

L'idéologie du jaugeage en recherche *Alfred Kieser*

Classement, évaluation, bibliométrie : il y a de nombreuses possibilités pour renvoyer tous jugements sur les productions scientifiques, de cesser de les lire ou même, il n'est plus nécessaire d'en prendre directement connaissance. Pourtant la revendication d'en décider la qualité à partir des chiffres de production et des listes de classement, est naïve.

Dans son cours d'introduction intitulé « Que signifie et à quelle fin étudie-t-on l'histoire universelle ? », en 1789 à l'Université de Iéna, Friedrich Schiller distingua deux types d'érudits : le « savant qui cultive la science par état » et la « tête philosophique ». Seul ce dernier fait de la recherche pour l'amour du connaître. Richard Feynman, le prix Nobel de physique américain, met prosaïquement au point la motivation du scientifique : « La science c'est comme le sexe. Assurément cela peut donner quelques résultats pratiques, mais ce n'est pas la raison pour laquelle on la pratique. » C'est d'abord l'activité qui motive le chercheur et, en second lieu seulement, le succès. Car la réussite en recherche est hautement incertaine. La plupart des scientifiques ne font jamais de découvertes sensationnelles, ne publient jamais d'articles dans des revues absolument de pointe, n'écrivent jamais d'ouvrages qui font fureur et ne sont pas invités à faire une conférence principale lors de congrès internationaux. Seuls quelques-uns empochent la célébrité.

Comment la production scientifique est-elle cependant mesurée ? Quelle production scientifique est à classer comme supérieure : la découverte de la pénicilline par Alexander Fleming ou la thèse de Max Weber selon laquelle le protestantisme fit avancer le capitalisme ? Cette question est purement insensée. Tout aussi insensée est celle de savoir si un article dans le « *Journal of Marketing* » serait à juger supérieur à un article dans le « *Journal of Finance* ». Même à l'intérieur d'une spécialité, les productions scientifiques s'avèrent souvent non comparables. Cette impossibilité de classement est fréquemment la cause originelle de luttes laborieuses dans les commissions de recrutement. Les estimations subjectives sont inévitables.

Efforts d'objectivation

Dans de telles situations on a volontiers recours à des critères quantitatifs, par exemple en indiquant qu'un candidat a une liste de publications plus riche à présenter qu'un autre. Fréquemment, on argumente aussi que si les articles d'une candidate sont mesurés à l'aune du rang des revues correspondantes, elle est clairement à classer au-dessus des autres postulants.

Si le rang d'une revue scientifique est simplement affirmé, une telle argumentation est toujours entachée nonobstant d'une tonalité subjective. Cette impression disparaît pourtant de la discussion lorsque tous les représentants d'une spécialité sont invités à donner leur avis sur l'importance des revues scientifiques et que ces jugements sont ensuite communiqués. Eugène Garfield a réalisé un tour de force qui va encore plus loin, avec son facteur d'impact. Cet indicateur sur l'impact de l'influence scientifique des revues est aujourd'hui établi par Thomson Science, une filiale de l'agence Reuter, pour un grand nombre de revues choisies de diverses disciplines et placé à la disposition des banques de données, contre rémunération : l'estimation scientifique c'est un *Big Business*.

Comparaisons insensées

Le facteur d'impact repose sur l'hypothèse que des scientifiques citent les travaux d'autres scientifiques pour la raison qu'ils construisent sur leurs résultats. Une revue, dont les articles sont fréquemment cités, produirait ensuite une contribution supérieure à la science et pourrait revendiquer une qualité supérieure à celle d'une revue aux articles de laquelle les scientifiques se réfèrent moins. La production scientifique d'un chercheur peut ensuite être comprise comme la somme des facteurs d'impact mesurés de ses articles. Sur cette base, on peut aussi instaurer un classement des scientifiques et la position d'une faculté entière dans un classement des facultés à

partir des rangs des chercheurs qui y travaillent. Ainsi procède en Allemagne la « *Handelsblatt* [feuille commerciale, ndt] », à vrai dire non pas sur la base des facteurs d'impact, mais au contraire sur la base d'une estimation des revues par les membres de l'association des enseignants de l'université pour la gestion des entreprises.

