



HAL
open science

Statistiques d'utilisation des ressources électroniques : le projet COUNTER

Chérifa Boukacem, Joachim Schöpfel

► **To cite this version:**

Chérifa Boukacem, Joachim Schöpfel. Statistiques d'utilisation des ressources électroniques : le projet COUNTER. Bulletin des bibliothèques de France, 2005, 50 (4). sic_00001543

HAL Id: sic_00001543

https://archivesic.ccsd.cnrs.fr/sic_00001543

Submitted on 6 Sep 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistiques d'utilisation des ressources électroniques : le projet COUNTER

Chérifa Boukacem

URSIDOC

boukacemc@yahoo.fr

Joachim Schöpfel

INIST-CNRS

schopfel@inist.fr

La connaissance précise de l'utilisation des ressources électroniques en ligne est un enjeu majeur dans les négociations avec les éditeurs et pour la répartition des coûts d'accès. La question du « qui utilise quoi quand et comment » détermine de plus en plus la politique tarifaire des éditeurs et le prix des licences, en particulier à un moment où l'on commence à progresser des « big deals » vers un achat plus sélectif des ressources (Roberts et al. 2004). De l'autre côté, cette question influence la répartition des dépenses et coûts entre organismes, unités de formation, laboratoires, facultés, services etc. – un aspect crucial pour les consortia et établissements multi-sites/multi-services.

Les chiffres d'utilisation ont également servi à justifier, surtout dans les premières années d'accès, les dépenses et demandes budgétaires pour les revues et bases de données en ligne. Mais comment mesurer cette utilisation, comment la suivre dans le temps, comment aussi comparer les chiffres donnés par les différents éditeurs, hébergeurs et autres agrégateurs ?

Vers un renouvellement des savoir-faire bibliothéconomiques ?

La conduite d'une politique documentaire au sein d'une bibliothèque se fait par la confrontation de considérations bibliothéconomiques (constitution, cohérence et suivi des collections, demande des usagers,...) et de considérations économiques (budget d'acquisition et coûts de la documentation scientifique). Pour les bibliothèques universitaires et de recherche, compte tenu des coûts de la documentation scientifique et de recherche, cette confrontation a souvent été à la source de crises budgétaires.

Dans le contexte papier, peu de données autour de la consultation des collections ont pu apporter un éclairage supplémentaire, capable de mieux orienter les décisions de la politique documentaire. Dans les domaines STM, l'activité de fourniture de documents a parfois donné des indices sur les tendances de la consultation : elle a permis d'identifier les titres les plus demandés ou au contraire, les moins demandés, en vue d'un réabonnement ou bien d'un désabonnement. Cette « méthode » d'évaluation de l'intérêt économique pour un réabonnement, répandue dans le milieu professionnel, est à l'origine de l'articulation économique « doit-on accéder ou doit-on acquérir ? » (Line 1996). La politique documentaire des bibliothèques a pu donc, le cas échéant, être mieux ciblée.

Aujourd'hui, la conduite d'une politique documentaire se fait en tenant compte des contraintes du cadre contractuel des licences qui se sont imposées comme les modalités d'accès à la documentation scientifique. Dès lors, les bibliothèques n'acquièrent plus des documents physiques, mais un accès à distance à des documents en ligne.

Cet accès se fait par l'intermédiaire du réseau, qui devient un support d'enregistrement et de recueil possible des données de consultation. Cette reconfiguration laisse entrevoir la possibilité du recueil des données chiffrées autour de l'utilisation des ressources en ligne et de leur exploitation statistique (pour une synthèse récente, voir le rapport de Tenopir 2003). Pour les bibliothèques, les enjeux de cette question sont de deux ordres, indissociables et majeurs : l'un relatif à la politique documentaire, l'autre aux ressources financières à investir dans le cadre de cette politique documentaire.

Il nous semble toutefois pertinent de préciser que les vecteurs de diffusion étant différents d'un domaine à un autre, les enjeux décrits ci-dessus ne revêtent donc pas la même importance d'un domaine à un autre et a fortiori d'une bibliothèque à une autre ; le face à face STM – SHS, avec le droit et les sciences économiques en position médiane, forment une illustration typique (Boukacem 2004). Les travaux de Christine Okret-Manville (Okret-Manville 2002) et de Claude Jolly (Jolly 2003) concourent à pointer l'inégalité de la couverture documentaire électronique, d'un domaine à un autre et parfois d'une bibliothèque à une autre à l'intérieur d'un même domaine.

