

Du “Benchmarking” aux indicateurs de citations: de la difficulté d’évaluer la recherche par des métriques automatisées

Mohamed Najim,

Professeur Emérite, Université de Bordeaux et ENSEIRB-Matmeca

mohamed.najim@u-bordeaux.fr

Je me réjouis que l’un des premiers orateurs de cette rencontre ait été mon collègue le Pr Yves Gingras¹. J’ignorais, au moment où j’avais fourni le résumé de mon exposé en juillet 2017, qu’il serait des nôtres ici. Il me coupe ainsi l’herbe sous les pieds, dans un sens, car je devais faire longuement référence à ses analyses. Mais à la réflexion, c’est une excellente chose qu’il intervienne en premier car il balise très bien le thème de cette rencontre et je vais tirer avantage de son exposé dans la mesure où, en tant que scientifique qui travaille dans le secteur de la recherche depuis plus de 50 ans, je partage totalement son analyse de sociologue J’aborderai les points suivants dans mon exposé :

- classification des stratégies nationales en matière de R&D et leur intégration dans l’économie libérale (ou néo libérale).

- la notion ou le concept de « Benchmarking ».

- les indicateurs de citation et la question des grands classements internationaux.

En conclusion, je ferai quelques propositions pour le Maroc.

1-R&D et stratégies nationales

Je dégagerai quatre grandes catégories de pays à travers leurs politiques de la R&D, qu’on peut aussi appeler politiques de la science et de la technologie :

1.1 Les pays industrialisés ou en forte croissance, ceux qui ont su intégrer une stratégie de développement de la recherche dans leur développement socio-économique : les USA, les pays scandinaves, Israël, la Corée du sud, l’UE, le Japon, l’Australie.

1.2 La Chine

1.3 Les BRICS : l’Afrique du Sud et le Brésil en particulier. Deux cas particulièrement intéressants sont ceux de Cuba et de la Malaisie. On peut également s’interroger sur la stratégie mise en place, sur des créneaux particuliers, pour la mobilisation des scientifiques en Corée du Nord.

1.4 Les pays où toute stratégie cohérente de la R&D est absente, auxquels appartiennent la quasi totalité des pays arabo-musulmans indépendamment du niveau de leur revenu par capita et les pays dits en développement.

Cette classification reflète la part du PIB que ces pays consacrent à leur R&D ainsi que les choix stratégiques qui ont permis de mobiliser les chercheurs pour piloter la recherche.

Si cette part est importante, elle contribue à un niveau élevé de productivité scientifique attesté par les indicateurs macroscopiques suivants :

¹ Y. Gingras :” Les dérives de l’évaluation de la recherche », Raisons D’Agir,Editeur,2014

- Le nombre absolu de publications
- Le nombre de brevets déposés
- Le nombre de brevets ayant donné lieu à une licence
- Un taux de croissance d'emploi élevé où se détachent très nettement les pays des deux premières catégories, à savoir, les pays industrialisés et la Chine.

2- L'Economie mondialisée et son implication dans la stratégie de la R&D

L'économie libérale a su tirer profit à la fois en période de paix ou de conflits² des fruits de ce développement de la R&D et a cherché à en assurer le contrôle pour gérer une communauté scientifique par essence, rebelle, libre, critique qualités qui la rendent pourtant innovante. Elle s'est donc dotée d'outils qui lui permettent de se donner un cahier de bord pour gérer cette communauté et les fonds qu'elle alloue à la recherche.

3- Le « Benchmarking »

L'économie libérale triomphante s'est donné une métrique pour s'auto-évaluer. Je ferai donc un bref historique sur le *Benchmarking* qui est la matrice génératrice de ces indicateurs.

Le monde des affaires nous a légué ce mot qui vient de bench : banc de mesure, qui a été construit en y ajoutant le mot mark qui signifie évaluer par comparaison.

C'est par analogie avec ce concept de la métrologie que les managers ont introduit le Benchmarking. On a ainsi cherché, sous couvert de « réforme » à exiger l'alignement de la gestion des affaires publiques, en l'occurrence ici la recherche, sur les standards du *Business*³.

En 2008, le patronat français, organisé au sein du MEDEF, équivalent de la CGEM au Maroc, applaudissait sa présidente Mme Laurence Parisot quand elle déclarait : « Benchmarking c'est la santé ». Elle vantait comme l'écrit notre collègue Isabelle Bruno, universitaire à Lille, « les mérites de cette nouvelle panacée pour réformer tous azimuts » : la fiscalité, l'administration, l'université, les services de sécurité, c'est à dire le pays tout entier⁴ y compris les hôpitaux⁵.

