

علم سنجی، ضعفها و قوت ها

محمد صالح مصلحیان

گروه ریاضی محض، دانشگاه فردوسی مشهد

شاخه ریاضی، فرهنگستان علوم ایران

در دهه ۱۹۶۰، از مفهوم تحلیل استنادی که بر چهار محور اصلی یعنی ارزیابی مؤلفان، انتشارات علمی، مراجع، و ارجاعات تأکید داشت به کتابداران کمک کرد بتوانند با توجه به بودجه کتابخانه‌ها، کتاب‌ها و مجلاتی که بیشترین استفاده و کاربرد را برای مخاطبان کتابخانه دارند شناسایی و خریداری کنند. از طرفی، همسو با گسترش پژوهش‌های علمی، لزوم یافتن روش‌ها و ابزارهایی برای ارزیابی نظاممند فعالیت‌های تحقیقاتی احساس می‌شد. تا این که در دهه ۱۹۹۰، همراه با شکل‌گیری نمایه استنادی معتبر بین‌المللی تامسون رویترز (ISI) (که امروزه به پایگاه Web of Science (WOS) معروف است و توسط شرکت Clarivate Analytics منتشر می‌شود)، سنجه ضریب تأثیر مجله (JIF) Journal Impact Factor معرفی شد که به کمک آن، بر پایه میزان استنادات به یک مقاله، کیفیت آن مقاله، نویسنده آن و حتی نشریه‌ای که آن مقاله را چاپ کرده است، مورد ارزیابی قرار می‌گیرد. به عبارت ساده، ضریب تأثیر JIF یک مجله در سال n برابر نسبت تعداد ارجاعات به تعداد مقالات مجله در طی دو سال است که متوسط تعداد ارجاعات به ازای هر مقاله را به دست می‌دهد. شاخص JIF در سال n برای یک نشریه یعنی نسبت تمامی استنادهایی که در سال n توسط ۱۳۰۰۰ مجله در پایگاه داده WOS به مقالات مجله در دو سال $n-1$ و $n-2$ داده شده است به تمامی مقالات قابل استناد آن مجله در همان دو سال $n-1$ و $n-2$.

پس از مدتی، پایگاه استنادی رقیب WOS، یعنی اسکوپوس (Scopus) و وابسته به شرکت الزویر (Elsevier)، سنجه دیگری را به نام Scimago Journal Rank (SJR) معرفی کرد که برخلاف JIF، وضعیت استنادی مجلات را در بازه زمانی سه‌ساله و به صورت وزنی بر مبنای استنادات از حدود ۳۰۰۰۰ مجله Scopus منتشر می‌کند. به این معنا که در SJR، برای استناداتی که از نشریات معتبر به یک مقاله داده می‌شود، وزن بیشتری در نظر گرفته می‌شود. به دلیل کیفی‌تر بودن این شاخص و اشکالاتی که به شاخص JIF وارد است، بعضی مطالعات نشان می‌دهد که احتمالاً، در سال‌های آینده، اطلاعات علم سنجی ارائه شده در Scopus بیشتر از WOS مورد استفاده قرار می‌گیرد.

این نوشتار ترجمه‌ای است با اضافه‌ها از

M.S.Moslehian, Impact factor, an inadequate yardstick!, The Newsletter of European Mathematical Society (to appear).

که به صورت سخنرانی در گروه علوم پایه در دی ماه ۱۳۹۹ ارائه گردید.

از جمله اشکالات JIF می‌توان به موارد زیر اشاره کرد: عدم حساسیت به تفاوت گرایش‌ها در یک موضوع علمی، عدم حساسیت به اعتبار مقالات ارجاع‌دهنده، نامناسب بودن دوره محاسبه دو سالانه ارجاعات برای بعضی موضوعات علمی مانند ریاضیات (برای ریاضیات دوره پنج ساله مناسب‌تر است)، در نظر گرفتن تعداد مطلق ارجاعات در فرمول محاسبه بدون لحاظ کردن کیفیت مجلات ارجاع‌دهنده و نقش ارجاع در مقاله ارجاع‌دهنده.

توجه بیش از حد به ضریب تأثیر و فشار نامتعارف بر پژوهشگران برای چاپ مقاله در مجلات با ضریب تأثیر بالا می‌تواند آن‌ها را به تحقیق در حوزه‌هایی سوق دهد که نه تنها در حاشیه علم روز بوده و مسائل ساده و غیرمهم را در برمی‌گیرند، بلکه در کوتاه‌مدت نتایجی قابل چاپ به دست می‌دهد. بنابراین عده زیادی از افراد، به‌خصوص در کشورهای در حال توسعه که عموماً "از آموزش علمی مناسبی برخوردار نیستند، روی چنین موضوعاتی کار می‌کنند. به‌علاوه، همین افراد به یکدیگر ارجاع داده و به این ترتیب، تعداد ارجاعات در چنین حوزه‌هایی بسیار زیاد می‌گردد که خود به خود ضریب تأثیر مجلاتی که در چنین حوزه‌هایی مقاله چاپ می‌کنند را بالا می‌برد. افزون بر اینها، امکان دستکاری توسط مسئولان مجله مانند تقاضا از نویسندگان برای افزودن بعضی مقالات مجله به فهرست مراجع یک مقاله و نیز کنترل تعداد مقالات در هر شماره می‌تواند بر ضریب تأثیر مجله موثر باشد.

