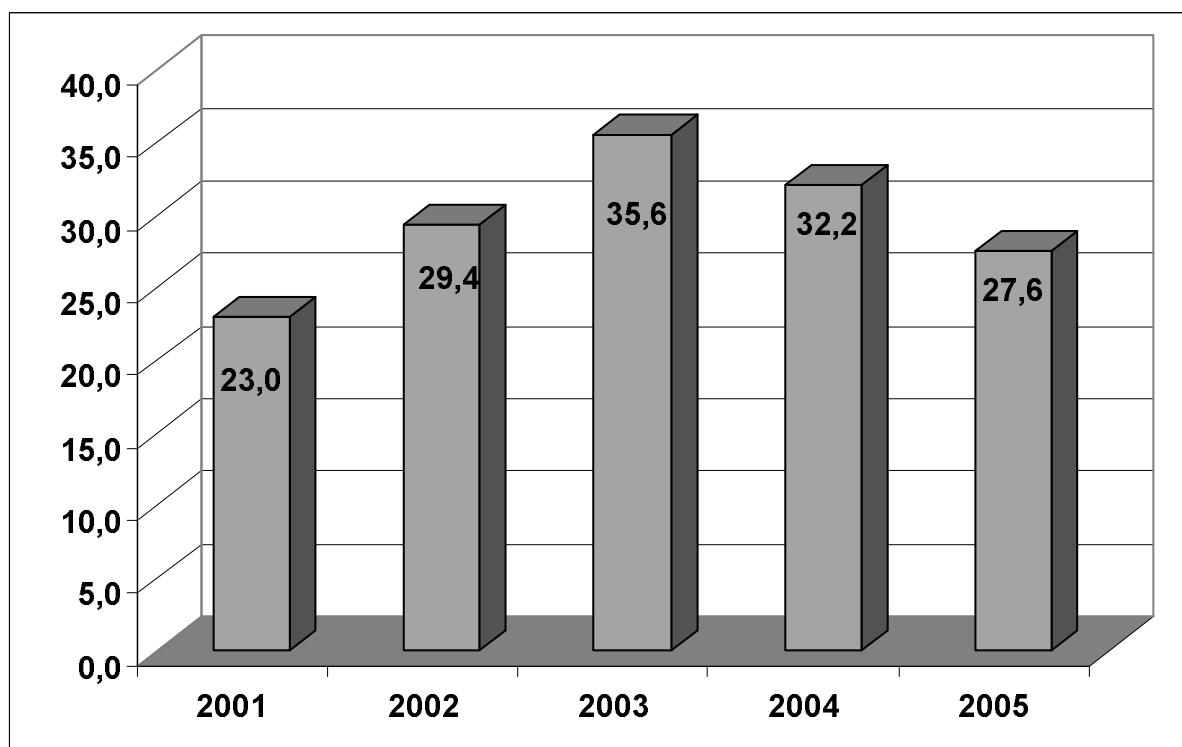


Figure 2: Full-text article requests per title 2001-2005

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, these data seem to be consistent with other findings (cf. Bevan et al. 2005 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%.

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).

