



World Library and Information Congress: 69th IFLA General Conference and Council

1-9 August 2003, Berlin

Code Number: 015-F
Meeting: 127. National Libraries & Statistics and Evaluation
Simultaneous Interpretation: -

Benchmarking: vue d'ensemble et contexte

Rowena Cullen

School of Information Management
Victoria University of Wellington
Wellington
New Zealand

« le management, c'est à la fois le changement et le maintien d'un taux élevé de changement »
Sir Harvey Jones

Résumé :

Cette communication passe en revue quelques points relatifs au benchmarking et à la réussite de sa mise en œuvre dans la gestion de la qualité dans les bibliothèques publiques, d'enseignement supérieur, et spécialisées. Elle traite du recours nécessaire à des processus de benchmarking et de revue par les pairs pour calculer la mesure et l'analyse des résultats, ainsi que les données plus classiques telles que les entrées, les processus et la production, tout en examinant l'importance du contexte social. Elle met en avant la pertinence du benchmarking pour les bibliothèques nationales, et certaines méthodes de mise en œuvre en vue d'améliorer la qualité. Ce texte tente d'expliquer comment prendre en compte la spécificité des bibliothèques nationales et la difficulté qu'elles ont à trouver des organismes avec lesquelles pratiquer le benchmarking. Il présente quelques éléments de « meilleure pratique » de benchmarking dans ce secteur.

Introduction

A l'époque d'intense changement où nous vivons, où il est parfois difficile de concilier l'obligation d'affronter ce changement et la réalisation de certains de nos objectifs permanents sur le long terme – pour offrir le meilleur niveau de service au plus grand nombre d'utilisateurs et aux décideurs. Pour y parvenir, nous devons identifier les changements extérieurs et ceux qui sont prévisibles en interne, et gérer le processus de changement afin de concilier les deux. Les changements extérieurs auxquels nous devons réagir comprennent les technologies de l'information et des communications (TIC) qui affectent tous les secteurs de notre activité, les

évolutions dans les attentes des usagers et des clients, les exigences toujours changeantes en matière de gestion, de fourniture de services et de coopération nationale et internationale.

Dans de nombreux types de bibliothèques, et dans certains secteurs des bibliothèques nationales, le benchmarking a été perçu comme un moyen de répondre à la gestion externe, et de stimuler le changement en interne. Dans ce document, j'examine quelques points concernant le benchmarking et ses applications réussies dont je donne quelques exemples, avant d'aborder la situation des bibliothèques nationales, en explorant la pertinence et la valeur de l'évaluation du benchmarking dans ce secteur. Enfin, puisque le benchmarking est souvent utilisé pour définir ce que l'on appelle les « bonnes pratiques », j'indique en quoi consistent les « bonnes pratiques » de benchmarking dans ce secteur.

Les débats concernant le benchmarking

La définition de performance donnée par Foot (1998), auprès d'un auditoire de responsables locaux et de bibliothécaires, à savoir « un point fixe par rapport auquel vous pouvez vous mesurer, comme un objectif local ou un norme », est plutôt restrictive. Il est vrai que de nombreuses applications de benchmarking en bibliothèque se sont dans le passé appuyées sur une norme statistique, soit à partir d'un groupe de bibliothèques comparables, soit à partir d'un tableau de bord national, se référant à une série de mesures d'entrées (ressources financières ou effectifs), de mesures de processus (efficacité) ou de mesures de résultat (tels que les prêts ou les communications), et les ont utilisés pour faire des comparaisons (Pritchard, 1995). Ceci correspond encore à l'approche prise au Royaume Uni par LISU [Unité statistique pour les bibliothèques et les sciences de l'information à l'Université de Loughborough]¹, par l'IPF [Institut de finances publiques]² et par SCONUL [Association des bibliothécaires des bibliothèques nationales et d'enseignement supérieur]³. Mais dans l'industrie et dans le secteur privé, où le benchmarking a pris naissance et où il a encore de fervents supporters, le benchmarking est beaucoup plus orienté vers les procédures commerciales, et moins vers la production ou les activités de services (Camp, 1998). Le site Internet du Benchmarking Network [Réseau de l'Évaluation des performances] fournit des exemples nombreux et variés de benchmarking ; sur les dix processus les plus couramment soumis à l'évaluation dans les entreprises, « la formation continue des salariés », « les technologies des systèmes d'information », « les ressources humaines » et « le développement des mesures de performance » arrivent en tête⁴. Bien que ces activités fassent aussi l'objet d'évaluation dans le monde des bibliothèques, cela est beaucoup moins fréquent que la comparaison des statistiques d'entrées et de sorties de documents, et beaucoup moins utilisé. Ceci contraste avec l'intérêt suscité dans le secteur privé par le développement et la mise en œuvre du benchmarking.

