

Usages

Usages des revues électroniques

Annaïg Mahé

Site web du guide
des ressources
pour l'édition de
revues numériques

Coordination :
Ghislaine Chartron
et Jean-Michel Salaun

Annaïg Mahé
Doctorante en science de
l'information

Bilan sur les études d'usage des revues électronique

La grande majorité des revues électroniques mises en ligne jusqu'à aujourd'hui émanant des domaines des sciences de la nature, les études d'usages se sont principalement basées sur ces domaines et la synthèse proposée ici reflète ce biais.

Cette synthèse est basée sur deux types de sources : les études d'usages des revues électroniques principalement publiées dans la littérature anglo-saxonne, et les résultats d'un travail de thèse étudiant ces usages auprès d'échantillons de chercheurs français.

Une évolution rapide

Préambule

Première période : le non-usage des revues électroniques

Deuxième période : des usages en croissance

Les facteurs d'intégration des revues électroniques dans les pratiques des chercheurs

Différenciation de l'intégration dans les pratiques

Complémentarité papier/électronique

La lecture à l'écran : éléments d'appréciation pour les revues

Avantages/inconvénients des revues électroniques

Références

Evaluation des usages

L'évaluation des usages peut être réalisée à partir de méthodes d'investigation quantitatives et/ou qualitatives. Cependant, il est nécessaire de garder à l'esprit que ces différentes méthodes n'ont pas les mêmes objectifs et qu'aucune n'est complète en soi. Une étude approfondie croisera différents types de données et d'analyses. La triangulation des méthodes permet effectivement de consolider les résultats obtenus selon une approche en les rapprochant de ceux obtenus selon un autre type d'approche.

Les méthodes quantitatives

Intérêts et limites

Traitement des données brutes

Définition des données à analyser

Fourniture et format des données

Références

Les méthodes qualitatives

Intérêts et limites

Description de différents types de méthodes

L'analyse des données

Références

Droits d'auteurs
Réservés



Bilan sur les études d'usage des revues électroniques

Usages des revues électroniques

Annaïg Mahé

Une évolution rapide

Préambule

Les études d'usages ayant suivi le développement de l'offre, il est possible d'en distinguer deux périodes bien distinctes : une première période débute avec la mise en ligne des toutes premières revues vers la fin des années 70 et se poursuit jusqu'au début des années 90 ; une seconde période commence vers les années 1994-95 avec l'arrivée du Web et le foisonnement de l'offre électronique que ce support permet. Les études d'usages se multiplient alors rapidement. Actuellement, nous pouvons considérer qu'une troisième période d'études d'usages se met en place, basée sur les possibilités croissantes d'effectuer des analyses statistiques sur plusieurs années d'usages effectifs et non plus dans des cadres exclusivement expérimentaux.

Première période : le non-usage des revues électroniques

Cette première période se caractérise principalement par un foisonnement d'expérimentations ayant pour but de tester la faisabilité technique et la viabilité économique des revues électroniques naissantes. Ces études sont menées par des équipes universitaires ou par de grands éditeurs commerciaux, et souvent dans le cadre de partenariats entre ces éditeurs et plusieurs sites universitaires.

En ce qui concerne les usages par les chercheurs, le constat majeur est celui du non-usage, et celui-ci s'explique assez aisément :

- l'offre mise à disposition ne dispose pas d'une masse critique suffisante pour entraîner le besoin d'apprentissage d'un support nouveau ;
- les barrières techniques sont encore trop nombreuses : beaucoup d'utilisateurs n'ont pas d'accès direct aux terminaux informatiques ; lorsque c'est le cas, les possibilités d'affichage et d'impression sont trop limitées, rendant l'usage particulièrement fastidieux ;
- les auteurs ne sont pas suffisamment motivés pour diffuser leurs résultats sur un nouveau support uniquement électronique : non seulement ce support n'offre pas de garanties suffisantes quant à la pérennité des publications, mais les chercheurs préfèrent surtout



publier dans des revues reconnues par la communauté scientifique, ce qui n'est pas le cas des nouvelles revues électroniques, souffrant fortement d'un manque de reconnaissance institutionnelle.

Usages des revues électroniques

Annaïg Mahé

De fait, si certaines de ces premières expériences sont des prouesses techniques, elles constituent une sorte de « faux départ » de la publication électronique par le manque d'auteurs et de lecteurs. La dimension économique reste, de même, peu probante du fait de ce non-usage. Leur intérêt du point de vue des usages aura été de remettre au premier plan l'importance des facteurs socio-cognitifs dans le processus d'appropriation d'une nouvelle technologie.

Deuxième période : des usages en croissance

Le contexte est alors complètement différent : les technologies ont mûri et sont plus facilement accessibles, et surtout le Web s'est imposé comme interface d'accès. En ce qui concerne les études d'usages, une des priorités de la majorité de ces recherches est de mieux comprendre les facteurs socio-cognitifs sous-estimés jusque-là, en replaçant l'usage particulier des revues électroniques dans le contexte plus général de l'activité des chercheurs. Certaines études se basent sur des méthodes uniquement quantitatives, d'autres privilégient les approches qualitatives, les plus complètes croisant plusieurs approches (questionnaires, entretiens, études de panels, etc.).

Les premiers résultats ne font que confirmer ceux de la première vague d'études : l'usage des revues électroniques reste marginal et contraste avec l'usage fortement croissant des outils de communication électronique, notamment la messagerie. Les premières raisons de cet usage faible, voire du non-usage, restent les mêmes que celles préalablement analysées :

- une offre encore peu développée et difficilement accessible : la non-disponibilité de titres électroniques ainsi que le manque d'informations sur les titres en ligne existants renforcent dans un premier temps l'attachement des chercheurs au support papier et le non-usage des revues électroniques ;
- des barrières techniques qui restent encore fortes : lorsque une offre est disponible, les difficultés liées à la technique rebutent souvent les usagers potentiels. L'essai n'est alors pas transformé et l'usage ne devient pas récurrent.
- la mauvaise perception des revues électroniques persiste : vers 1994-95, les grands éditeurs scientifiques n'ont pas encore effectué le transfert de leurs collections au format électronique et les revues électroniques continuent de souffrir d'un manque de reconnaissance, sans doute accentué par la confusion qui caractérise leur définition par les chercheurs.