Le terme a été utilisé par Robert C. Camp ingénieur chez Xerox, aux US, auteur en 1988 de l'ouvrage : « Benchmarking, The search for industry best practices that lead to superior performances ».

Ainsi le mot « Benchmarking » est-il rentré dans la langue française, comme dans toutes les autres langues, avec des qualificatifs : « comparer, étalonner, mesurer » dans une optique concurrentielle et pour s'améliorer. Se « Benchmarking » c'est être réaliste, c'est se donner les moyens du pragmatisme, c'est savoir qu'on n'est pas seul dans le monde, ni le centre du monde. On va Benchmarking un pays c'est à dire le comparer à d'autres pays, rubrique par

² Effort des USA et de l'Allemagne durant la 2^e Guerre Mondiale et effort continu d'Israël de la Corée du Nord,

³ Luc Boltanski et Eve Chapello : « Le Nouvel Esprit du Capitalisme » Paris, Guillaume, 1999, cité par Isabelle Bruno ci dessous.

⁴ Isabelle Bruno et Emmanuel Didier : « Benchmarking, l'Etat sous pression statistique », Editions Zona, Paris 2013

⁵ Ph. Grimbart et A. Grimaldi « Non aux conflits d'intérêts dans les hôpitaux » Le Monde, 15 février 2018.

rubrique. Quelle est la fiscalité la plus avantageuse, au regard de son efficacité pour l'emploi ? Quelle est l'administration la moins pesante ?⁶

Ainsi les experts ont imaginé une panoplie de méthodes et d'instruments, qui constituent une boîte à outils dont l'une des pièces maîtresses est le Benchmarking.

Le Benchmarking présuppose une procédure de quantification de l'activité que l'on veut évaluer c'est-à-dire de dégager une métrique adéquate et la première difficulté que l'on rencontre est de convertir en chiffres, une organisation et ses différentes séquences,...

Autre difficulté : dégager des indicateurs. Ainsi par exemple pour un enseignant:

- sa présence devant les élèves
- la réaction des étudiants à ses cours
- le taux de succès aux examens
- le succès dans l'obtention d'un emploi
- le nombre de ses publications (articles, livres..)

La disponibilité de masses de chiffres donne l'illusion aux décideurs que l'on peut ainsi quantifier, évaluer, « scorer » une activité. Pour ce faire il suffit de mobiliser des experts.

Les experts, sorte de force supra nationale de l'avis de laquelle dépend souvent l'obtention de financements, n'ont le plus souvent aucune expertise dans le domaine expertisé: en d'autres termes ce ne sont pas des pairs qui évaluent.

4- Les indicateurs de citation tels que le Hindex.

Si même on pouvait prêter quelque crédit au « scoring » par le Hindex ou au classement des universités soumises au « Benchmarking », comment comparer, par exemple, une université française, quelle qu'elle soit (bénéficiant ou non des fonds dits Investissements d'Excellence) qui souffre d'un manque d'accompagnement administratif, qui ne bénéficie que de peu de moyens pérennes⁷, et qui a un ratio nombre d'étudiants/nombre d'enseignants élevé vis à vis à d'une université où ce ratio est beaucoup plus faible. Comment peut-on comparer les performances d'un enseignant chercheur français surchargé de tâches administratives, à un enseignant exerçant dans une grande université américaine se consacrant quasi exclusivement à ses activités académiques (enseignement et recherche) ? Son aura et celui de son université, lui procurent aisément des fonds et font converger vers lui les meilleurs étudiants/doctorants du monde entier !

En conclusion : le Hindex en plus des limitations conceptuelles soulevées plus haut est *culture/dependent* ou *environment/dependent*. Il est encore moins pertinent quand on veut comparer des chercheurs exerçant dans des contextes socio-économiques différents.

De nombreuses voix se sont élevées contre cet indicateur et contre celui dit du « Facteur d'Impact ». Ainsi le 8 décembre 2017, c'est-à-dire le lendemain de la clôture de notre Colloque, trois académies, française, allemande et britannique ont remis une déclaration à Mr Carlos Moedas, commissaire européen à la recherche⁸ où l'on peut lire: « *The use of bibliometrics measures has become an additional tool ; However, there has been too much reliance on bibliometric indices and indicators-based tools as measures of performances by many evaluation committees and exercises, leading to the danger of superficial, over-simplified and unreliable methods of evaluation. This bad practice involving the misuse of metrics has become*

⁶ idem

⁷ Par opposition à ceux obtenus par AAP (Appel A Projets)

⁸ « Statement by three national academies (Académie des Sciences (France), Leopoldina (Allemagne) and Royal Society (UK) « On good practice in the evaluation of researchers and research programmes » in <http://www.academie-sciences.fr/fr/>

a cause of serious concern. ». Et plus loin, on peut encore lire dans cette déclaration: « *Of particular concern are the used impact factors (IF) which are an estimate of the impact factor the journal itself rather than the intrinsic scientific quality of a given article published within it* »

5-Qu'en est-il au Maroc ?