Ainsi donc, pour une bibliothèque, il n'est donc plus question de réajuster son périmètre documentaire avec de simples tendances, mais avec des données précises, qui permettent aux professionnels d'utiliser au mieux leurs budgets pour répondre à la demande de leurs usagers. De manière corollaire, ces données fournissent aux professionnels des bibliothèques des arguments concrets lors des négociations avec les éditeurs, et peut-être de réussir à infléchir ces derniers sur leurs prix.

La mise à disposition de données précises, véritables arguments lors des négociations avec les éditeurs, nous semble d'autant plus importante aujourd'hui que l'actualité des négociations des consortiums en Amérique du Nord et en Europe montre que les bibliothèques voient leur poids s'amoinrir dans le bras de fer inflexible des éditeurs.

L'ARL tire la sonnette d'alarme et rapporte qu'en 2003 (Case, 2004), « face aux renouvellements de contrats, sur les 57 des bibliothèques interrogées, 5 ont déjà supprimé au moins un package, 22 bibliothèques envisagent de le faire et 16 autres bibliothèques pourraient le faire au vu des difficultés budgétaires qu'elles rencontrent ».

Le consortium de l'OhioLINK pour sa part, avait publié une étude (Sanville, 2001) faisant état de l'explosion de la consultation avec l'intégration massive des packages électroniques. Cependant, après quelques années de fonctionnement, et une fois l'effet « nouveauté » passé, les membres du consortium éprouvent les limites de leurs budgets. Le consortium revient donc à une politique d'évaluation pour une meilleure sélection des titres à partir de 2005 (« unbundling »). Par ailleurs, des institutions prestigieuses et richement dotées, telles que Harvard, envisagent de se désengager de certains packages afin de se donner le temps de mieux penser leur ligne directrice à venir.

Car principale clientèle des éditeurs, les bibliothèques sont la première cible des augmentations des coûts des périodiques ; si elles souhaitent continuer dans leur voie de regroupement pour l'acquisition de ressources électroniques, elles doivent prendre conscience des enjeux de l'exploitation des données quantitatives de la demande faite à leurs collections.

En France, le dispositif d'acquisition de ressources électroniques le plus significatif du réseau universitaire, COUPERIN (www.couperin.org), ne permet pas encore d'envisager une

démarche d'évaluation similaire à celles entreprises dans les pays anglo-saxons. De fait, l'indisponibilité des données de consultation ne permet pas encore de se prononcer sur la pertinence de l'offre de COUPERIN. Le traitement de données exploitables permettrait d'identifier des tendances de la consultation et de se pencher sur leur signification.

Le cadre d'interprétation tiendrait à la fois des spécificités de l'offre du consortium, de la politique documentaire des Services Commun de Documentation (SCD) et de leurs objectifs de couverture documentaire et de diffusion de l'information. Ceci permettrait en même temps de dégager les insuffisances de l'offre électronique commerciale pour la réajuster aux priorités des SCD. Le but ultime serait de donner des arguments concrets aux professionnels des bibliothèques lors des renégociations de bouquets, en terme d'offre et/ou de prix.

Il est toutefois quelque peu surprenant de constater qu'en comparaison avec les débats et discussions animées dans d'autres pays, cette question reste peu soulevée dans la communauté professionnelle française. On a pu constater quelques messages diffusés sur la liste Biblio.fr qui manifestent une préoccupation certaine, mais point encore de réflexion qui laisserait entrevoir une mobilisation nationale.