Usages des revues électroniques

Annaïg Mahé

Cependant, la plupart de ces études notent une acceptation croissante de ce nouveau support. A partir des années 1997-98, les éditeurs principaux dans les domaines des sciences de la nature sont en mesure de proposer la plus grande partie de leurs collections au format électronique et la grande majorité des revues électroniques disponibles sont donc des versions de revues papier reconnues par la communauté scientifique. Ceci permet une plus grande confiance des usagers envers ce support qui conserve, d'une certaine manière, une forme traditionnelle. A cela s'ajoute la multiplication des projets expérimentaux (dans de nombreuses universités anglo-saxonnes) qui joue un rôle non négligeable dans la familiarisation de nombreux usagers. Dans ce contexte, les études d'usages approfondissent considérablement la compréhension des facteurs d'intégration des revues électroniques dans les pratiques des chercheurs. Par ailleurs, les études quantitatives basées sur les données de consultation commencent à se multiplier et les premiers résultats permettent déjà de titrer la troisième phase : « une explosion des usages ».

Les facteurs d'intégration des revues électroniques dans les pratiques des chercheurs

Différenciation de l'intégration dans les pratiques

La discipline de recherche

Un des constats majeurs des études d'usages est l'importance de la culture de communication de la discipline : à ce titre, on retrouve dans les usages du support électronique les différences disciplinaires, voire sous-disciplinaires, relevées pour le support papier. Celles-ci sont notamment fonction de :

- la place de la revue scientifique comme support de communication, et la perception de l'accès à l'information scientifique et technique : rôle de la validation par les pairs, niveau d'urgence du suivi de l'actualité, besoins en antériorité, concentration ou dispersion des titres suivis, flexibilité par rapport au prix, etc.
- les types de contenus : texte, graphiques, images, formules mathématiques, etc.
- la familiarité avec l'outil informatique et les ressources électroniques

L'environnement

Le contexte dans lequel évoluent les usagers joue aussi un rôle important dans la différenciation de l'appropriation des revues électroniques. Dans cette catégorie, nous considérons :



Usages des revues électroniques

Annaïg Mahé

- l'équipement informatique : si le niveau de familiarité avec les outils informatiques est en partie lié à la culture disciplinaire, le niveau d'équipement (et de maintenance) dépend directement des ressources disponibles en local et de leur répartition entre et à l'intérieur des laboratoires de recherche ;
- les habitudes locales : le laboratoire est un lieu de travail où se cristallisent et se transmettent des façons de faire, des habitudes de gestion du travail, du matériel, de communication. Chaque laboratoire est un micro-système qui organise ses ressources à sa façon et le contexte spatial a son importance, notamment en ce qui concerne les échanges d'information entre chercheurs et les types d'ambiance de travail (plus ou moins propices à une activité de recherche bibliographique et de lecture) ;
- les ressources à disposition et leur promotion : en premier lieu, la proximité ou l'absence de bibliothèques de référence influe sur l'usage des ressources électroniques. Quant à ces dernières, un des premiers facteurs de non-usage est la non-connaissance de l'offre à disposition. Dans ce contexte, une attention particulière doit être accordée à leur promotion pour s'assurer que tous les publics potentiels aient accès à cette information ;

La motivation et les préférences personnelles

Au-delà des considérations disciplinaires et contextuelles, les chercheurs ont des préférences individuelles plus ou moins marquées quant à leurs pratiques informationnelles. Certains conservent des habitudes papier fortes tandis que d'autres intègrent rapidement le support électronique. L'âge n'est pas toujours un critère significatif, malgré la perception par de nombreux chercheurs d'une différence générationnelle : les utilisateurs les plus assidus ne sont pas forcément les plus jeunes (même si effectivement les chercheurs débutants arrivent avec un bagage de plus en plus conséquent en ce qui concerne l'utilisation des ressources électroniques). De plus, les étudiants et les jeunes chercheurs sont généralement moins bien lotis que leurs aînés en ce qui concerne l'accès aux ressources électroniques (ainsi qu'aux informations concernant la disponibilité de ces ressources).

Complémentarité papier/électronique

L'accès électronique aux articles scientifiques s'apparente pour le moment plus à une alternative et/ou un complément du support papier qu'à une substitution. Les revues électroniques sont en effet principalement utilisées de manière « traditionnelle », c'est-à-dire sur le modèle de consultation des revues imprimées. Ainsi, les fonctionnalités les plus utilisées sont d'abord les plus basiques : feuilletage des titres, des numéros et des tables des matières, impression des articles, et dans une moindre mesure la recherche à partir d'un moteur (certains chercheurs utilisant ainsi les sites des revues



Usages des revues électroniques

Annaïg Mahé

ou des éditeurs à la manière d'une base de données bibliographiques). Le format PDF est, de loin, le plus utilisé car il permet la reproduction du format classique de l'article scientifique. Par contre, les fonctionnalités avancées, telles que le suivi des liens vers d'autres articles, la sauvegarde des références ou les possibilités d'interaction, sont utilisées de façon plus marginale. Les services d'alertes par messagerie sont de même utilisés de façon très inégale selon les types d'activités scientifiques : plus un chercheur est autonome dans sa recherche d'information et plus il préférera consulter directement la revue en ligne plutôt que d'encombrer sa messagerie. A l'inverse, plus un chercheur fait appel à des services intermédiaires pour le seconder dans sa recherche et plus il appréciera ce type de service le forçant à suivre l'actualité de son domaine.

Parmi les usages les plus avancés, l'on constate notamment une consultation croissante de la part de chercheurs traditionnellement attachés à la consultation des articles validés des « prépublications » sur les sites des éditeurs (articles mis en ligne sans pagination ni date de publication). Une autre pratique émergente « forte » est celle de la mise en place de véritables bibliothèques électroniques personnelles. Les chercheurs pratiquent en effet de plus en plus le « double archivage » : ils impriment une copie de l'article et sauvegardent le fichier PDF sur leur disque dur.