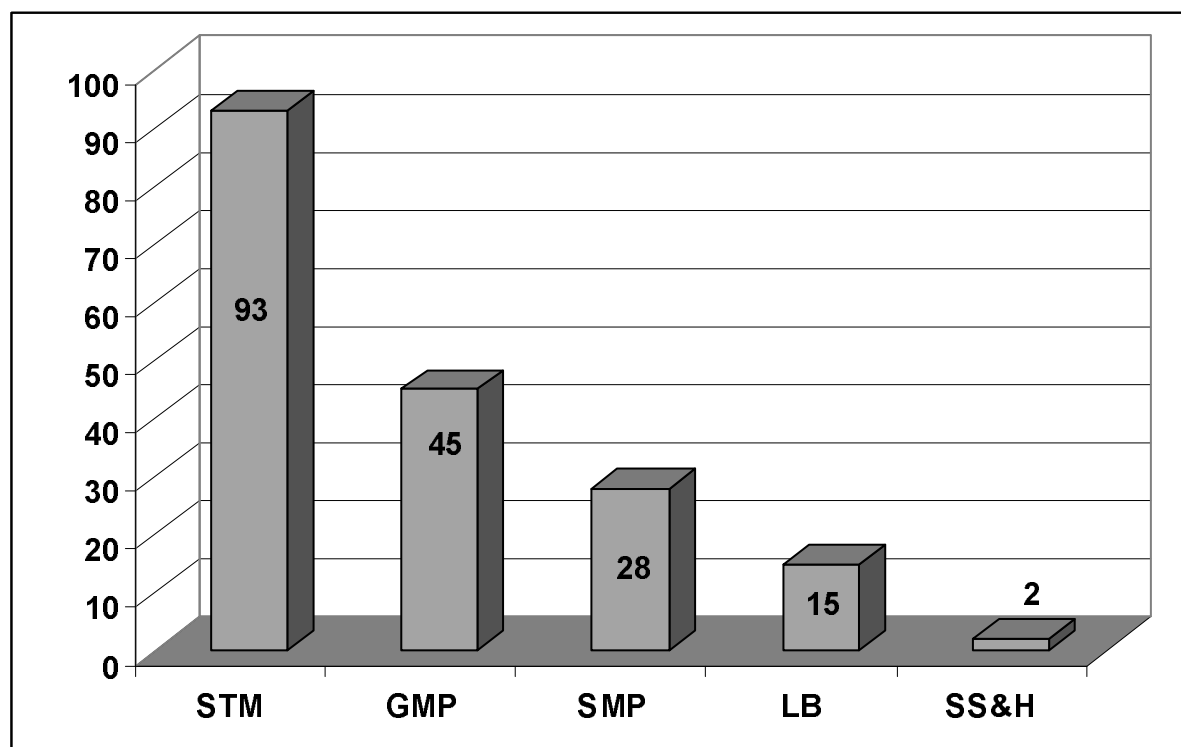


Figure 3: Full-text article requests per user and university type 2005 (acronyms cf. note 6)

Academics from STM faculties use electronic resources negotiated by the academic consortium 6 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 2006) received 12% of the overall traffic. The same analysis for SS&H universities indicates a more “centralised” usage, with the 10 most requested titles obtaining nearly 19%. Like Conyers (2006), we observed that STM titles predominated in the “high usages ranges”.

All categories together, the three international titles “in the charts” in 2005 were *Tetrahedron Letters*, *Tetrahedron* and *Cell*. The first 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) as reported by Bevan et al. (2005), Franklin (2005) or Shepherd (2008). But again, the subject-related differences are significant (see figure 4).