La deuxième définition du benchmarking donnée par Foot est plus exploitable : « processus permettant de mesurer les procédures et la performance d'un secteur en les comparant systématiquement aux résultats d'autres secteurs afin de rechercher la meilleure pratique ». Cette dernière définition met l'accent sur des notions utiles, telles que la comparaison systématique, les

¹ <http://www.lboro.ac.uk/departements/dils/lisu/lisuhp.html>

² <http://www.ipf.co.uk/benchmarking>

³ <http://www.sconul.ac.uk>

⁴ <http://www.benchnet.com>

procédures, la performance et la recherche de la meilleure pratique. Elle se rapproche de l'usage du benchmarking dans le secteur privé, pratique encore toute nouvelle dans les bibliothèques.

C'est pourquoi, le benchmarking, tel qu'on l'utilise dans les bibliothèques, doit être considéré comme mettant l'accent sur :

- i) la collecte et la comparaison de données – en premier lieu les entrées et les sorties. Ceci comprend les statistiques de bibliothèques nationales, universitaires et publiques, la comparaison budgétaire, du nombre d'acquisitions et de leur coût, du nombre de personnel qualifié, du nombre de recherches documentaires et, plus récemment, des services électroniques ;
- ii) l'évaluation des procédures, qui concerne l'analyse des fonctions et des opérations entre une ou plusieurs institutions comparables, avec pour objectif le gain qu'en retireront les participants et les leçons qu'ils tireront les uns des autres.

La première catégorie d'évaluation s'apparente à un outil de diagnostic, même si l'évaluation des procédures peut aussi servir dans ce cas. Cependant, seule l'approche par évaluation des procédures est valable pour identifier la meilleure pratique. Tandis que l'évaluation des procédures est généralement conduite par les pairs, avec une ou plusieurs institutions ou départements, elle peut également être réalisée par des consultants indépendants qui apportent la connaissance d'autres pratiques professionnelles, avec lesquelles comparer, ou à partir desquelles tirer des exemples de meilleures pratiques.

La collecte et la comparaison de données de plusieurs bibliothèques différentes posent le problème de la définition des données, question qui est documentée et bien connue des membres de la Section Statistiques de l'IFLA. L'attention portée à cette question est un des points forts de LISU, et constitue un objet d'attention pour l'Institut de finances publiques, qui pilote le benchmarking pour les bibliothèques des collectivités territoriales au Royaume Uni. Alors que la collecte des données s'oriente désormais des entrées vers les processus et les résultats, ces questions deviennent encore plus complexes. Les différences de calcul des dépenses et des frais généraux s'amplifient lorsque l'on compare les ratios entre ces données, de même que s'accroissent les différences entre les services et les résultats. Cependant, la qualité du diagnostic portant sur la comparaison de statistiques brutes d'entrées, de sorties et de processus, et de ratios entre elles, se limite au signalement des points qui méritent plus ample investigation ; les données ne révèlent rien en soi sur les causes. On ne saurait tirer de conclusions sur les relations entre les entrées, les processus et les résultats sans examiner la validité des mesures utilisées, ni se baser sur elles pour en tirer des hypothèses.

Ceci a été mis en avant par Pratt et Altman (1997) qui ont démontré, à partir d'une simple analyse des statistiques des 24 plus grandes bibliothèques des Etats-Unis, dans le rapport annuel des bibliothèques publiques américaines de 1996, que d'une part il n'y avait aucune corrélation entre les entrées et les sorties, ou entre les dépenses et les résultats, comme démontré par les 5 mesures de résultat utilisées (nombre d'emprunts par personne, total des prêts, recherches bibliographiques, taux de renouvellement et nombre de documents en stock), mais que d'autre part les données de nombreuses rubriques avaient été calculées sur des bases différentes, aussi bien pour les entrées que pour les sorties. En outre, même si on avait éliminé ces différences dans la définition et la collecte des données, les chiffres ne montreraient pas toujours si des différences visibles de résultat tiendraient à une utilisation plus ou moins efficace des ressources, ou à des différences structurelles, culturelles ou stratégiques au sein des diverses organisations, de leur environnement ou de leur contexte. L'évaluation des processus est

indispensable pour fournir une analyse plus approfondie de la question, si vraiment c'est perçu comme un problème.