La lecture à l'écran : éléments d'appréciation pour les revues

Dans le discours des chercheurs, la lecture à l'écran apparaît généralement comme un des obstacles majeurs de l'usage des revues électroniques. Ceci est effectivement à considérer en fonction des différents types de lecture que demande l'activité de recherche : une lecture exhaustive et approfondie d'un document se fait plus facilement sur une copie papier que l'on peut transporter et annoter que sur un écran d'ordinateur. Cependant, certains chercheurs reconnaissent « lire » de plus en plus à l'écran, c'est-à-dire qu'ils y effectuent la plus grande part du processus de sélection des documents pertinents : lecture de la table des matières, du résumé et consultation « en diagonale » de l'article (introduction, graphiques, tableaux et images, résultats, conclusion). Ce type de « lecture » est aussi fonction de la taille et de la qualité de la résolution de l'écran, ainsi que du type de présentation du contenu.

Par ailleurs, les analyses statistiques qui commencent à être de plus en plus disponibles sur les données d'usage, permettent de faire d'ores et déjà un premier constat majeur : une grande partie des articles consultés en ligne par les usagers proviennent de titres de périodiques non disponibles à la bibliothèque de l'institution. On peut donc supposer que de nombreux usagers vont se familiariser rapidement avec ce type de lecture plus ou moins restreinte à l'écran (généralement corrélée avec une impression papier dès que l'article présente un intérêt de lecture plus approfondie).



Usages des revues électroniques

Annaïg Mahé

Avantages/inconvénients des revues électroniques

- Avantages : globalement, les revues électroniques constituent un gain d'efficacité substantiel dans une activité de recherche aux contraintes de plus en plus fortes : gain de temps, lecture en avance sur le papier, impression directe, données supplémentaires en ligne, accès à des titres non disponibles en local (ne serait-ce que pour la consultation des tables des matières pour identifier les articles pertinents). Selon les disciplines, les chercheurs apprécient plus ou moins les possibilités multimédias (et selon la qualité de leur matériel informatique ou d'impression).
- Inconvénients : les inconvénients perçus par les chercheurs évoluent en fonction de leur degré d'adoption du nouveau support. Ainsi, les moins familiers seront plus sensibles aux problèmes de pérennité et d'archivage, de difficultés techniques, d'un risque d'isolement et d'invasion de la technique. Les plus assidus feront part d'une perception de perte du butinage lié au papier, d'un renforcement de l'hyperspécialisation, et, globalement, d'un manque d'organisation de ce type de ressources. Ces inconvénients s'apparentent de fait à des exigences de plus en plus fortes face à un outil de plus en plus intégré dans les pratiques quotidiennes. Ces exigences sont à prendre en compte pour l'amélioration de l'offre des revues en ligne, et ce en fonction des besoins spécifiques à chaque communauté scientifique, notamment par l'amélioration des fonctions hypertextes qui, de l'avis de la majorité des chercheurs, ne permettent pas une consultation à la fois aussi libre et structurée que celle d'une collection de bibliothèque.

Références

Abels Eileen G., Liebscher Peter et Denman Daniel W., « Factors that influence the use of electronic networks by science and engineering faculty at small institutions. Part 1 : Queries », *Journal of the American Society for Information Science*, t. 47, n° 2, 1995, p. 146-158.

Baldwin Christine et Pullinger David, « What readers value in academic journals », *Learned Publishing*, n° 13, 2000, p. 229-239.

Brockman William S., Neumann Laura, Palmer Carole L. et Tidline Tonya J., *Scholarly work in the humanities and the evolving information environment*, Digital Library Federation, Council on Library and Information Resources, Washington, D.C., 2001, <http://www.clir.org/pubs/reports/pub104/pub104.pdf> (une étude récente en sciences humaines et sociales dont les résultats ne démentent pas les



tendances observées jusqu'ici).

Eason Ken et Harker Susan, « Psychological processes in the use of electronic journals », *Serials*, t. 13, n° 2, 2000, p. 67-72.

Usages des revues électroniques

Eason Ken, Richardson Sue et Yu Liangzhi, « Patterns of use of electronic journals », *Journal of Documentation*, t. 56, n° 5, 2000, p. 477-504.

Annaïg Mahé

Eason Ken, Yu Liangzhi et Harker Susan, « The use and usefulness of functions in electronic journals : the experience of the SuperJournal Project », *Program*, t. 34, n° 1, 2000, p. 1-28.

Fondation Maison des Sciences de l'Homme, *Les usages et les besoins des documents numériques dans l'enseignement supérieur et la recherche*, Programme numérisation pour l'Enseignement supérieur et la Recherche, étude dirigée par Jean-Michel Salaün et Alain Van Cuyck, 1999, <http://www.pner.org>

Kling Rob et McKim Geoffrey, « Not just a matter of time : field differences and the shaping of electronic media in supporting scientific communication », *Journal of the American Society for Information Science*, t. 51, n° 14, 1999, <http://xxx.lanl.gov/ftp/cs/papers/9909/9909008.pdf>

Mahé Annaïg, Andrys Christine et Chartron Ghislaine, « How French researchers are making use of electronic journals : a case study conducted at the Pierre et Marie Curie and Denis Diderot Universities », *Journal of Information Science*, t. 26, n° 5, 2000, p. 291-302.

Mahé Annaïg, thèse de doctorat en cours sur les usages des revues électroniques par les chercheurs, Université Claude Bernard Lyon 1 (soutenance 2002).

Mahé Annaïg, « L'intégration des revues électroniques dans les pratiques : un processus d'appropriation observé auprès de chercheurs du Commissariat à l'Energie Atomique », in Chartron, Ghislaine (dir), *La mutation de l'édition de recherche, Internet et les chercheurs*, Cercle de la Librairie, Paris, à paraître, printemps 2002.

Tenopir Carol et King Donald W., *Towards electronic journals. Realities for scientists, librarians and publishers*, Washington, Special Library Association, 2000.