Soulignons que les ressources électroniques, particulièrement les périodiques en ligne, sont en acquisition croissante : les résultats de l'enquête sur les ressources électroniques pour l'année 2001 (www.sup.adc.education.fr/bib) repris dans l'annuaire des bibliothèques universitaires de l'année 2001 montrent que le nombre des titres de périodiques en ligne dans le réseau universitaire français ont augmenté de 148%. De façon corollaire, les dépenses documentaires sont en augmentation croissante avec plus de 82% entre 1998 et 2001 avec 19,4% d'augmentation chaque année pour les périodiques (Annuaire, 2001). Comme le précise Claude Jolly, cette augmentation ne donne « qu'une image amoindrie de la réalité : on sait en effet que la facturation des périodiques électroniques ne prend le plus souvent en compte qu'un surcoût calculé en fonction de la charge des abonnements sur papier » (Jolly 2003).

Après avoir porté tout leur effort sur l'offre et sur l'intégration de cette offre dans leurs fonds, il serait temps pour les bibliothèques de prendre conscience des enjeux de leur implication dans la compréhension de la demande. Aujourd'hui, le rôle des bibliothèques ne se limite plus à une offre de services. C'est aussi une confrontation de l'offre à la demande des communautés scientifiques pour mieux cerner les exigences de leurs publics et éventuellement savoir si elles peuvent ou veulent y répondre (Duy 2004).

Approches pour une harmonisation des statistiques

Pour résumer, les professionnels de la bibliothèque ont besoin de statistiques pour quatre raisons : évaluer la valeur des différents produits et services en ligne, mieux argumenter les décisions d'achat, planifier l'infrastructure technique et promouvoir les services de la bibliothèque. Mais ces données sont livrées avec une périodicité, un format et un traitement variables, hétérogènes, parfois à l'état brut. En outre, leur présentation varie d'un éditeur à l'autre. Et certains éditeurs ne communiquent tout simplement aucune information sur l'usage de leurs ressources. Et quand il y en a, est-on sûr de bien comprendre ce qui a été compté ? Quel est le sens exact par « search », « request », « turnaway », « full-text download » ou « session » ? Que faut-il entendre de « item » ou « article » ? Tous les éditeurs définissent-ils ces notions de la même façon, les comptent-ils de manière comparable ?

Pour être utiles, les statistiques doivent être crédibles, cohérentes et compatibles (= les « trois C » de Shepherd 2004). Or, le plus souvent elles ne le sont pas, et il est pratiquement

impossible de comparer les chiffres de différents éditeurs, agences etc. Si chaque acteur produit les statistiques à sa façon ou pas du tout, les chiffres deviennent plus ou moins inutilisables voire inutiles pour les besoins de la bibliothèque. Et puisque cela ne mène nulle part, il y a eu à partir du milieu des années 90 une prise de conscience du problème et plusieurs efforts pour harmoniser et normaliser la collecte et la présentation des statistiques (Duy 2004, Shepherd 2003, 2004).

Citons parmi les tentatives les plus significatives les recommandations de JSTOR pour mesurer l'usage des ressources sur le Web et, basées sur le travail de JSTOR, celles d'ICOLC pour la collecte et diffusion de ces données (recommandations de 1998 mises à jour 2001) qui restent aujourd'hui encore une sorte de standard « de facto » pour les éditeurs – même si beaucoup d'éditeurs ne les appliquent pas.

De son côté, l'association des bibliothèques de recherche américaines a financé dès 1998 une étude sur les mesures pour évaluer d'une manière exhaustive et efficace l'usage des ressources électroniques (ARL e-metrics).

D'autres démarches ont suivi, de portée plus limitée, parfois aussi au niveau d'un seul pays (Pesch 2004). Une seule initiative cependant a réussi à intégrer les différents projets et approches, à fédérer les principaux protagonistes et acteurs et à susciter suffisamment d'adhésion et de dynamisme pour devenir en deux ans une sorte de norme internationale pour les statistiques d'utilisation : le projet COUNTER (Counting Online Usage of Networked Electronic Resources).

Création et objectif du projet COUNTER

COUNTER a vu le jour en mars 2002. Parmi les membres fondateurs on trouve le gotha de l'information scientifique et technique : Elsevier, Taylor & Francis, Oxford University Press, Blackwell Publishing, Lippincott Williams & Wilkins, Nature Publishing Group, Institute of Physics Publishing, Association of Learned & Professional Society Publishers (ALPSP), Association of American Publishers, International Association of Scientific Technical & Medical Publishers (STM), Association of Subscription Agents and Intermediaries (ASA), EBSCO, Ingenta, ProQuest, l'Association of Research Libraries (ARL), le Joint Information Systems Committee (JISC) et la UK Serials Group (UKSG). Dès le début, l'organisme international de normalisation NISO (National Information Standards Organization) est associé au projet. Ces différents groupes et acteurs sont représentés au comité de direction et au comité consultatif international du projet.

