

Recherche

PAR Martin Clavey PAR Adria Fruitos

VICES ET VERTUS DE L'ÉVALUATION PAR LES PAIRS

Les chercheurs rêvent tous d'être à l'origine d'une découverte majeure. Pour s'assurer de la validité de leurs travaux, les articles sont soumis à l'expertise de leurs confrères. Ce système, qui repose sur des critères scientifiques autant que sur un principe de confiance, n'est pas exempt de failles.

C'était en février 2014. La très sérieuse maison d'édition allemande de publications scientifiques Springer annonçait dans un communiqué le retrait de ses archives de seize articles consacrés à des travaux qui n'avaient de scientifiques que le nom. Totalement dénuées de sens mais présentant l'aspect et le jargon des publications spécialisées, citations et références à l'appui, ces « études » avaient pourtant figuré au programme de conférences sur les sciences de l'informatique et l'ingénierie entre 2008 et 2013, avant d'être publiées. Un « tissu d'absurdités » de l'aveu même de l'éditeur, alerté par Cyril Labbé, un informaticien français de l'université Joseph-Fourier à Grenoble (Isère) prompt à débusquer les fausses enquêtes générées par informatique. « Malheureusement, l'édition scientifique n'est pas plus immunisée contre la fraude que contre les erreurs », avait reconnu Springer.

Comment l'absurdité des propos contenus dans ces publications a-t-elle pu échapper à l'expertise de spécialistes ? Depuis les années 1950, le monde de la recherche s'est organisé collectivement, à l'échelle internationale, pour se prémunir contre les annonces farfelues, un peu trop optimistes, voire clairement erronées. Face à la multiplication des champs d'études, un système a été instauré au sein des revues et des conférences scientifiques spécialisées, qui repose sur l'évaluation par ses pairs... et la confiance.

Pour la plupart des disciplines, les revues scientifiques, telle la prestigieuse *Nature*, sont le lieu privilégié pour exposer ses travaux. Le chercheur soumet une première

version de son article à l'éditeur de la publication, qui charge un comité d'experts, appelés relecteurs, d'évaluer rigoureusement la qualité et la nouveauté de la recherche. Après analyse, chaque relecteur rend un rapport. Sur la base de ces avis, l'éditeur choisit soit de refuser l'article, soit de l'accepter sous réserve de modifications mineures ou majeures. Dans ce dernier cas, un cycle d'allers et retours s'enclenche dans le but d'enrichir la qualité de l'article. Ces échanges peuvent durer des semaines, voire plusieurs mois et nécessiter des expérimentations complémentaires.

« Sur la base de l'avis exprimé par les relecteurs, l'éditeur choisit de refuser ou d'accepter l'article, sous réserve de modifications »

Dans le domaine de l'informatique, les scientifiques cherchent tout autant à être publiés mais « le plus important est de présenter ses travaux dans des conférences internationales, comme la Constraint Programming Conference (CP) ou l'International Joint Conference on Artificial Intelligence (IJCAI), raconte Charlotte Truchet, maître de conférences à l'université de Nantes (Loire-Atlantique) et membre du comité scientifique de la CP Conference. Contrairement aux éditeurs, les organisateurs de conférences n'ont pas

le temps de multiplier les échanges avec l'auteur. Le comité scientifique prend sa décision après une à deux semaines de discussions. S'il est accepté, l'article sera exposé puis publié dans les actes de la conférence. »

Préserver l'anonymat des relecteurs

Ce mode d'évaluation échappe aux universités, instituts de recherche ou entreprises pour lesquels travaillent les chercheurs afin que les seuls critères purement scientifiques soient pris en compte. Pour éviter les conflits d'intérêts, la plupart des revues ne transmettent pas les noms des relecteurs aux auteurs qui demandent à être publiés. Certaines ont même mis en place un système dit en « double aveugle ». « Dans ce cas, le chercheur ne connaît pas l'identité de ses évaluateurs, mais ces derniers ignorent aussi le nom de l'auteur tant que son article n'est pas publié », explique Christine Kosmopoulos, ingénieur de recherche au CNRS et rédactrice en chef de la revue *Cybergéo*. Un procédé loin de faire l'unanimité. « Le double aveugle est très rare en chimie, souligne ainsi François-Xavier Coudert, chargé de recherche du CNRS à l'institut de recherche de Chimie Paris et éditeur pour la revue de chimie *Adsorption Science & Technology*. Personnellement, je n'y ai pas recours. Cette méthode suppose qu'une partie du contexte de la recherche ne soit pas expliquée pour préserver l'identité de l'auteur, ce qui est problématique pour la compréhension du texte. » Chaque revue scientifique ou conférence fixe ses propres critères de sélection des



relecteurs, mais l'objectif reste de vérifier l'originalité et la pertinence du travail sur la base de critères scientifiques très stricts.

Un processus nécessaire mais insuffisant

Ainsi *Nature*, l'une des plus importantes publications pluridisciplinaires, soumet-elle ses relecteurs à une série de questions précises : Y a-t-il des raisons de s'interdire de publier cet article ? La méthodologie est-elle valide ? Les résultats sont-ils statistiquement fiables ? Cet article fait-il référence à la littérature précédente de manière appropriée ? Pour Arnaud Saint-Martin, sociologue du CNRS au Centre européen de sociologie et de science politique à Paris et codirecteur de publication de la revue *Zilsel*, « ce processus d'évaluation par les pairs est fondamental. Il demande, en tant que relecteur, un important travail de

contre-argumentation pour justifier le refus d'une publication – ce qui arrive fréquemment –, mais il est essentiel. » La publication d'études fantasmées par Springer démontre les limites du procédé d'évaluation, basé sur la confiance. Au sein des conférences qui les avaient sélectionnées, « le processus de "revue par les pairs" a été très mal fait. Ou bien les relecteurs n'étaient pas compétents et se sont laissés impressionner par le jargon », avançait en 2014 Cyril Labbé, l'informaticien français à l'origine de la découverte de la supercherie. « Ce n'est pas parce qu'un travail a été évalué par des pairs qu'il est forcément juste. C'est un premier crible, une première étape de validation nécessaire, mais qui n'est pas suffisante pour assurer la validité d'une recherche », explique François-Xavier Coudert, l'éditeur de *Adsorption Science & Technology*. Il faudra, entre autres, que les travaux soient repris par d'autres

équipes de chercheurs pour les confirmer et exclure ainsi toute erreur, volontaire ou involontaire. Mais la reproduction d'une expérience peut s'avérer difficile, voire impossible dans certaines disciplines. Les scientifiques doivent régulièrement faire paraître des articles pour démontrer la réalité et le sérieux de leurs recherches. L'évolution de leur carrière et l'obtention de financements dépendent souvent de ces parutions et de la notoriété des revues qui les accueillent.

L'expression du sociologue américain Clarence Marsh Case, « publish or perish » (« publier ou périr »), qui remonte à 1917, reste régulièrement usitée pour pointer à la fois la nécessité de publier mais aussi les effets pervers de ce mode d'évaluation par ses pairs. Pour François-Xavier Coudert, ce système « génère des frustrations, mais globalement, c'est le moins mauvais que l'on ait trouvé. » ♦