Pullinger David, « Academics and the new information environments : the impact of local factors on use of electronic journals », *Journal of Information Science*, t. 25, n° 2, 1999, p. 164-172.



Evaluation des usages

Les méthodes quantitatives

Usages des revues électroniques

Intérêts et limites

Pour une évaluation quantitative des usages, deux **types de données** peuvent être utilisées :

Annaïg Mahé

- les **indicateurs internes** classiques recueillis directement par les différents services de la bibliothèque/documentation par des mesures de comptage ou par questionnaires : il est ainsi possible de comparer l'évolution de l'usage des services de revues électroniques avec celle de la fréquentation de la bibliothèque, de la fourniture de documents, et des autres services.
- les **données de consultation en ligne** recueillies auprès des services d'accès aux revues électroniques, soit directement sur les serveurs en local, soit indirectement sur les serveurs des éditeurs.

Ce deuxième type de données étant encore très récent et leur usage mal défini, nous proposons ici une synthèse des avancées dans ce domaine, basée sur une traduction synthétisée des guides d'évaluation et des travaux réalisés ou en cours.

Un intérêt partagé

Jusqu'à présent, les bibliothèques ne disposaient que de moyens très empiriques et limités pour connaître le degré de consultation de leurs collections. La consultation des revues en ligne va bouleverser les possibilités d'analyse de ce type de consultation grâce au degré de précision bien supérieur des données recueillies, toutes les données relatives à la consultation en ligne étant enregistrées dans des fichiers spécifiques.

Si, au départ, les éditeurs pouvaient craindre l'utilisation de ces données par les bibliothèques pour annuler les abonnements à des titres à faible consultation, les différents acteurs sont actuellement parfaitement conscients de l'intérêt partagé de ce type de données. Les bibliothèques sont devenues dépendantes des fournisseurs pour obtenir des données d'usage qui ne leur appartiennent pas et dont elles ont besoin pour justifier leurs investissements. Pour les consortiums, notamment, ces données sont primordiales pour connaître les types de consultation pour les membres individuels, pour tout le consortium, ou encore pour réaliser des comparaisons d'usage entre les institutions. De façon générale, cela leur donne une meilleure base de négociation. Pour les éditeurs, elles leur permettent de mieux mesurer la santé de leurs collections. Il s'agit surtout d'un nouveau rôle pour les éditeurs et les agrégateurs, et qui leur demande un temps d'apprentissage ainsi qu'un investissement financier. Ceci



Usages des revues électroniques

Annaïg Mahé

explique qu'actuellement encore moins de la moitié des éditeurs soient en mesure de proposer ce type de données. Un certain nombre d'intermédiaires proposent aussi des données de consultation de leur service, selon des formats différents (pour en citer quelques uns : OCLC, HighWire Press, Ebsco, RoweCom, Swets-Blackwell).

Les premières analyses basées sur ce type de données indiquent toutes l'usage d'une collection de titres bien plus large que celles offertes en local par les bibliothèques. Ces analyses ont donc un impact important sur la sélection et la gestion des collections : en permettant une couverture la plus large possible, les usagers ont une latitude de choix plus importante, ce qui permet aux responsables de mieux cerner les besoins. L'analyse des données statistiques donne des indications précises sur les usages et permet de les comparer avec le discours tenu par les usagers (une étude a ainsi constaté un usage relativement élevé du format HTML tandis que les usagers déclaraient une forte préférence pour le format PDF), ou encore de comparer les usages du papier et des ressources en ligne. Il est aussi possible de mieux définir les pratiques selon les domaines, notamment pour les différences de besoins en antériorité.

Des possibilités d'analyse encore incomplètes

Les possibilités d'interprétation de ce type de données restent cependant encore très incomplètes : les bibliothèques ont des demandes différentes et il n'y a pas d'homogénéité dans le type et le format des données fournies. Dans le cadre de travaux sur les indicateurs de performance des bibliothèques, un certain nombre d'initiatives travaillent à la mise en place de standards qui permettraient d'exploiter la richesse des renseignements apportés par ce type de données. L'éditeur Elsevier aborde ces questions dès 1996 dans le cadre de son projet Tulip, et une des premières initiatives de standardisation provient de l'ICOLC qui, dès 1997, propose une liste, récemment mise à jour, d'éléments de base à prendre en compte. Les consortiums membres de l'ICOLC ont, en effet, la responsabilité de fournir des données d'usage aux bibliothèques membres, et les fournisseurs sont encouragés à aller au-delà des données minimum requises (cette liste s'adresse cependant plus aux données concernant l'usage des bases de données de façon générale qu'à celles concernant l'usage des périodiques en ligne en particulier).

Selon l'étude de Judy Luther, il est nécessaire de prendre en compte, en premier lieu, le contexte d'interprétation des données pour éviter de tirer des conclusions trop hâtives. La revue reste l'unité de vente et baser le prix sur l'usage peut fonctionner pour les titres les plus populaires, mais cela ne prend pas en compte l'importance de l'impact sur la recherche de titres moins utilisés. Ainsi, le niveau d'activité (les téléchargements de fichiers, par exemple) n'indique pas la valeur d'un article et d'autres facteurs sont à prendre en compte dans l'analyse :



Usages des revues électroniques

Annaïg Mahé

- le temps d'intégration du service dans les pratiques (évalué à une période de 16 mois à 3 ans) ;
- la publicité et la promotion du service qui influe aussi sur les taux d'usage ;
- le type de contenu (présence d'archives ou non) ;
- les barrières à l'usage qui ont aussi leur importance : nécessité d'une inscription, prix, type d'interface, etc.

Il existe, par ailleurs, de nombreux biais relatifs à la collecte des données de base. Les résultats obtenus de ces évaluations quantitatives restent à prendre comme des indicateurs de tendances plus ou moins prononcées, et qui permettent de nombreuses analyses comparatives.

Traitement des données brutes

Le matériau de base pour la production de rapports d'usage des produits en ligne est le **fichier log** du serveur web. Il existe des logiciels qui permettent d'analyser les données contenues dans ce type de fichier.