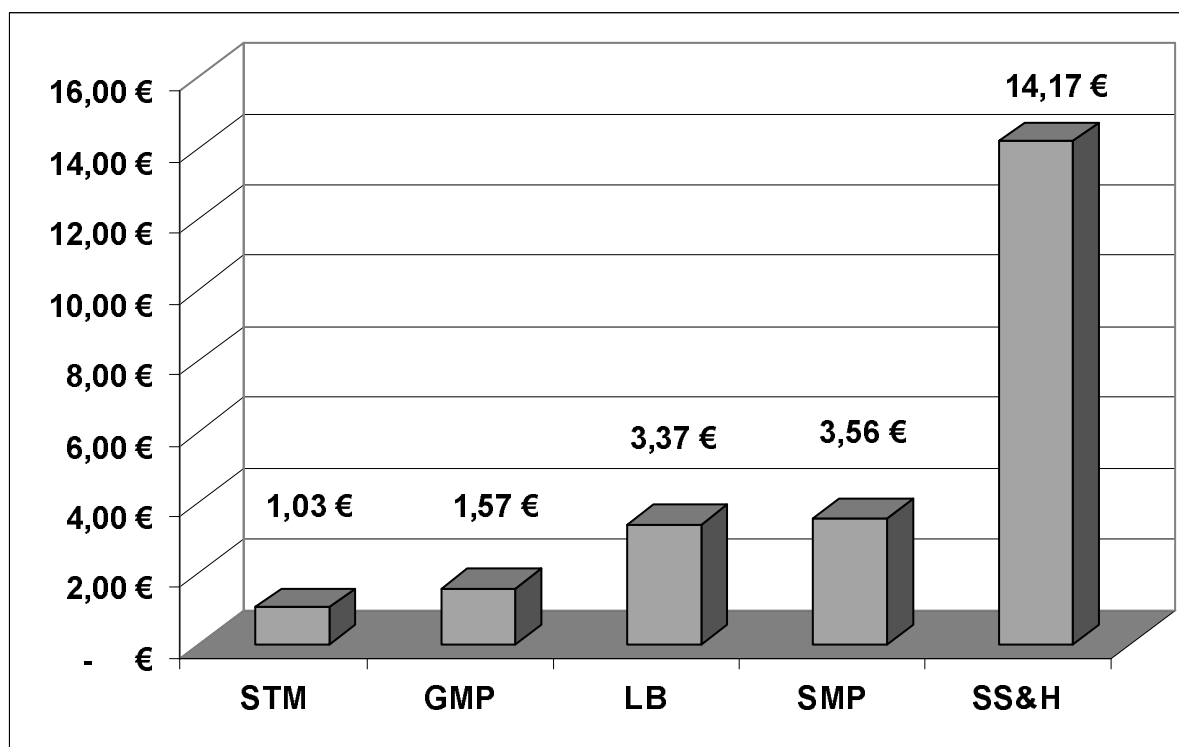


Figure 4: Cost per full-text article request and university type 2005 (acronyms cf. note 6)

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. (2005) 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 4 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 GMD universities.

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

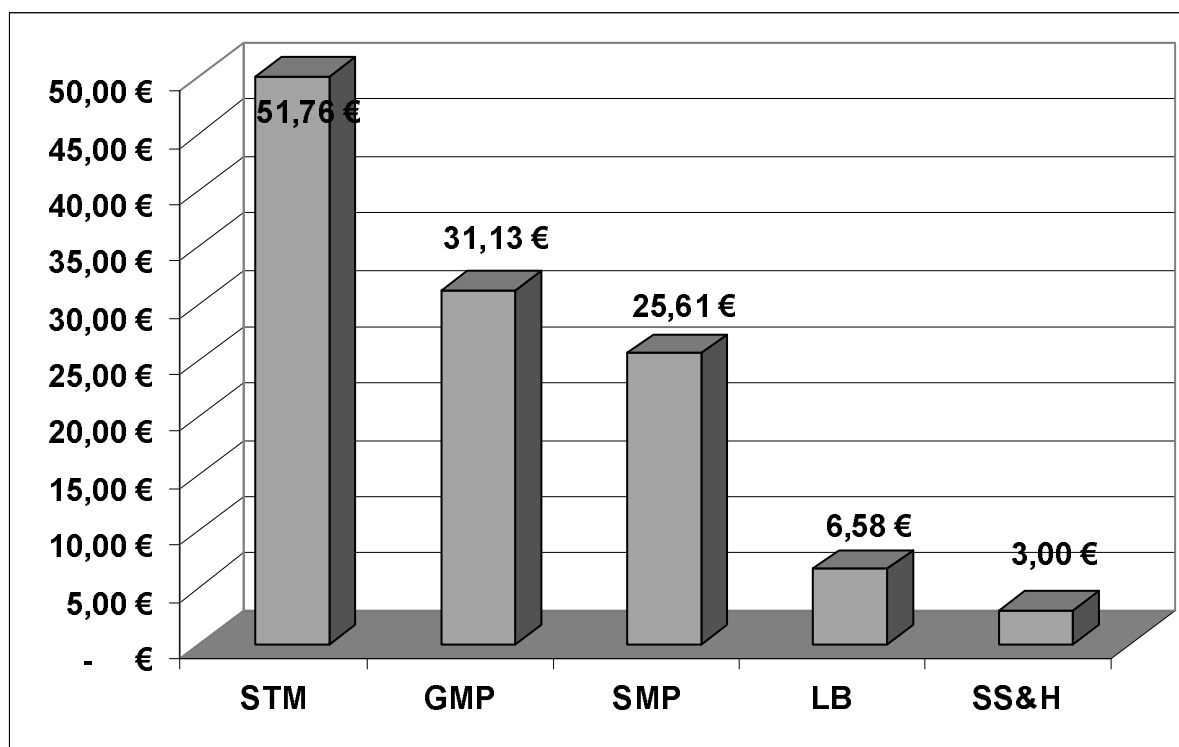


Figure 5: Investment in e-journals per user and university type 2005 (acronyms cf. note 6)

Expenditures for e-journals per user in STM universities were 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).