Même si les risques de telles comparaisons sont connus depuis un certain temps, le Royaume Uni a pris un ensemble très approchant de mesures de résultat pour définir les « meilleures pratiques » dans les BVPIs [Indicateurs de Performance de la Meilleure Valeur] appliquées aux bibliothèques publiques. « A partir d'avril 2000, l'obligation de « meilleure valeur » impose aux collectivités territoriales d'offrir des services respectant des normes claires, portant à la fois sur le coût et sur la qualité, par les moyens les plus économiques, les plus efficaces et les plus efficaces dont elles disposent » (Favret 2000). En ce qui concerne les bibliothèques, en plus d'un grand nombre d'indicateurs de performance figurant dans le Annual library Plan, les BVPIs doivent mesurer :

- le coût d'une visite à une bibliothèque publique ;
- le nombre de visites par tête d'habitant ;
- le pourcentage d'utilisateurs de la bibliothèque ayant trouvé le document souhaité ;
- le nombre de documents fournis par tête d'habitant.

Favret poursuit : « le gouvernement fixera des objectifs au quartile supérieur pour certains indicateurs, pour que, d'une façon générale, les coûts se trouvent dans le quartile le plus bas, tandis que les résultats de la prestation de service se trouvent dans le quartile supérieur ».

D'une part il faut avoir une grande confiance dans l'intégrité, dans la validité et la valeur comparative des données utilisées pour ces indicateurs, d'autre part, il faut noter que ceux-ci comportent également deux problèmes qui leur sont inhérents. Le premier est qu'il est en fait extrêmement difficile de mesurer avec un quelconque degré de précision, le « pourcentage d'utilisateurs de la bibliothèque ayant trouvé le document souhaité », pourcentage qui devrait en principe refléter l'objectif de base que la bibliothèque cherche à remplir ; ce en dépit de l'indicateur en apparence transparent. Le deuxième problème tient au fait que le processus lui-même attribue à la majorité des bibliothèques une place inacceptable dans les 3 quartiles les plus bas (ce qu'on a désigné par l'expression « pointer pour mortifier »).- ce qui amène difficilement à la meilleure performance. Un troisième problème, sans doute plus important, tient au fait que les données contextuelles - l'environnement interne et externe - sont totalement absentes de ces indicateurs. Ces derniers ne peuvent donc être utilisés que comme des éléments de diagnostic, mais ne sont pas forcément significatifs pour rendre compte de la performance, de la meilleure pratique ou du meilleur rapport qualité/coût. Ces problématiques ne sont pas nouvelles pour les bibliothécaires au Royaume Uni, qui débattent déjà depuis un certain temps de la pertinence et de la recevabilité des principaux indicateurs de BVPI, et de la liste complète des indicateurs inclus dans le Plan Annuel des Bibliothèques.

En Europe, de récentes initiatives pour intégrer les services des bibliothèques électroniques mettent en avant les mêmes problèmes. Le projet EQUINOX a été mis en place pour développer les données concernant les services électroniques dans les bibliothèques, et pour faciliter la comparaison des résultats, puisque d'énormes efforts ont été faits pour arriver à un accord sur les modes de calculs eux-mêmes, et sur la définition des données et leurs méthodes de collecte. Les indicateurs de performance élaborés couvrent un domaine familier : pourcentage de la population cible atteint, nombre de sessions enregistrées par le système, ressources propres, et nombre de documents visualisés [par rapport à la population cible], ainsi que le coût par session, et le nombre de sessions de formation suivies par les utilisateurs. Geoffrey Ford a comparé de manière fructueuse le projet EQUINOX avec d'autres méthodes d'évaluation des services électroniques

en bibliothèque, tels que le projet e-Lib, le projet MIEL de Brophy et Wynne, le travail de Bertot et de McClure pour l'ARL [Association of Research Libraries], et il écrit, en faisant référence au projet, «en avoir pour son argent » :

Le document final comporte une série d'indicateurs de performance dérivés d'un croisement de 20 institutions ... [ces] indicateurs financiers ne sont pas utiles à l'évaluation ; que ces indicateurs soient bas ou élevés n'est qu'une question d'opinion. .. [ils] ne sont que des ratios de chiffres qui se trouvent disponibles (Ford 2001, 9).