La première année a été consacrée à faire connaître le projet, surtout auprès de l'industrie de l'information, et à préparer des recommandations. La version 1 du « Code de bonnes pratiques de COUNTER » (code COUNTER) a été publiée en janvier 2003 et s'adresse aux bibliothécaires, éditeurs et autres fournisseurs d'information. Il donne des orientations sur les données à mesurer, leur définition, le contenu et format de leur présentation (rapports) ainsi que sur l'exploitation et l'évaluation des données.

L'information sur COUNTER est assurée principalement via le site du projet hébergé par l'Institut of Physics (www.projectcounter.org). Promotion du code et formation se font à travers de conférences (en 2003 à New York, Munich, Nîmes etc.), ateliers et stages sur site. Il existe également une liste de diffusion (usage@jiscmail.ac.uk).

En 2003, l'équipe du projet a créé « Counter Online Metrics », une société à but non lucratif de droit anglais, afin de pérenniser le travail du projet.

Afin de faire connaître ce projet en France et faciliter la compréhension des recommandations, l'INIST a traduit avec l'accord du directeur de projet les documents de référence en Français. Ces traductions sont mises en ligne sur le site de l'INIST (<http://www.inist.fr/INIST/counter.php>) et sont régulièrement mises à jour.

Contenu du code COUNTER

Le projet COUNTER est conçu comme un « code de bonnes pratiques » unique, international, flexible permettant de mesurer l'utilisation des produits et services en ligne d'une manière crédible, cohérente et compatible, en se fondant sur les données fournies par le fournisseur.

Le code COUNTER spécifie le format, le contenu, la périodicité et le mode de transmission des statistiques pour des revues et bases de données sous forme de sept « rapports d'utilisation » (« usage reports »). Le niveau de base pour mesurer l'utilisation des revues est le « journal report 1 » qui chiffre le nombre de requêtes réussies d'un article en texte intégral, par revue et par mois :

	ISSN print	ISSN online	Janv - 01	Févr - 01	Mars - 01	Cumul année
Total revues			6637	8732	7550	45897
Revue AA	1212-3131	3225-3123	456	521	665	4532
Revue BB	9821-3361	2312-8751	203	251	275	3465
Revue CC	2464-2121	0154-1521	0	0	0	0
Revue DD	5355-5444	0165-5542	203	251	275	2978

Chaque revue est identifiée par son titre et ISSN (print et/ou online). Le tableau contient également le cumul annuel pour chaque revue et le cumul mensuel pour l'ensemble des titres.

Les autres « rapports » incluent entre autre les interrogations sans accès au texte intégral, le format du document (HTML, PDF) ou le nombre des séances et les chiffres pour les bases de données.

Pour la transmission des données, COUNTER préconise deux formats (fichiers .csv ou .xls), la mise à disposition sécurisée sur un serveur, une périodicité au moins mensuelle, une mise à jour dans les quatre semaines après la fin du mois et une information sur l'année en cours N et l'année N-1.

Bilan intermédiaire du projet COUNTER et perspective

Aujourd'hui, le « journal report 1 » tend à devenir un standard pour les éditeurs. Les plus importants éditeurs et fournisseurs d'information produisent déjà des statistiques d'utilisation "conformes à COUNTER" ou sont en train de se mettre en conformité avec le code : Elsevier, Blackwell, Wiley, OUP, Nature, Muse et HighWire figurent sur la liste mais également l'American Chemical Society, ISI, EBSCO et récemment Swets (cf. la liste des « compliant vendors » à www.counterproject.org/articles.html). La production des éditeurs « conformes » représenterait 40-50% des articles de revues avec facteur d'impact.