Le fichier log contient les informations suivantes :

- **Qui** : qui a demandé une URL (adresse IP de la machine de l'utilisateur, le nom d'utilisateur (quand il existe un login personnel) et le cookie)
- **Quoi** : quelle URL a été demandée (URL de la page demandée)
- **Quand** : à quel moment l'URL a été demandée (tampon de la date et de l'heure, et zone de temps correspondant à l'utilisateur en fonction de sa localisation)
- **Comment** : comment la demande a été faite (type et version du navigateur utilisé, type du système d'exploitation)
- **Informations supplémentaires** (le développement des navigateurs permettra par la suite à d'autres type d'éléments d'être enregistrés dans le fichier log) :
 - code de retour (indique le statut de la demande : accomplie, manquée, refusée, etc.)
 - nombre total de bits transférés par demande
 - référent ou URL précédente (d'où la demande ou le clic a été fait)

Cependant, les données enregistrées dans ce type de fichiers ne peuvent être traitées telles qu'elles : près de 50% des informations du fichier sont à exclure et il est nécessaire d'effectuer un premier traitement consistant à éliminer les données multiples. Pour cela, on applique des **filtres** permettant d'exclure toutes les autres données contenues dans le fichier log et qui ne concernent pas les demandes des utilisateurs :



Usages des revues électroniques

Annaïg Mahé

- un filtre de base permet d'abord d'exclure les enregistrements concernant les images ou objets contenus dans la page demandée, les enregistrements contenant un code autre que « 200 » (demande accomplie), « 301 » (demande redirigée) ou « 304 » (utilisation d'une copie cache), et de ne conserver ainsi que les demandes intentionnelles et accomplies.
- d'autres filtres permettent :
 - d'exclure les enregistrements multiples générés lors de la consultation de fichiers PDF (avec Microsoft Internet Explorer versions 4.x et 5.x). Ces enregistrements multiples sont aussi possibles avec d'autres navigateurs et il est difficile de suivre les changements du marché ;
 - d'exclure les demandes accomplies mais non-intentionnelles des usagers : double-clic au lieu d'un simple clic, bouton « retour » du navigateur, bouton de rechargement de la page

Certains sites interdisent l'accès aux moteurs de recherche et aspirateurs de sites, mais lorsque ce n'est pas le cas les statistiques peuvent être artificiellement gonflées. Par ailleurs, la multiplicité des moyens d'accès à un titre rend difficile une analyse globale : ainsi, les accès obtenus par l'intermédiaire d'une copie cache ne sont pas comptabilisés dans le fichier log et cela peut donc fausser l'interprétation de l'usage des documents souvent consultés.

Les informations sur les demandes qui échouent peuvent être conservées à part. Cela permet notamment d'analyser les tentatives d'accès à des titres non-souscrits.

Définition des données à analyser

A partir de ce premier traitement des données brutes, il est nécessaire de définir le type d'éléments que l'on cherche à analyser. Trois types de mesures sont à prendre en compte et donnent des **informations sur l'activité et l'identité des utilisateurs** : les « hits » (chaque action réalisée sur le site), les sessions, les téléchargements. Selon les éditeurs, les données fournies sur ces éléments diffèrent et sont plus ou moins précises, ce qui complique l'analyse ultérieure.

Les données minimum fournies donnent des indications sur le nombre total, et par titre, de pages visualisées (tables des matières, résumés, et articles en texte intégral, type de format), ces données pouvant se décliner par année et par mois, et éventuellement par volume et par numéro pour chaque titre, ainsi que par groupe d'adresses IP.

La liste ci-dessous présente un certain nombre des éléments qui peuvent faire l'objet d'une analyse quantitative :



Usages des revues électroniques

Annaïg Mahé

Contenu, éléments utilisés

- nombre total d'actions (« hits » ou clics) réalisées sur le site ;
- nombre total de pages visualisées par type de page (page de titre de périodiques, table des matières, résumé, article en texte intégral, review, letter, page de recherche) ;
- type de format des pages en texte intégral (PDF, HTML) ;
- nombre d'unités en texte intégral visualisées, téléchargées ou fournies d'une autre manière (impression ; envoi par messagerie à partir du site, les envois par messagerie réalisés à partir de fichiers téléchargés ne pouvant être comptabilisés) ;
- éventuellement : nombre de requêtes refusées (permet de connaître le nombre de demandes de connexion pour des documents dont l'accès n'est pas permis) ; nombre maximum d'utilisateurs simultanés et autres éléments pertinents pour le modèle de prix appliqué

Pour l'analyse de l'usage par titre, il faut prendre en compte le nombre de documents téléchargés en fonction du nombre total d'articles publiés pour chaque titre. Les articles fortement utilisés donnent des indications sur les thèmes d'actualité. Certains éditeurs fournissent des informations de type « top 10 » des articles visualisés (et indication du format et âge de l'article en nombre de jours). Par ailleurs, comme nous l'avons déjà mentionné, les documents accédés par un cache / proxy (copies souvent demandées et stockées sur une mémoire/un serveur intermédiaires pour un accès plus rapide) ne sont pas comptés dans les statistiques : celles-ci sont donc sous-estimées pour les documents populaires. De même pour les impressions ultérieures de fichiers : il est possible de connaître le nombre de fichiers téléchargés mais une fois ces fichiers sauvegardés sur la machine de l'utilisateur, il n'est plus possible de connaître le nombre de consultations ou d'impressions supplémentaires, ou d'envois de fichiers par messagerie en dehors des envois à partir du site de l'éditeur.

Utilisateurs

- nombre total de sessions (logins) : les éléments d'une session doivent être définis (début, fin, durée ; généralement la session s'arrête lorsqu'un temps défini de connexion est dépassé) ;
- nombre de sessions par adresses IP (ou plus généralement par tranche/classe d'adresses IP) ;
- nombre d'adresses IP uniques qui se sont connectées (si ce type d'informations a été précisé).