Exemples d'indicateurs de performance, et ce que nous pouvons en apprendre

La littérature professionnelle sur l'utilisation du benchmarking en bibliothèque et dans le secteur de la documentation reflète les nombreuses manières dont il a été utilisé. A l'exclusion du domaine de la santé (Marshall et Buchanan 1995, Robbins et Daniels 2001, Todd-Smith et Markwell 2002), et du domaine des bibliothèques spécialisées (Henczel 2002, Polling 2002), l'évaluation des performances a été particulièrement encouragée dans le secteur des bibliothèques d'enseignement supérieur au Royaume Uni, en premier lieu par la SCONUL [Association des bibliothécaires des bibliothèques nationales et d'enseignement supérieur], qui pilote l'évaluation des performances au sein d'un petit groupe de bibliothèques universitaires, ainsi qu'en Australie, où l'intérêt se manifeste au sein d'une association similaire, le CAUL [Comité des bibliothécaires d'université]. Malgré cela, une étude sur l'utilisation du benchmarking dans l'enseignement supérieur et dans les bibliothèques spécialisées au Royaume Uni a montré qu'à cette date seulement 7,4% des personnes ayant répondu l'utilisaient effectivement comme outil d'évaluation, alors que presque 25% d'entre elles comparaient leur institution à d'autres au moyen d'indicateurs de performance ou de statistiques standards (telles que dépenses, acquisitions, et/ou transactions fournies par étudiant en équivalent temps plein). L'étude a ensuite soumis l'évaluation à l'essai dans 3 institutions. L'essai a élaboré une définition fonctionnelle de l'évaluation des performances « L'objectif du benchmarking est de documenter et de mesurer un processus clé, et ensuite de comparer le résultat avec celui de processus semblables dans d'autres organisations » (Brockman 1997, 345). Brockman et ses collègues ont identifié 5 catégories d'évaluation des performances :

- concurrentielle (comparaison avec des entreprises leaders dans le même domaine) ;
- générique (comparaison de processus d'activités semblables, quelque soit le secteur d'activité) ;
- interne (comparaison d'opérations internes au sein d'une même entreprise) ;
- fonctionnelle (comparaison de fonctions semblables au sein d'une même activité) ;
- cliente (comparaison du degré de satisfaction des attentes du client).

Douze étapes, issues d'un projet américain impliquant une bibliothèque spécialisée dans le domaine scientifique, ont été identifiées (Allen 1993), et corrélient les approches normalisées utilisées par les experts en management de la qualité (Bullivant 1994, 7). Les étapes portent sur l'identification, la définition et la compréhension d'un processus clé pour qu'il fasse l'objet de benchmarking, la recherche d'un partenaire adéquat, l'identification, la définition et la collecte de données dans les deux institutions, l'analyse des données, les recommandations de changement, de mise en place et de contrôle, étapes qui se passent toutes en concertation avec sur des sites de test dans chacune de institutions. L'étape clé, qui consiste à obtenir l'implication du management, se situe en tout début de processus, et insiste fortement sur la consultation du personnel, étape qui se situe bien plus en amont qu'il est d'usage dans le secteur privé, où le management est impliqué dès le début, et conduit en fait le processus d'évaluation

(Brockman 1997, 39-5). Une approche générique du benchmarking a été mise en place, puisque ceci « peut se faire au rythme des besoins organisationnels », et puisque ceci réduit le montant des ressources (surtout le temps et l'effort) qui doivent être investies dans cette activité en une seule fois. L'attention portée à l'engagement du management et à la consultation de la base est la preuve que le benchmarking est lié au changement et à la mise en place de nouveaux processus et que la gestion de le benchmarking comme processus de gestion du changement a des chances de conduire à un changement plus efficace et plus durable, et à une amélioration des résultats.

Les organismes participant au test étaient déjà considérés comme des cas de « meilleure pratique », et le dernier consensus de l'exercice était que « peu de changements étaient nécessaires ». Cette conclusion de type « pas besoin de changement » s'est aussi révélée à l'issue de l'étude de benchmarking conduite par Robbins et Daniels (2001) au concernant la manière dont étaient perçus par les usagers les services de référence dans les bibliothèques du secteur de la santé. Robbins et Daniels ont utilisé une variante de SERVQUAL (SERVPERF) comme enquête de base et outil d'analyse pour leur étude – un exercice de benchmarking des usagers. La question soulevée par ces deux conclusions est : les études ont-elles été faussées par l'absence d'implication du management dans le benchmarking (et dans le processus de gestion du changement) ? ; le modèle SERVQUAL/SERVPERF est-il mal adapté au benchmarking, ou, comme dans l'étude de Brockman, le processus de benchmarking est-il mal compris ou mal conduit ?