En marge de la conférence annuelle de l'UKSG fin mars 2004 à Manchester, Counter Online Metrics a tenu sa première assemblée générale. A cette date, la société comptait 114 membres dont 42 éditeurs et intermédiaires (« vendors »), 29 consortia, 29 bibliothèques et 14 autres (associations professionnelles, lobbys...). L'objectif est d'augmenter le nombre à 150-200 membres. Il s'agit d'une participation internationale ; néanmoins il n'y pas encore de membres francophones.

Les principes d'audits ont été déterminés, une DTD XML pour les rapports d'utilisation a été publiée sur le site du projet en juillet 2004, et cinq sites sont en phase d'évaluation pour analyser comment le code COUNTER fonctionne et est réellement appliqué. La prochaine assemblée générale aura lieu 2005 aux Etats-Unis.

La première version du code se concentre sur l'utilisation des revues et bases de données, produits qui représentent traditionnellement la majeure partie des budgets documents des bibliothèques. Les versions ultérieures élargiront le champ d'application de COUNTER, pour y intégrer non seulement de nouveaux types de contenus, mais aussi des niveaux plus détaillés de mesures pour chaque type de contenu.

Une 2^e version de ce code rentrera en vigueur en 2006 ; le texte validé se trouve sur le site du projet depuis mars 2005. En fait, il s'agit d'une modification légère du premier code en tenant compte de l'évaluation en cours (ex. : ajout du nom d'éditeur et de la distinction HTML/PDF dans le « journal report 1 »).

Le document de travail d'un code pour des livres électroniques (« books and reference works ») se trouve sur le site du projet depuis janvier 2005 et est ouvert à commentaire et critique jusqu'en décembre 2005, pour entrer en vigueur en 2006 ou 2007.

Le principe d'une harmonisation et simplification des statistiques d'utilisation rencontre un vif succès. Les ateliers COUNTER figurent depuis 2003 parmi les formations les plus demandées des conférences annuelles du UK Serials Group. Début octobre 2004, la NFAIS a mis les statistiques d'utilisation à l'ordre du jour d'une réunion internationale à New York. Dans les négociations, la conformité avec les recommandations COUNTER est de plus en plus utilisée par les éditeurs comme argument de vente, par les bibliothèques comme critère d'évaluation et de sélection.

En France, l'accès aux revues et bases de données en ligne progresse rapidement. Les négociations se poursuivent dans un rythme annuel et tri-annuel, et les éditeurs français commencent à « aller sur le Web ». Dans ce contexte, les statistiques revêtent une importance grandissante et stratégique. Supposons que les professionnels des bibliothèques rattrapent très vite le retard, prennent conscience des enjeux, organisent des formations aux statistiques et exigent des éditeurs et fournisseurs des données d'usages fiables et utiles ; et espérons que les éditeurs qui aujourd'hui commencent à publier des revues en ligne incluent dès le début ces statistiques dans leur cahier des charges.

Juillet 2004 (révisé février 2005)