Rappelons tout d'abord ce postulat : « L'évaluation nécessite qu'il existe en premier lieu une activité de recherche à évaluer ».

Je ne peux personnellement qu'applaudir l'initiative des organisateurs de ce colloque pour avoir programmé un exposé sur le Programme Urgence qui devait amener, d'après ceux qui l'ont initié, vers un bond en avant de l'université marocaine.

On pouvait ainsi lire dans le quotidien marocain, Le Matin du Sahara daté du 6 octobre 2009, à propos des contrats, objets du programme Urgence, signés entre les quinze présidents d'université et le gouvernement marocain, en présence de SM le Roi Mohamed VI, à Agadir :

« Ils serviront de cadre contractuel intégré pour le financement des programmes de développement de l'Université, qui visent la mise à niveau et l'extension des infrastructures, l'augmentation de la capacité d'accueil des universités, le renouvellement des équipements et l'amélioration de l'offre de l'enseignement supérieur, la promotion de la recherche scientifique et l'amélioration de la qualité des prestations des œuvres sociales en faveur des étudiants »⁹.

Je me limiterai à la place de la recherche dans un tel Programme où l'Etat a mis en jeu des fonds importants. Mais pourquoi aborder le programme Urgence sous l'angle de la recherche dans un colloque sur l'évaluation ?

L'évaluation est difficile si le diagnostic de ce qui existe en tant que recherche n'a pas été fait.

De nombreux rapports sur cette question ont été rédigés par des institutions étrangères ou par des organismes internationaux mais leurs auteurs ne connaissent pas toujours la réalité marocaine ou ne l'abordent pas clairement de peur d'être taxées d'intrusion dans la politique intérieure marocaine.

J'avais en ce qui me concerne rédigé un rapport¹⁰ sur « La structuration de la recherche au Maroc », pour l'IRES¹¹ sur l'état de la situation de la recherche au Maroc et proposé des mesures à prendre pour sa structuration. Je signalais déjà dans ce rapport que les objectifs en

⁹ https://lematin.ma/journal/2009/Activites-Royales_S-M--le-Roi-preside-a-Agadir-la-ceremonie-de-signature-de-17-contrats-pour-le-developpement-de-l-universite-marocaine/120702.html

¹⁰ https://www.ires.ma/wp-content/uploads/2015/11/structuration_de_la_recherche_scientifique_-_analyse_de_la_situation_et_recommandations.pdf

¹¹ Institut Royal d'Etudes et de Recherches Stratégiques, Dirigé à l'époque par Mr Méziane Belefquih, conseiller de SM le Roi Mohamed VI, qui m'avait personnellement sollicité pour effectuer cette étude.

matière de recherche, concernant en particulier une forte croissance du nombre de publications, ne seraient pas atteints par le programme Urgence.

J'écrivais en effet : « Ces prévisions sont irréalistes car aucun changement prévisible sur la stratégie de la recherche ne semble être envisagé pour justifier un taux de croissance aussi élevé ».

Le tableau ci dessous fournit le nombre de publications effectives, celui qui était prévu et le pourcentage d'erreur dans l'estimation du nombre de publications prévues dans le cadre du programme Urgence :

Années	Publications effectives¹²	Publications prévues¹³	Erreur %
2009	1298	2400	85
2010	1370	2750	100
2011	1506	3200	112
2012	1597	3494	118

Ainsi quand on compare ces chiffres, on ne peut que constater la stagnation du nombre de publications contrairement au saut que prédisait le programme Urgence. Il ne pouvait en être autrement car il n'y a eu aucune mesure pouvant induire une inflexion dans l'évolution de la recherche et par ricochet sur l'évolution du nombre de publications.

Par ce constat, je ne veux pas lancer la pierre aux décideurs qui, à l'instar de nombreux responsables politiques ou syndicaux pensent, souvent, qu'on peut améliorer le système d'ESR en augmentant les moyens mis à sa disposition. Certes c'est une condition nécessaire mais hélas non suffisante. L'élaboration du projet Urgence qui a probablement été pilotée par un bureau de consultants c'est à dire un bureau d'experts qui a omis de faire appel aux pairs.