Les deux premières places du classement de la « *Handelsblatt* » pour les sciences économiques relatives à la gestion d'entreprise sont occupées par deux professeurs de la faculté des sciences économiques de gestion d'entreprise de l'Université de Mannheim. C'est très réjouissant du point de vue d'un membre de l'Université de Mannheim, mais c'est tout de même insensé. Que tous deux, dans leur spécialité, sont des scientifiques de haute réputation internationale, tout connaisseur peut en assurer le bien-fondé. Mais ce que chacun des deux recherche et ce qu'il publie, n'est pas comparable et n'est donc pas non plus classable. L'un explore le marketing, l'autre les banques et le financement. La constatation que l'un est numéro un et donc meilleur que le numéro deux, est aussi insensée que d'affirmer que Tiger Woods est un meilleur sportif que Roger Federer.

Recherche justifiée par le marché

Mais si les classements de scientifique sont insensés, alors les classements d'universités fondés sur eux sont tout aussi insensés. En dépit de leur absurdité, les classements de scientifiques et d'institutions scientifiques sont prodigieusement populaires. De nombreuses commissions professionnelles de recrutement, de doyens et présidents d'universités, s'orientent d'après eux. Leur popularité se fonde avant tout sur le fait qu'ils raccourcissent le processus de jugement. On multiplie simplement les articles du candidat par les facteurs d'impact correspondants et on additionne les points ainsi établis. Pour faire cela on ne doit pas avoir été un jour membre de la spécialité concernée, ni même, en effet, scientifique. Les scientifiques sont bien conseillés de mener un type de recherche qui leur permette d'élaborer des articles lesquels, avec une grande vraisemblance, seront publiés par des revues les plus hautement classées.

Comme le manager en chef d'une entreprise diversifiée connaît quels montants de gains produisent les divers domaines d'affaires, les doyens et directions d'universités connaissent désormais quels facultés et instituts engendrent une recherche « faisant de bonnes opérations sur le marché » et orientent d'après cela leurs décisions sur l'attribution des moyens d'équipement ou bien sur les attributions de postes. La stratégie des universités économiques vise là-dessus à investir un « portefeuille optimal de création de valeur recherche ». Cette stratégie détermine dans une haute mesure le travail des scientifiques. On attend d'eux qu'ils explorent ce qui leur rapportera, à eux et à leurs universités, des points de classement, mais pas ce qu'ils considèrent eux-mêmes comme des projets faisant avancer la science.

À vrai dire, des facteurs d'impacts signalent quelques limitations sensibles. (1) Des publications ne sont prises en compte que dans des revues déterminées, ce sont presque uniquement des revues anglophones, américaines de manière prépondérante. (2) Des publications de manuels ou des contributions dans des recueils spécialisés ou encyclopédies, ne sont pas prises en compte. (3) Seules les citations sont prises en compte, qui surviennent dans les deux ans suivant la publication. Des revues de domaines spécialisés, qui montrent une circulation moindre, se voient donc désavantagées.

Avant tout il est plus que discutable de savoir si des citations expriment de manière prépondérante une reconnaissance pour la reprise de résultats importants d'autres chercheurs qu'elles citent. Un chercheur en cite un autre, non seulement, lorsqu'il en reprend les résultats, mais au contraire aussi lorsqu'il en critique les résultats, tente de se ranger à son école, ou bien au contraire, de s'en délimiter ou encore d'impressionner par sa culture. Volontiers, il cite ceux qui sont généralement et fréquemment cités, car cela augmente la vraisemblance qu'il est lui-même cité. Il est étonnant à présent de rencontrer des revues dotées d'un facteur d'impact élevé, qui sont peu ou pas citées du

tout. Cela laisse soulever la question de savoir si le système du rapport d'expertise d'articles fonctionne de telle manière qu'effectivement les meilleurs surviennent dans les revues les plus renommées.

Rapports d'expertises contestables

Déjà le contingent des refus typiques aux revues renommées, qui atteignent plus de 90% des articles proposés, laissent filtrer un doute. Une grande partie des manuscrits soumis sont, pour préciser, rapidement parcourus en travers par l'éditeur principal ou co-éditeur et refusés par une procédure de « *desk-rejet* » [rejet par le bureau d'édition, *ndt*]. On peut faire l'hypothèse d'une vraisemblance d'erreur au niveau de ce premier criblage. Pour un jugement en détail des articles, qui franchissent cette haie, l'éditeur choisit en général deux, parfois trois ou quatre, experts [*referees*, à savoir « arbitres » *ndt*]. Ceux-ci recommandent le refus ou l'acceptation du manuscrit. L'accord entre les experts sur les articles soumis à la revue est extrêmement mince, ce qui laisse maints auteurs arriver à la conclusion que l'on pût tout aussi bien aussi confier la décision d'accepter ou de refuser un manuscrit à un jeu de dés.