La communication, la formation et l'absence de compréhension des procédures d'évaluation par le management ont été identifiés par Brockman et son équipe comme étant les facteurs clés du manque de succès relatif du test. Ils ajoutent que pour que le benchmarking fonctionne, une institution doit d'abord être centré sur le management de la qualité. « Il s'agit d'un outil tourné vers la qualité, qui est donc mieux utilisé là où la culture et les pratiques sont déjà centrées sur les meilleures pratiques » (Brockman 1997, 372). Pour résoudre la question du choix du partenaire, et pour éviter les situations où le partenaire de benchmarking serait à la traîne derrière celui qui a pris l'initiative en matière de meilleure pratique, ils suggèrent plutôt de mettre en place une base de données de partenaires potentiels pour du benchmarking, « en indiquant quelles institutions ont évalué quels processus, avec les résultats de ces évaluations. Ceci permettrait, d'une part l'identification de partenaires, et d'autre part de désigner « les meilleurs parmi les meilleurs ». (Brockman 1997, 357). La définition du benchmarking donnée par Foot, qui concerne des partenaires égaux cherchant à apprendre les uns des autres, semble s'être perdue avec le désir de montrer sa supériorité et avec le besoin de ranger les institutions dans des cases toutes faites par rapport à leurs évaluations passées plutôt que de chercher à améliorer la performance à venir.

Une base de données des bibliothèques qui pratiquent le benchmarking existe déjà, bien sûr, dans les bibliothèques publiques du Royaume Uni qui appartiennent au Club de l'Évaluation IPF (dont sont actuellement membres plus de 100 bibliothèques) - bien que le benchmarking paraisse centré avant tout sur les indicateurs de performance du BVPI et de ceux du Rapport annuel des Bibliothèques. Malgré l'accent mis sur la consultation avec des bibliothèques pour définir et normaliser les données collectées, l'approche du CIPFA/meilleure valeur pourrait jouer contre les bénéfices réels du benchmarking de processus, ou de l'évaluation par les pairs, comme relevé dans l'industrie (Camp 1998), où l'accent, comme nous l'avons vu, est mis sur l'amélioration des processus commerciaux et non sur les mesures d'entrées/sorties.

Une étude d'évaluation très réussie sur la fourniture de documents à distance a été conduite par le groupe de travail sur le Partage des Ressources nationales (2001) et par la Bibliothèque

nationale d'Australie, à partir des Opérations de Mesure de la performance du Prêt entre bibliothèques de ARL, également très réussies. L'étude australienne a utilisé des données statistiques sur les coûts unitaires, sur le taux de réponse (nombre de documents en mesure d'être prêtés), et sur le taux de rotation pour les documents prêtés et empruntés dans 97 bibliothèques d'Australie. D'énormes écarts de coût et de performance sont apparus entre les pourcentages les plus hauts et les plus bas, et entre les différents secteurs de bibliothèques. Là aussi, comme dans les statistiques des bibliothèques publiques américaines, aucune bibliothèque ne figurait dans les 10 pour cent supérieurs pour l'ensemble des indicateurs utilisés. Malgré l'apparente supériorité de certaines institutions, des améliorations apparaissent possibles dans toutes les institutions, ainsi que des leçons à tirer, y compris d'institutions à la réussite moindre. L'étude a été suffisamment étendue et en même temps suffisamment rigoureuse pour permettre l'investigation derrière ces processus, et la comparaison entre institutions. Des conclusions bien argumentées et bien perçues abordent un certain nombre de points qui ont paru décisifs pour les meilleures pratiques, tels que l'informatisation, les pratiques de catalogage, la formation, et les accords de coopération entre bibliothèques. L'étude a utilisé un outil bien rodé, avec des questions d'une profondeur suffisante pour tester diverses pratiques opérationnelles et managériales, et a été mise en place en tenant compte de l'environnement et de la culture des différentes bibliothèques. – facteurs qui ont conduit, dit-on, du moins de manière anecdotique, à des résultats décevants dans d'autres études d'évaluation dans la région. D'où la leçon qu'il faut peut-être tirer, que l'identification des meilleures pratiques par l'interrogation d'ensembles statistiques, requiert un échantillon suffisamment grand, et une validité de construction solide, pour que les données fassent l'objet d'un examen rigoureux de causalité. Il est peu probable que l'on puisse identifier la meilleure pratique en comparant les entrées et les résultats de deux ou trois institutions seulement. Des institutions désireuses de pratiquer le benchmarking avec un nombre réduit d'autres partenaires devraient se concentrer sur la « méthode la plus souple du spectre » et choisir un processus d'évaluation par les pairs pour lui permettre d'en savoir plus sur ses propres valeurs et ses processus. L'évaluation par les pairs est l'approche suivie par les sociétés qui réussissent le mieux dans le secteur concurrentiel, et qui font état d'une satisfaction générale concernant ce processus (Camp 1998).