Ces éléments sont à décliner selon différents degrés de granularité de l'ensemble de l'institution concernée à l'utilisateur individuel (selon le degré de confidentialité appliqué), ce qui permet éventuellement de comparer les usages en fonction des types d'utilisateurs (étudiants, jeunes chercheurs, chercheurs confirmés, etc.). Si les données ne sont pas connues pour les adresses IP individuelles, les données par tranche/classe d'adresses permettent éventuellement de connaître le département des usagers. Au



Usages des revues électroniques

Annaïg Mahé

sein d'un consortium, cela permet d'obtenir des données pour l'ensemble des institutions, pour chaque institution, pour les différentes unités, etc. Il est, par contre, plus difficile d'obtenir des données pour les usagers distants ayant une adresse IP dynamique ; de même si un proxy est utilisé par l'institution.

A l'analyse, cela permet de confronter les données d'usage aux spécificités des différents sites (et par exemple de constater si les titres consultés correspondent à des titres déjà disponibles sur place sous forme papier ou au contraire à des titres non disponibles). Cela permet aussi de connaître l'origine des usages réguliers, à répétition (le projet Tulip propose aussi une mesure du degré de pénétration : nombre des usagers à répétition divisé par le nombre d'usagers potentiels du service).

Types d'utilisation

- source du document : chemin d'accès suivi, URL de la page précédente (utile pour les accès multiples) ;
- nombre de sélections de menus et accès au texte intégral (feuilletage par titres, numéro, table des matières, thèmes) ;
- nombre de recherches et éventuellement types de recherches : par titre, auteur, sujet (les modalités de recherches sont complexes à collecter)

Comparer les statistiques des résumés et des tables des matières avec celles des téléchargements donne des indications sur la façon dont la collection est utilisée.

Période d'utilisation

- les mesures d'activités (hits, sessions, téléchargements) sont résumées en heure, jour, semaine, mois et année

Un certain nombre d'études ont ainsi pu constater la fréquence des usages en dehors des horaires d'ouverture de la bibliothèque.

Fourniture et format des données

Il est important de définir les modalités de fourniture des données a priori, si possible dans la licence d'accès elle-même. Des définitions doivent aussi être fournies pour chaque élément. Il a été suggéré qu'une DTD soit définie pour ce type de rapport.

- formats : les formats utilisés sont généralement de type CSV (format texte délimité), HTML et XML. Le format CSV permet de télécharger les données pour les importer dans un tableur.
- mode de fourniture : selon les préconisations de l'ICOLC, la fourniture par email ne doit pas être obligatoire et les clients doivent pouvoir :



Usages des revues électroniques

Annaïg Mahé

- accéder à des données en temps réel ou mises à jour chaque jour
 - spécifier la période couverte
 - accéder deux années de données en ligne pour mesurer la croissance
 - visualiser les données par graphiques par année, titre ou pour comparer avec d'autres bibliothèques
 - personnaliser le contenu, établir un profil et recevoir les résultats régulièrement par mail
 - importer les fichiers en local
- confidentialité de l'utilisateur : les informations portant sur les utilisateurs peuvent poser des problèmes de confidentialité (par exemple, les adresses de messagerie, mais aussi les thématiques suivies pour certains types de recherche (à ce propos, le degré de contrainte sur la protection des données est plus fort en Europe qu'aux Etats-Unis)
 - confidentialité du consortium ou de l'institution : la confidentialité de l'institution cliente doit être garantie. Dans le cadre d'un consortium, un membre peut demander à ce que les données d'usage le concernant ne soient pas accessibles par les autres membres.
 - garantie de l'accès aux données : l'accès doit être garanti et défini, de préférence dans la licence. Il est aussi nécessaire de définir la responsabilité au sein de l'institution cliente pour l'accès à ce type de données.
 - format des rapports : des résumés doivent être fournis pour un consortium, une bibliothèque individuelle et par titres. Les fournisseurs sont encouragés à fournir des données supplémentaires.

Références

Exemples d'études

Elsevier Science, Tulip Final Report, 1996, <http://www.elsevier.nl/homepage/about/resproj/tulip.shtml>

Guthrie, Kevin M., « Revitalizing older published literature : preliminary lessons from the use of JSTOR », The economics and usage of digital library collections, March 23-24 2000, Ann Arbor, Michigan, <http://www.si.umich.edu/PEAK-2000/guthrie.pdf>

Mackie-Mason, Jeffrey K. ; Riveros, Juan F. ; Bonn, Maria S. ; Lougee, Wendy P.A., « Report on the PEAK experiment. Usage and economic behavior », D-Lib Magazine, July/August 1999, vol. 5, n° 7/8, <http://www.dlib.org/dlib/july99/mackie-mason/07mackie-mason.html>



Usages des revues électroniques

Annaïg Mahé

Rusch-Feja, Diann ; Siebeky, Uta, « Evaluation of usage and acceptance of electronic journals. Results of an electronic survey of Max Planck Society researchers including usage statistics from Elsevier, Springer and Academic Press », D-Lib Magazine, October 1999, vol. 5, n° 10, <http://www.dlib.org/dlib/october99/rusch-feja/10rusch-feja-summary.html>

Sanville, Thomas J., « Use of electronic journals in OhioLINK'S Electronic Journal Center », 67th IFLA Council and General Conference, August 16-25, 2001, <http://ifla.inist.fr/IV/ifla67/papers/118-141e.pdf>

Linda S. Mercer, « Measuring the use and value of electronic journals and books », Issues in Science and Technology Librarianship, Winter 2000, <http://www.library.ucsb.edu/istl/00-winter/article1.html> (présente un exemple des statistiques fournies par titre par HighWire Press)

Guides et initiatives sur les données statistiques

Judy Luther, « White Paper on Electronic Journal Usage Statistics », octobre 2000, <http://www.clir.org/pubs/abstract/pub94abst.html>

Marthyn Borghuis, « What to count and what not ? A white paper on the filters to be applied to a Web-server log file before usage-analysis and reporting can start », September 20, 2000, ScienceDirect, Elsevier, End-user & Library Research