Cette analyse est totalement en résonance avec l'exposé d'Yves Gingras sur le rôle des experts et celui des pairs.

Ainsi donc, toute évaluation doit au préalable, s'asseoir sur une refondation de l'ESR par la définition d'une stratégie de la politique de recherche. Cette stratégie dégagera un cahier de charge pour les priorités en matière de recherche, la définition des sources de financement, comment les abonder, la place des enseignants chercheurs. Cette stratégie dégagera également des mesures concrètes pour remobiliser les enseignants chercheurs en valorisant ceux qui se mobilisent pour leur établissement, leurs étudiants, ceux qui encadrent des étudiants, ceux qui s'ouvrent au monde socio-économique, ceux qui créent des formations répondant aux besoins de leur environnement. Ainsi et de manière endogène il sera possible aux décideurs de dégager les critères d'évaluation adaptés à la situation spécifique du Maroc.

Il faut laisser aux pairs le soin de dégager des critères appropriés à l'environnement du Maroc, pour évaluer leur activité. Ces critères qui reflèteraient des activités qui ne sont pas

¹² Données disponibles sur le site de l'IMIST : <http://www.imist.ma/>

¹³ Programme Urgence : 17 contrats pour le Développement de l'Université Marocaine, Ministère de l'Education Nationale, Octobre 2009

toujours considérées pour évaluer l'excellence mais qui seraient plus adaptés au stade de développement où se trouve le Maroc.

6-Rôle de la diaspora marocaine

L'évaluation peut inquiéter les enseignants chercheurs au regard de la grande segmentation des domaines couverts par la recherche scientifique. Il existe, dans de nombreux domaines, peu de personnes actives pouvant effectuer l'évaluation en tant que pairs. Ceci peut conduire à des cas où les évaluateurs peuvent être juges et partie. Il y a là en effet un biais dommageable au crédit que l'on peut porter à l'évaluation.

Par ailleurs, le Maroc semble faire face à une pénurie d'enseignants pour le supérieur. Ainsi le vivier des pairs est-il restreint et certains, à raison, peuvent penser que les pairs de ce fait risquent d'être taxés de juge et partie. Pour éviter un tel biais on peut faire appel à des membres de la *diaspora* marocaine reconnus en premier lieu pour leur notoriété scientifique dans le pays où ils exercent. Le Maroc ne serait pas le premier à utiliser la ressource que représente la *diaspora*. D'autres pays en ont fait un usage structuré et intensif : Chine, Inde, Corée du Sud, Israël..etc. Il est intéressant d'analyser de près la manière dont la Chine a utilisé d'abord les ressources financières de sa *diaspora*, comme une source de financement direct provenant de l'étranger appelé *FDI*¹⁴. Ainsi de 1990 à 2000, la diaspora chinoise a contribué pour 50 à 70 pour cent du FDI. Mais ultérieurement la Chine a davantage mis l'accent sur des compétences telles que celles d'universitaires jouissant d'une grande expérience ou d'entrepreneurs, en créant en 2008 le programme appelé «*Thousand Talents Programm*¹⁵». Le Maroc quant à lui est resté à la première étape de ce processus.

Ainsi, par exemple, pour la définition de nouveaux statuts des personnels enseignants chercheurs, qui sont le socle de la recherche, il serait utile de dégager des mesures fortes de mobilisation à la fois des personnels exerçant au Maroc et des mesures spécifiques envers la *diaspora* en lui fournissant des situations matérielles et la valorisation de son expérience qui rendraient attractif l'exercice des métiers de l'enseignement et de la recherche au Maroc.

7-Remerciements

Je tiens à remercier le Conseil Supérieur de l'Education qui, par son invitation, m'a permis d'approfondir ma réflexion sur l'analyse que je viens de présenter. Les questions soulevées après mon exposé oral m'ont conduit à détailler et à illustrer les analyses que j'avais présentées.

Je voudrais remercier ma collègue Isabelle Bruno, universitaire à Lille et Pascal Ragouet, sociologue à Bordeaux, ainsi que Yves Gingras de l'UQAM pour les échanges fructueux que nous avons pu avoir ainsi que mes collègues marocains qui ont accepté de relire la version finale de cet article et me faire part de leurs remarques.

¹⁴ (FDI: Foreign Direct Investment)

¹⁵ <http://www.1000plan.org/en/>