Résultats du benchmarking

L'Association des bibliothèques de recherche (ARL) qui a contribué à l'élaboration du projet australien très réussi de benchmarking a déplacé son centre d'intérêt sur la qualité et sur l'évaluation des bibliothèques vers le projet LIBQUAL +tm. S'appuyant sur la méthodologie de SERVQUAL, utilisée pendant deux décennies dans les études de management et de marketing pour étudier la qualité des services (Zeithaml, Parasuraman et Berry 1990), la méthodologie de LIBQUAL+tm découle d'un projet de recherche et développement qui a débuté en 2000 entre ARL et l'Université A et M du Texas, pour « mesurer la qualité de service minimum, perçue ou attendue, des usagers de bibliothèques, dans 4 secteurs clés : Accès à l'information, Atteinte de service, Bibliothèque comme lieu, et **Contrôle personnel (personal control)**⁶. Le projet, lancé par 12 institutions en 2000, s'est étendu à 164 en 2002. Bien qu'au départ conçue comme outil de mesure de la qualité de service d'une bibliothèque individuelle, et comme outil de diagnostic des secteurs présentant des difficultés, la base de données est peu à peu devenue une base nationale par rapport à laquelle les institutions prises individuellement peuvent mesurer leur performance, sous la forme d'un benchmarking client.

Contrairement à l'étude de benchmarking australiennes sur le PEB, LIBQUAL+tm a été critiquée (Hernon et Nitecki 2001, Cullen 2001) au motif qu'elle prend peu en compte les

⁶ voir le site ARL à : <http://www.arl.org/pr/libqual>

différences entre bibliothèques dans des environnements pourtant très hétérogènes, et des situations administratives diverses à travers les Etats-Unis. Il faut reconnaître que des variantes de la méthode SERVQUAL ont été utilisées dans différents contextes de bibliothèques avec succès (Calvert 2001, Hernon et Calvert 1996). L'étude de Calvert (2001) montre que les éléments de qualité de service, tels que le comportement du personnel, l'environnement de la bibliothèque et les services qui permettent à l'utilisateur de trouver et d'utiliser efficacement les documents de la bibliothèque, sont partagés par les étudiants chinois et néo-zélandais. Toutefois, des essais pour tester la méthodologie LIBQUAL+tm au Royaume Uni sont en cours, subventionnées par SCONUL, avec l'idée que l'outil d'enquête très bien testé LIBQUAL+tm et les méthodes d'analyse des données pourront s'appliquer à des institutions au Royaume Uni comme ils l'ont été aux Etats-Unis. Ceci est probablement vrai, mais il reste à voir si les données pourront être versées dans une base de données commune, ou plutôt si celles provenant du Royaume Uni présenteront des schémas de réponse différents, basés sur des différences culturelles, et sur des points comme les variations dans les systèmes de notation entre les universités américaines et britanniques, qui ont un impact sur la manière dont les usagers des bibliothèques (étudiants et enseignants) « notent » les questionnaires d'enquête. Si la réactivité au contexte joue un rôle dans le benchmarking, cela devra être confirmé, nous l'espérons, par ces nouvelles études de LIBQUAL+tm au Royaume Uni.. Ainsi, nous pouvons nous attendre à voir que, tandis que la composante qualité de service constitue un socle robuste relativement cohérent entre divers pays et cultures, l'utilisation des ensembles de données d'où découle cette composante à des fins de benchmarking ne fait peut-être pas un bon usage de ces données.