ICOLC Guidelines for statistical measures of usage of web-based information resources, <http://www.library.yale.edu/consortia/2001webstats.htm>

ARL E-metrics project, <http://www.arl.org/stats/newmeas/emetrics/index.html>

NCLIS, National Commission on Libraries and Information Science, Electronic access and use related measures, <http://www.nclis.gov/statsurv/statsurv.html>

PALS : Publishers and Libraries Solutions Group (ALPSP, JISC, PA). Code de pratiques version 1 prévue pour avril 2002, <http://www.usagestats.org/> (voir aussi le Forum on Vendor-Based Usage Statistics, <http://www3.oup.co.uk/central/pals>)

Exemples de logiciels d'analyse des fichiers log

NetTracker, : <http://www.sane.com/products/NetTracker/>

Sawmill : <http://www.sawmill.net/>

FlashStats : <http://www.maximized.com/products/flashstats/>

Analog (gratuit) : <http://www.ftp.asi.fr/analog/>



A propos du questionnaire

De Singly François, *L'enquête et ses méthodes : le questionnaire*, Paris, Editions Nathan, 1992.

Usages des revues électroniques

Annaïg Mahé

Les méthodes qualitatives

Cette partie est basée sur l'expérience du travail de doctorat en cours et les références utilisées.

Intérêts et limites

Les méthodes d'enquêtes qualitatives sont souvent dévalorisées par rapport aux méthodes quantitatives : on leur reproche généralement de se baser sur des échantillons trop faibles pour légitimer les résultats obtenus, mais c'est un reproche principalement basé sur une méconnaissance de ce type de méthodes. Les méthodes qualitatives sont, au contraire, particulièrement bien appropriées pour l'étude des opinions, des comportements et des pratiques des individus, et se justifient donc pleinement pour l'étude des usages. Elles permettent de les comprendre du point de vue de l'utilisateur, de se familiariser avec son environnement, ses pratiques, ses besoins, en prenant en compte les contextes propres à chacun. A ceux qui questionnent la crédibilité de telles méthodes en regard de celles des sciences dites « exactes », Bruno Latour répond :

« Pour combler ce déséquilibre (...) il faudrait environ une centaine d'observateurs pour ce seul terrain, chacun ayant le même pouvoir sur leurs sujets que vous en avez sur vos animaux. En d'autres termes, il aurait fallu enregistrer sur bande vidéo ce qui se passait dans chaque bureau; il aurait fallu dissimuler des micros dans les téléphones et dans les bureaux; nous aurions dû prendre la liberté de faire des électro-encéphalogrammes; et nous nous serions réservé le droit de couper les têtes des participants si un examen interne s'était avéré nécessaire. Si nous avions disposé de ces libertés, nous aurions pu produire des données dures. » (Latour et Woolgar, 1988).

De fait, « les méthodes qualitatives ont pour fonction de comprendre plus que de décrire systématiquement ou de mesurer » (Kaufmann, 1996). Comprendre, c'est-à-dire rendre intelligible, en se donnant les moyens de connaître la complexité subjective des actions d'un individu dans un contexte particulier.

C'est la raison pour laquelle la recherche de la représentativité statistique a peu de sens dans les méthodes qualitatives qui doivent au contraire



Usages des revues électroniques

Annaïg Mahé

privilégier la variété : « *A valeur significative, ce qui est qualitatif, ce qui est spécifique* » (Jo Link-Pezet, 1999). C'est justement ce type de connaissance que l'on cherche à recueillir : l'expérience, la spécificité d'un individu dans un processus collectif, mais c'est aussi la plus implicite, la plus proche de la personne enquêtée et, de ce fait, la plus difficile à exprimer.

Un autre reproche courant, et justifié cette fois, envers ce type de méthodes est le fait que le discours recueilli ne constitue pas une base légitime de connaissances des pratiques : il existe un écart entre le discours d'un individu sur ses pratiques et la réalité des pratiques décrites, et le contexte dans lequel se déroule une enquête qualitative peut avoir un effet majeur sur cette différenciation. Il est possible d'éviter en partie ce biais, inhérent à la pratique de l'entretien classique, individuel ou en groupe (étude de panels, « focus group »), grâce à une certaine maîtrise de ce type d'enquête, mais aussi grâce à l'utilisation en parallèle d'autres méthodes qualitatives, notamment l'observation participante ou l'utilisation de « carnets de route ». En effet, ce qui fait l'intérêt de ces méthodes est aussi une de leurs limites principales : il faut du temps pour s'imprégner d'un contexte tout en recevant un discours qui n'est qu'une partie de ce qu'il décrit. L'utilisation de ces différents types de méthodes permet ainsi une certaine flexibilité, et une grande richesse des résultats. Une enquête qualitative peut soit être réalisée en amont pour préparer l'élaboration d'un questionnaire, ou en aval pour approfondir les résultats d'une enquête quantitative.

Description de différents types de méthodes

L'entretien individuel

L'entretien est un « classique » des enquêtes qualitatives. Il est utilisé sous différentes formes : de l'entretien libre sur un thème large (cf biographies, histoires de vie) à l'entretien semi-structuré qui définit un certain nombre de thèmes que l'on veut aborder tout en laissant une grande liberté de réponse à la personne enquêtée. Pour l'étude des usages des revues électroniques, c'est ce type d'entretien qui s'applique le mieux. En voici quelques éléments d'application :

- un guide d'entretien préparé au préalable définit les thèmes principaux à aborder, et prévoit des relances-types qui peuvent aider à approfondir certaines questions au cours de l'entretien. Le thème principal peut rester suffisamment large pour ne pas biaiser le contenu du discours et donne un cadre dans lequel l'entretien reste libre. Par exemple, pour comprendre comment la pratique des revues électroniques s'intègre dans le contexte plus large d'une activité d'information, le thème spécifique des revues électroniques n'est jamais mentionné avant que le chercheur ne l'ait abordé lui-même de façon explicite. Cela permet :