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 were 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 revealed often to be of lesser value, in particular concerning data on investment and staffing. This again induces risk of error¹⁰ 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 did not address so far is the possible overestimation of usage by double downloads (HTML and PDF, double clics etc.).

Even so, taking into consideration the potential impact of these aspects, the collected data of the first phase of the EPEF project appear 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.

¹⁰ E.g. a latent overestimation of the two metrics “full text article requests per user” and “costs per user”.

- 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 rather 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 probably can be explained by four factors: (1) With explicit support by the government, the national academic consortium in the first years clearly gave preference to licenses 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 develop only recently¹¹. (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 (Revue.org, Persée, CAIRN) have not been available. Nevertheless, as the statistical data show a significant growth of usage 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 by Boukacem & Kamga (2008). 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 factors¹².
- 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, consortium- (nation-) 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, nationwide data are not sufficient for an accurate analysis of return on investment (scholarly outcomes related to access and information-

¹¹ Cf. the digital collections from CAIRN, Revue.org and Persée.

¹² Cf. <http://www.uksg.org/usagfactors>

seeking behaviour) and need to be completed by local information, even if the latter may be biased, too¹³.

Monitoring the usage of digital resources gains increasing importance 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 organisations (CNRS, INSERM etc.) in order to provide realistic understanding of the “who uses what and how, and who pays for it”.

Actually the French academic landscape is changing, 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

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>

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>

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/

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>

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/

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

Franklin, B, Managing the Electronic Collection with Cost per Use Data, *IFLA Journal* 2005, 31(3), 241-248. Available at: <http://ifl.sagepub.com>

¹³ E.g. underestimation of off campus counts.

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

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

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

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.

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.

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

Annex 1: Most requested titles by university type (2005 sample)

STM (sciences, technology, medicine)

<i>Tetrahedron Letters</i>
<i>Tetrahedron</i>
<i>Cell</i>
<i>American Journal of Cardiology, The</i>
<i>Angewandte Chemie International Edition</i>
<i>Lancet, The</i>
<i>Journal of Chromatography A</i>
<i>Biochemical and Biophysical Research Communications</i>
<i>NeuroImage</i>
<i>Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with</i>

<i>Materials and Atoms</i>

GMD (grand multidisciplinary)

<i>Tetrahedron Letters</i>
<i>Lancet, The</i>
<i>Tetrahedron</i>
<i>Biochemical and Biophysical Research Communications</i>
<i>Journal of Membrane Science</i>
<i>Bioorganic & Medicinal Chemistry Letters</i>
<i>Polymer</i>
<i>Archives de Pédiatrie</i>
<i>Earth and Planetary Science Letters</i>
<i>Annales Francaises d'Anesthesie et de Reanimation</i>

SMD (small or medium-sized multidisciplinary)

<i>Tetrahedron Letters</i>
<i>Tetrahedron</i>
<i>Journal of Sound and Vibration</i>
<i>Journal of the European Ceramic Society</i>
<i>Solid State Ionics</i>
<i>Water Research</i>
<i>Polymer</i>
<i>Thin Solid Films</i>
<i>Composites Science and Technology</i>
<i>Catalysis Today</i>

SS&H (social sciences and humanities, including arts)

<i>Personality and Individual Differences</i>
<i>Annales Medico-psychologiques, revue psychiatrique</i>
<i>International Journal of Heat and Mass Transfer</i>
<i>Social Science & Medicine</i>
<i>Neuropsychologia</i>
<i>Journal of Archaeological Science</i>
<i>Animal Behaviour</i>
<i>Neuropsychiatry de l'Enfance et de l'Adolescence</i>
<i>Annals of Tourism Research</i>
<i>NeuroImage</i>

LB (law and business, including economics and management)

<i>Journal of Public Economics</i>
<i>Economics Letters</i>
<i>European Economic Review</i>
<i>Journal of Economic Theory</i>

<i>Journal of Financial Economics</i>
<i>Journal of Econometrics</i>
<i>Journal of Business Research</i>
<i>Journal of Banking & Finance</i>
<i>Journal of Corporate Finance</i>
<i>Journal of Accounting and Economics</i>