Un des points les plus intéressants de la méthodologie de LIBQUAL+tm réside dans le fait qu'elle ne cherche pas à évaluer des entrées, des processus ou des résultats, mais un impact supplémentaire de la satisfaction des services fournis par les bibliothèques, et la qualité de service. Même si le protocole britannique de Meilleure Valeur oblige la remontée des résultats des enquêtes de satisfaction dans les Rapports annuels des bibliothèques, on sait trop peu de choses sur les éléments qui conduisent à la satisfaction des usagers (Hernon et Altman, 1998) pour faire confiance au benchmarking de satisfaction des usagers dans différents environnements géographiques, démographiques et culturels. Cela est dû au fait qu'il y a un certain nombre de concepts de base qui sont cruciaux pour le processus de benchmarking. L'un d'eux, comme nous l'avons vu, est le contexte. Un autre est la validité du système, qui pose la question de savoir si notre instrument de mesure calcule en fait bien les facteurs auxquels nous pensons, et quelles autres variables pourraient avoir un impact sur le calcul, mais que nous n'aurions pas prises en compte. Avant d'en être sûr, des tests rigoureux de mesure restent à faire, et les affirmations des développeurs de LIBQUAL+tm sur la validité de leur produit reste à prouver (Hernon et Nitecki 2001). Pour rechercher la validité du système, il est nécessaire de s'appuyer sur des méthodes à la fois empiriques et ethnographiques, pour se défaire d'affirmations qui peuvent être erronées, et pour faire ressortir les différences de contexte et de culture qui affectent à la fois les attentes des usagers des bibliothèques et la performance des bibliothèques. Tous les instruments d'évaluation et toutes les procédures de collecte de données doivent être soumis au même examen minutieux pour devenir des outils de management fiables.

Le point concernant la validité du système s'applique à un certain nombre de domaines où nous souhaiterions voir se développer le benchmarking. Dans ce contexte, considérons plusieurs points. Tout d'abord, comme noté ci-dessus, nous avons tendance à considérer comme fiables les ratios entre les mesures, du simple fait de leur existence, sans tenir compte de la validité de ces chiffres, ou sans voir s'il y a un véritable rapport entre eux . Les données, les ratios entre les données et tout chiffre utilisé à des fins de comparaisons, doivent provenir d'une compréhension

claire du raisonnement sous-jacent, et non de la commodité à utiliser des chiffres disponibles comme l'indique Ford.

Deuxièmement, si nous sommes capables d'identifier le changement nécessaire via le benchmarking, nous devons employer une combinaison judicieuse de benchmarking (d'abord pour servir de diagnostic) et de benchmarking des processus, de façon à extraire des commentaires pertinents au sujet des données, du contexte et des processus qui génèrent ces données.

Troisièmement, j'ai la ferme conviction que si l'on veut que le benchmarking ait un sens pour l'encadrement, on ne peut l'isoler des autres formes d'évaluation et de mesure de performance pratiquées dans l'institution. L'accent mis actuellement sur la mesure de performance ne porte pas sur les entrées, les processus et les résultats, comme cela nous a occupés par le passé, mais sur les conséquences. Dans tous les types de bibliothèques, le résultat et les enjeux ne sont encore pas bien définis, ou à tout le moins, si on peut les définir, on n'a encore développé aucun outil de mesure satisfaisant. Toutefois, si le benchmarking doit s'avérer un véritable outil pertinent, il faudra alors inclure les résultats dans les résultats de l'évaluation dans les activités de benchmarking – probablement de manière souple, à savoir en utilisant l'analyse par des pairs.

Un modèle d'évaluation des performances pour les bibliothèques nationales

Si nous nous tournons à présent vers les bibliothèques nationales, et leur usage du benchmarking pour l'évaluation, il y a un certain nombre de points à prendre en considération. Tout d'abord, les bibliothèques nationales constituent un secteur très particulier, avec des missions nationales bien définies, qui varient de pays à pays, et qui dépendent d'échéances politiques qui évoluent souvent plus rapidement que pour les bibliothèques publiques ou universitaires. L'importance de l'environnement national et politique leur rend la tâche de trouver des partenaires avec lesquels s'évaluer difficile. Il leur est également difficile de trouver des bases de données suffisamment importantes dans lesquelles s'insérer pour faire du benchmarking sur des données classiques. Il en va de même pour recenser les meilleures pratiques, même si ces dernières s'avèrent valables pour les activités de base communes à tous les types de bibliothèques. Elles peuvent aussi conduire du benchmarking générique, en confrontant les divers services qu'elles offrent avec ceux d'institutions de taille comparable, ou de type semblable, comme le montre l'exemple australien. Elles peuvent également envisager des procédures internes de benchmarking, en comparant des secteurs tels que les ressources humaines, la gestion de personnel, ou les procédures budgétaires et comptables ; elles peuvent aussi envisager du benchmarking avec d'autres établissements publics, en analysant les processus concernant leurs activités et en évaluant le service à l'utilisateur. Si l'on songe que le benchmarking dans le secteur privé touche des activités telles que la formation du personnel, les ressources humaines, les technologies de l'information, la gestion des connaissances, l'archivage des données, la communication externe et interne, et que ceci est transverse à tous les secteurs d'activité, la recherche de partenaires pour effectuer des revues par des pairs pourrait s'étendre au-delà du secteur des bibliothèques ou du secteur public, à toute structure qui pourrait partager ses meilleures pratiques. Des entreprises de pointe pourraient se réjouir de se lancer dans du benchmarking et de la revue par des pairs sur leurs processus de travail avec une institution aussi grande et prestigieuse qu'une bibliothèque nationale.