Usages des revues électroniques

Annaïg Mahé

- de voir comment le thème apparaît, s'il est mentionné immédiatement ou s'il apparaît en position périphérique,
 - de voir les autres termes qui s'y rapportent,
 - de saisir des attitudes et des représentations plus générales ;(Ghiglione et Matalon, 1985.
- un échantillon de quelques personnes permet déjà de dégager des thèmes particuliers ou récurrents. Un échantillon de 20 à 30 personnes constitue une bonne base pour connaître les usages d'un service. Comme nous l'avons déjà mentionné, il n'est pas utile de constituer un échantillon selon des normes de représentativité statistique. Si l'on cherche à connaître l'intégration d'un service auprès d'une population, il faut s'assurer d'intégrer des usagers comme des non-usagers dans l'échantillon. Si l'on cherche à étudier en détail les usages d'un service, il est plus intéressant de se concentrer uniquement sur des usagers de ce service ;
 - l'entretien se déroule de préférence sur le lieu de travail de l'enquêté. Cela permet ainsi de replacer le discours sur les pratiques dans leur contexte grâce à l'observation *in situ* de l'environnement quotidien de travail : organisation spatiale des locaux, ambiance de travail, présence et organisation de la documentation personnelle, proximité de la bibliothèque, etc ;
 - l'entretien doit avoir une durée suffisante pour faire le tour des questions à aborder et permettre à l'enquêté de s'exprimer librement. Selon la disponibilité des personnes, il peut durer de une à trois heures. Lorsque c'est possible, il est intéressant de réaliser des campagnes d'entretiens à plusieurs semaines ou mois d'intervalle. Cela permet notamment d'approfondir certaines questions, mais aussi de constater l'évolution des pratiques ;
 - il est fréquent, au cours de l'entretien, que la personne enquêtée soit amenée à donner une image de ses pratiques conformes aux attentes supposées de l'enquêteur. Il est donc nécessaire d'instaurer une confiance réciproque lors de l'entretien pour permettre l'expression d'une parole la plus sincère et la moins formelle possible ;
 - idéalement les entretiens sont enregistrés, si la personne accepte et sous garantie d'anonymat. Cependant, l'analyse des cassettes étant très preneuse de temps (et plus appropriée à une analyse de contenu), il est préférable de prendre des notes extensives au cours même de l'entretien (cela a en outre l'avantage de donner du temps de réflexion à la personne enquêtée) et qui serviront de base à l'analyse, les cassettes servant alors de garantie et permettant la vérification des notes, si besoin.

L'étude de panels

Les études de panels (« focus group ») sont des entretiens en groupe sur un thème particulier. Elles sont intéressantes pour compléter des entretiens individuels, ou lorsque la durée des enquêtes est restreinte, en permettant



Usages des revues électroniques

Annaïg Mahé

d'interroger plusieurs personnes simultanément. Elles sont aussi utiles pour confronter des personnes provenant de contextes variés, ce qui stimule la discussion. Ce type d'enquête demande néanmoins une préparation préalable assez importante et une bonne gestion de son déroulement.

L'observation participante

Cette méthode consiste à passer un temps plus ou moins prolongé et/ou répété sur le lieu de travail des personnes enquêtées. Elle a pour avantage principal de pouvoir observer et questionner directement, en temps réel, les pratiques, et constitue ainsi une intéressante contrepartie aux entretiens. Sa limite principale, en revanche, est la disponibilité requise des participants.

Le « carnet de route »

Le carnet de route est un substitut intéressant à l'observation participante. Un carnet préparé est remis à chaque personne qui doit elle-même remplir les différentes rubriques pendant une durée déterminée. Elle requiert, cependant, une rigueur assez forte de la part des personnes enquêtées.

L'analyse des données

Une analyse approfondie des entretiens permet de recenser les thèmes communs, suggérant des tendances larges, mais aussi de voir plus en détail les façons de faire et les particularités de chacun. Cela permet de dégager des variables pour la construction d'une typologie dans laquelle les profils individuels peuvent être replacés de manière intelligible, ce que ne permet pas la grande diversité des données de départ.

Sans aller jusqu'à une analyse approfondie des données ainsi recueillies, la pratique ponctuelle de ce type d'approches permet l'entretien de **relations de proximité** entre les responsables d'un service de documentation, et leur « clientèle » : effectuer des entretiens, se rendre sur les lieux de travail, sont des occasions privilégiées d'expression et d'échange dont chacun ressort bénéficiaire.

Références

Barry, Christine A., Critical issues in evaluating the impact of IT on information activity in academic research : developing a qualitative research solution, *Library & Information Science Research* , vol. 17, 1995, 107-134.

Cacaly, Serge, (dir.), *Dictionnaire encyclopédique de l'information et de la documentation* , Nathan, Paris, 1997.

Ghiglione, Rodolphe ; Matalon, Benjamin, *Les enquêtes sociologiques. Théories et pratiques* , Armand Colin, Paris, 1985.



Kaufmann, Jean-Claude, *L'entretien compréhensif*, Editions Nathan, Paris, 1996.

Latour, Bruno ; Woolgar, Steve, *La vie de laboratoire. La production des faits scientifiques*, La Découverte, Paris, 1988.

Usages des revues électroniques

Le Coadic, Yves F., *Le besoin d'information : formulation, négociation, diagnostic*, ADBS Editions, Paris, 1998.

Annaïg Mahé

Link-Pezet, Jo, De la représentation à la coopération : évolution des approches théoriques du traitement de l'information, *Solaris*, vol. 5, 1999, <http://www.info.unicaen.fr/bnum/jelec/Solaris/d05/5link-pezet.html>

Mahé Annaïg, thèse de doctorat en cours sur les usages des revues électroniques par les chercheurs, Université Claude Bernard Lyon 1 (soutenance 2002).

Mahé Annaïg, « L'intégration des revues électroniques dans les pratiques : un processus d'appropriation observé auprès de chercheurs du Commissariat à l'Energie Atomique », in Chartron, Ghislaine (dir), *La mutation de l'édition de recherche, Internet et les chercheurs*, Cercle de la Librairie, Paris, à paraître, printemps 2002.