Comme la recherche l'a prouvé, la corrélation des jugements d'experts sur la qualité d'un manuscrit d'avec les citations ultérieures, n'est que très mince. Conséquemment, la cour suprême américaine de justice l'a constaté, il y a peu de temps, que la circonstance que des revues sont réalisées par des pairs n'a pas permis de conclure sur la qualité des contenus des revues scientifiques.

Particulièrement instructive à cet égard est une étude de Peters et Ceci. Ils sont choisis, à partir de douze revues psychologiques, de prendre en compte un seul article de chacune de la dernière année et ont remplacé les noms des auteurs et des institutions très réputées par des anonymes. Ils en ont changé aussi le titre, le résumé et l'introduction. Puis ils ont envoyé l'article — non modifié par ailleurs — à la même revue qui l'avait déjà publié. Le résultat : seuls trois des douze manuscrits furent reconnus comme déjà publiés et huit des neuf restants furent désormais refusés.

Dans l'expérimentation de Peters et Ceci, le fait que les caractérisations des auteurs aient été changées en « anonymes » a certainement joué un rôle. Les auteurs des articles authentiques soumis étaient sans exception bien connus. Mais avec des auteurs, connus par la communauté, les éditeurs se donnent du mal pour trouver des experts qui apprécient les mêmes articles que les auteurs et ils se résolvent, lors d'avis contradictoires des experts, plutôt à accepter l'article. Outre une prédilection pour des théories et méthodes déterminées, les experts ont encore d'autre préjugés. Ils sont avant tout conservateurs, parce qu'ils veulent faire carrière : d'experts, ils veulent devenir co-éditeurs puis éditeurs en chef et enfin directeur de revues plus importantes et, pour cette raison, ils évitent donc le risque de recommander l'acceptation d'un article, avec lequel il faut compter qu'il se heurtera au refus de la *Scientific Community*. En tant qu'auteur, on doit, non seulement rédiger un bon article, mais avant tout avoir la chance que l'éditeur choisisse, non pas un, mais deux ou trois experts, chez lesquels l'article soumis n'éveillera ni jalousie, ni orgueil ou dogmatisme et qui auront une certaine sympathie pour l'article théorique choisi. Tout particulièrement à cause de la cohérence extrêmement faible entre jugements d'experts et citations d'un article, le facteur d'impact est donc un mauvais indicateur de qualité ; combien plus mauvaise encore l'addition des articles jugés par lui d'un auteur.

Le classement en tant que prophétie se réalisant elle-même

Mais le plus gros problème des systèmes de classement c'est qu'ils influencent ceux qui sont concernés par eux. Dans le classement très populaire de la *Handelsblatt* pour les professeurs d'économie de la gestion d'entreprise, deux cents professeurs sont listés, les plus couronnés de succès des pays germanophones. Parmi eux, aucun qui ait publié une monographie de ses résultats de recherche prépondérants, quand bien même ceux-ci puissent être aussi influents. Dans la *Handelsblatt*, on ne prend aussi en compte, pour préciser, que des publications dans des revues. La conséquence c'est des chercheurs rédigent peu de monographies, ce que personnellement je trouve

très regrettable, car lorsque j'étais étudiant, c'est avant tout à partir des biographies que j'ai beaucoup appris. Pour les revues germanophones, ils n'écrivent parce que cela ne rapporte pas de points ou seulement peu, au plus des « ressucées » de leurs publications anglaises ou de second rang. Chercheurs et institutions mal classées perdent toute considération auprès des étudiants, politiciens et administrateurs d'universités. Ils subissent des pertes considérables dans leur crédits d'équipement et leur classement dégringole éventuellement. Alors les classements réalisent eux-mêmes leurs prophéties.

Par ailleurs, le système dérive vers les combines. Par exemple des scientifiques citent leurs articles plus fréquemment que jusqu'à présent. Ils privilégient aussi la citation d'articles parus dans des revues avec un facteur d'impact élevé. Cela octroie de l'importance à leurs propres articles. Des revues bien cotées en terme d'impact, n'ont pas seulement un haut prestige pour la raison qu'elles sont souvent citées, mais elles sont fréquemment citées aussi parce qu'elle montre un grand prestige. Avec la promesse, qu'ils reprendront la plus grande partie du travail, ils donnent envie à quelques collègues renommés de faire des publications ensemble. Cela ne fait pas que rehausser la vraisemblance de l'acceptation des manuscrits proposés à la publication — pour les éditeurs, les auteurs ne sont pas anonymes —, mais plus tard au contraire la perspective de se voir cités. Lorsque autrefois, on demandait à un candidat au poste de professeur ce qui l'intéressait dans la recherche, il répondait : « la manière dont fonctionne un groupe de recherche m'intéresse », ou bien : « Je fais de la recherche parce que l'attrait rehausse la publication ». Aujourd'hui, il vous dit : « Je veux publier dans un *Journal* classé A »