- ARL e-metrics : www.arl.org/stats/newmeas/newmeas.html (visité le 12 juin 2004).
- Boukacem, Chérifa : *Du prêt entre bibliothèques à la bibliothèque virtuelle : une évolution contrastée des bibliothèques universitaires françaises selon la discipline*. Soutenue en mai 2004. Université Lyon 2.
- Case, M. : « Big deal or bad deal : libraries and aggregation ». The 2004 Allen Press Emerging Trends Seminar, Avril, 2004. www.allenpress.com (visité le 10 décembre 2004).
- COUNTER : www.projectcounter.org (visité le 15 février 2005).
- Duy, Joanna : « Usage data : issues and challenges for electronic resource collection management ». Fowler, D.C. (éd.) : *E-Serials Collection Management : Transitions, Trends and Technicalities*. New York : Haworth 2004, pp 111-138.
- ICOLC : www.library.yale.edu/consortia/2001webstats.htm (visité le 12 juin 2004).
- Jolly, Claude : « Documentation électronique à l'université », *Bulletin des Bibliothèques de France* 2003, vol 48, no 4, pp 5-8.
- JSTOR : www.jstor.org/news/newsletter/no2/iss2/taskforce.html (visité le 12 juin 2004).
- Line, Maurice B. : « Accéder ou acquérir : une véritable alternative pour les bibliothèques ? » *Bulletin des Bibliothèques de France* 1996, vol 41, no 1, pp 32-34.
- Ministère de l'Éducation Nationale : *Annuaire des Bibliothèques Universitaires et des Grands Établissements*. Paris : La Documentation Française 2003. – 119 p.
- Okret-Manville, Christine : « Le développement numérique : un facteur d'intégration des bibliothèques dans les établissements d'enseignement supérieur ». In : Chartron, Ghislaine : *Les chercheurs et la documentation numérique : nouveaux services et usages*. Paris : Editions du Cercle de la Librairie, 2002. – 268 p.
- Pesch, Oliver : « Usage Statistics : Taking E-Metrics to the Next Level ». *The Serials Librarian* 2004, vol 46, no1/2, pp 143-154.
- Roberts, Michael, Kidd, Tony, Irvine, Lynn : « The impact of the current e-journal marketplace on university library budget structures : some Glasgow experiences ». *Library Review* 2004, vol 53, no 9, pp 429-434.
- Sanville, Thomas J. : « L'utilisation des journaux électroniques dans l'OhioLINK electronic journal center ». <http://www.ifla.org/IV/IFLA67/papers/118-141f.pdf> (visité le 15 juillet 2001).
- Shepherd, Peter T. : « The COUNTER Code of Practice : implementing and adaption ». *SCONUL Newsletter* 2003, Summer/Autumn, no 29, pp 22-26.
- Shepherd, Peter T. : « Project COUNTER : Setting International Standards for Online Usage Statistics ». *Joho Kanri Journal of Information Processing and Management* 2004, vol 47, no 4, pp 245-257.
- Tenopir, Carol : *Use and Users of Electronic Library Resources: An Overview and Analysis of Recent Research Studies*. Council on Library and Information Resources 2003. www.clir.org/pubs (visité le 22 juin 2004).
- Annuaire des Bibliothèques Universitaires : 2001. – Ministère de l'éducation nationale, 2003.

Chérifa Boukacem est chercheur associé à l'équipe d'accueil URSIDOC (Enssib – Lyon 1). Elle a soutenu (mai 2004) une thèse dirigée par Jean-Michel Salaün, en sciences de l'information et de la communication à l'Université Lyon 2. L'intitulé de sa thèse est « Du prêt entre bibliothèques à la bibliothèque virtuelle : une évolution contrastée des bibliothèques universitaires françaises selon la discipline ».

Joachim Schöpfel a été de 1985-1990 enseignant-chercheur en psychologie à l'université de Hamburg (Allemagne). Thèse de 3^e cycle en psychologie d'éducation en 1992. Depuis 1992, il enseigne à l'université de Nancy 2 en sciences de l'éducation et en documentation d'entreprise. Il travaille depuis 1991 à l'INIST-CNRS, d'abord comme documentaliste dans les bases de données, depuis 1999 comme responsable de la bibliothèque et depuis 2004 aussi de la fourniture des documents. Il est membre de l'UK Serials Group et de l'association européenne EAGLE (littérature grise).

Résumé :

L'accès aux ressources électroniques se développe de plus en plus dans les bibliothèques et services de documentation français. La connaissance des statistiques d'utilisation des revues et bases de données en ligne devient un enjeu important pour l'analyse des besoins et les décisions d'achat. L'article présente les recommandations du projet COUNTER qui deviennent une norme internationale pour le recueil et le traitement de données chiffrées. .

The access to electronic resources in French libraries and documentation services is increasing rapidly. The knowledge of usage statistics of online serials and databases becomes more and more important for need analysis and acquisition decisions. The article describes the recommendations of the project COUNTER which are de facto going to be an international norm for the production of these statistics.

Der Zugang zu elektronischen Ressourcen erlebt eine rasante Entwicklung in den französischen Bibliotheken und Dokumentationszentren. Die Kenntnis der Nutzungsstatistiken von Online-Zeitschriften und Datenbanken wird zunehmend wichtig für Bedarfsanalyse und Kontentbeschaffung. Der Artikel beschreibt die Richtlinien des Projekts COUNTER, die dabei sind, zu einer internationalen Norm für die Produktion dieser Statistiken zu werden.