Là où les bibliothèques nationales rencontrent le plus de difficulté est le développement de mesures de performance et la nécessité de les associer à un calendrier social dicté par le gouvernement, comme par exemple participer à la formation et au bien être de la population dans

un sens bien défini, ou entretenir le patrimoine et la culture, ou enfin favoriser la connaissance. Qui sont ses usagers, quelle est sa population cible, quel devrait être son rendement, et quel est le niveau souhaité de rendement ? Plus difficile encore, quel devrait être leurs réalisations ? Il est déjà difficile d'essayer de mesurer le résultat et les réalisations des bibliothèques d'enseignement supérieur, à partir des résultats de l'apprentissage étudiant, ou des travaux de recherche des enseignants, ainsi que celui de bibliothèques de lecture publique, à partir de l'accroissement de la formation continue, et de l'élargissement des offres d'emploi ou des offres commerciales, ou à partir de l'enrichissement personnel grâce à la lecture de loisir, il est plus difficile encore de définir et de mesurer le résultat et les réalisations des bibliothèques nationales. Et pourtant, si le benchmarking doit aller de pair avec la mesure de la performance, c'est ce vers quoi nous tendons.

Si elles consentent à s'inspirer d'autres secteurs d'activité, les bibliothèques nationales pourraient mettre en place un processus de revue par les pairs pour partager des idées sur le développement des mesures de performance dans ce secteur difficile que sont les réalisations et résultats, avec d'autres établissements publics, ou des organisations non gouvernementales ayant une mission comparable dans les domaines sociaux et éducatifs. Une manière de procéder consiste à examiner les différentes missions remplies par une bibliothèque nationale pour ses différents partenaires et à rechercher des pairs remplissant un rôle semblable. Et, comme nous l'avons observé plus haut, la question de la validité du système est un point critique ici, et n'est pas un domaine auquel les politiques ont accordé beaucoup d'attention. L'examen des actes qu'un usager accomplit lors d'une visite, soit physique, soit virtuelle, à une bibliothèque nationale, puis l'examen du résultat de ces opérations, et la différence que l'objectif atteint apportera dans sa vie et dans son travail, devra d'abord être défini par des méthodes ethnographiques, puis confirmé par des études empiriques, qui fourniront une construction sous-jacente à partir de laquelle une évaluation des résultats pourra peu à peu être élaborée. Des partenariats et un dialogue avec des pairs dans le secteur des bibliothèques, dans les établissements publics, et dans les agences non gouvernementales et dans le secteur privé peuvent tous être utiles dans ce processus.

Ainsi, un processus de benchmarking pour une bibliothèque nationale sera moins basé sur une exercice comparatif de résultats entre bibliothèques nationales ou grandes bibliothèques étrangères de diverses cultures, qui toutes évoluent selon des calendriers différents, que sur l'analyse de l'excellence au sein d'un contexte culturel connu, et sur la poursuite d'un processus d'évaluation soutenu par les pairs pour aller vers une ouverture continue vers le changement et l'amélioration.

La meilleure pratique dans le benchmarking inclurait donc :

- une attention particulière accordée au benchmarking des processus et à l'évaluation par les pairs avec d'autres établissements publics et organismes au sein de la même culture ;
- le développement, par l'évaluation par les pairs, de mesures de performance couvrant les entrées, les processus, les produits et les résultats ;
- une attention particulière accordée au changement de principes managériaux, pour s'assurer que le changement est ancré dans l'organisme ;
- une base de recherche sous-jacente, faisant la preuve de la validité du système et une attention particulière au contexte culturel et organisationnel.

En intégrant ces principes, le benchmarking peut devenir un moyen très efficace d'aider les bibliothèques nationales à s'adapter à un changement d'environnement sans perdre de vue l'amélioration de la qualité.