Et il sait aussi ce qu'il doit faire pour cela : choisir une thématique, une théorie et une méthode qui sont à la mode. Autrement dit, il fait exactement ce qui s'est avéré bon de faire dans un passé récent, juste en changeant un petit peu. Au mieux, il répète une expérimentation avec une méthode quelque peu plus productive. Des progrès de méthodologie sont plus faciles à constater que des progrès dans les résultats. Pour préciser, si des chercheurs communiquent des découvertes réellement nouvelles, ils risquent que les experts ne puissent pas les mettre en lien avec les connaissances auxquelles ils sont habitués et donc les refusent. Si l'auteur ne fait pas cela foncièrement assez de lui-même, alors experts et éditeurs veillent à ce qu'un article soumis soit un article restant dans le *mainstream* [courant dominant]. Afin de ne pas simplement bousculer les experts ni provoquer de décision négative, les auteurs font tout ce que proposent les experts, même s'ils ont l'impression que de ce fait leur articles en est plutôt plus mauvais que meilleur. Bruno Frey parle pour cette raison de « prostitution de la publication ».

Des chercheurs sur le sentier battu

Et des scientifiques tentent de publier le plus d'articles possibles à partir de leurs projets de recherche. Cela procède comme avec ces livres de combinaisons d'images, avec lesquels en les feuilletant on peut engendrer divers chapeaux, visages, ventres et jambes de nombreux personnages divertissants. Avec cette méthode, un patron dans la gestion économique à peine trentenaire produit 36 publications internationales en trois ans. Éditeurs et maisons d'édition combinent pareillement : Ils poussent les auteurs, à citer le plus d'articles possibles qui sont parus dans leur revue : « *We have noticed that you cite « Leukemia » only once in 42 references. Consequently, we kindly ask you to add references of articles published in “Leukemia” to your present article. [Nous avons remarqué que vous citez « Leukemia » seulement une fois sur 42 références. En conséquence nous vous demandons aimablement d'ajouter des références d'articles publiées dans « Leukemia » à votre présente publication]* ». Facultés et universités participent aux combines. Ainsi doyens et présidents poussent les chercheurs de leurs institutions à publier dans des revues qui ont un classement comme celui à la base du « *Financial Times* ». Ils engagent volontiers des candidats — et aussi en tant que scientifiques invités — qui ont publié dans des revues « *high impact* » et promettent de nombreuses citations, pour ainsi améliorer le classement de leurs institutions. Ainsi les classements créent-t-il la réalité. Ils n'instaurent pas de manière transparente un « marché » déjà existant, bien plus ils

engendrent un pseudo-marché [truqué, *ndt*]. Une formation de profil selon l'échelle de mesure des classements ne veut pas dire construire un autre profil que celui de facultés en concurrence, mais au contraire s'efforcer au même profil avec une plus grande perfection : ainsi pour devenir des concurrents couronnés de succès mais dans leur « façon d'être » encore plus « hors pair ».

Les acteurs du système sont les éditeurs, les experts, et les directions d'université, qui tous s'efforcent à ce que leurs classements s'améliorent. Et parce qu'ils prennent toutes les décisions d'une importance décisive pour les carrières, les chercheurs veulent exactement les points conquis qui sont calculés dans le système. Ils ne s'orientent plus d'après les questions qui se posent à la recherche, qu'ils estiment comme importantes en terme de gain de connaissance qu'ils pourraient en retirer, mais ils accumulent des points pour leurs listes de classement. Ils ne s'adonnent plus au plaisir et à l'attrait de la recherche scientifique, mais suivent au contraire les sentiers battus et jalonnés par des listes de classement. « Des « têtes philosophiques » se mutent en « savants qui cultivent la science par état ».

Alfred Kieser — FAZ [Frankfurter Allgemeine Zeitung] du 11 juin 2010.
(Traduction Daniel Kmiecik)

Alfred Kieser enseigne la gestion économique générale d'entreprise et la théorie de l'organisation à l'Université de Mannheim. Ce texte est une rédaction légèrement écourtée de son cours d'adieu tenu à la fin du moins de mai 2010.