بعضی مجلات «خاص» در میان مجلات JCR (اینها مجلات منتخبی در WOS هستند که JIF ناصفر دارند و از کیفیت‌های ویژه‌ای برخوردارند) ضریب تأثیری بالاتر از مجلات فاخر در رشته خود دارند؛ لیکن کیفیت علمی مقالات آن‌ها در جامعه دانشمندان تراز اول جهان مورد شک است. این در حالی است که سنجه SJR که وزنی محاسبه می‌شود، برای این مجلات خاص به میزان زیادی از مجلات پراعتبار پایین‌تر است که نشان می‌دهد، در عمل، SJR شاخص مناسب‌تری است تا JIF.

دلیل این تفاوت در JIF را می‌توان ناشی از عوامل زیر دانست: در بعضی موضوعات، تعداد زیادی مقاله چاپ می‌شود، در بعضی حوزه‌ها، تعداد زیادی محقق وجود دارد، در بعضی حوزه‌ها، مجلات با دسترسی آزاد بیشتری وجود دارد.

شاخص‌های علم‌سنجی مانند JIF، نشانگر ویژگی‌های کلی یک مجله هستند، نه یک مقاله مشخص. ارجاع، بعد از چاپ مقاله ظاهر می‌شود و لذا تقاضای دانشگاه‌ها و مؤسسه‌های تحقیقاتی برای ارجاع زیاد در دوره زمانی کوتاه، ممکن است محققان را وادار کند از تحقیقات در حوزه‌های عمیق که نیاز به کار علمی سخت و طولانی تری دارند دست شسته و سوی تحقیقات خود را به سمت حوزه‌های علمی ساده‌تر و زودبازده‌تر عوض کند. اطلاعات عددی حاصل از علم‌سنجی، بخشی از حقایق را درباره یک مقاله می‌گوید، نه تمام حقایق را. در واقع، همه مقالات در یک مجله، تعداد ارجاعات یکسانی ندارند و حتی یک مقاله پر ارجاع (مثل یک مقاله مروری) می‌تواند عدم ارجاع به مقالات دیگر در آن مجله را جبران کند؛ بنابراین، صرف چاپ مقاله در یک مجله با ضریب تأثیر بالا را نمی‌توان دلیل بر خوب بودن مقاله دانست.

اینجا باید به نقش ارزیابی انسانی اشاره کرده و بر اهمیت آن تأکید نمود. نبودن منابع اطلاعاتی قابل اعتماد، عدم رعایت استانداردهای ارزیابی همتا (Peer Review)، و نیز تعداد بسیار کم داوران حرفه‌ای (یعنی داوران عادل و عالم) در ایران، مدیران دانشگاه‌ها و مؤسسه‌های تحقیقاتی را برای ارزیابی‌های استخدای، تمدید قرارداد، ترفیع، و ارتقای مرتبه به سوی استفاده از ابزارهای

علم سنجی سوق داده است. قطعاً ابزارهای علم سنجی بسیار عینی (Objective) هستند و در برابر ارزیابی‌های انسانی قرار می‌گیرند که اساساً غیرشفاف و ذهنی (Subjective) بوده و از حب و بغض و احساسات انسانی بری نیستند.

با این حال، ایرادهای وارد بر شاخص‌های علم‌سنجی از جمله ارجاعات، توسط جمع‌کثیری از محققانی که در حوزه‌های ساده‌تر و زودبازده‌تر کار می‌کنند و به مقالات یکدیگر ارجاع می‌دهند، ارزیابی انسانی در کنار شاخص‌های علم‌سنجی را ضروری می‌کند. به‌ویژه، به‌منظور تشخیص تحقیقات فاخر که در مقالات متأخر چاپ شده‌اند، نمی‌توان صرفاً بر پایه معیار مقاله داغ (Hot paper)، یعنی دریافت بیشترین تعداد استناد در دوماه اخیر، میزان تأثیرگذاری مقالات را تعیین کرد؛ چرا که در بعضی موضوعات مانند ریاضی، باید حدود ۵ سال صبر کرد. افزون‌براین، استفاده از شاخص‌های علم‌سنجی برای ارزیابی دانشمندان جوان یعنی افرادی که سابقه کاری کمتر از ۶ سال دارند، بسیار نامناسب است و نوعی بازی با سرنوشت آن‌ها است. از سوی دیگر، مقاله‌ای که در بیش از ۵ سال، ارجاعات بسیار کمی از دانشمندان معمولی و در مجلات معمولی گرفته است را نمی‌توان به‌عنوان مقاله‌ای تأثیرگذار تلقی کرد. در نقطه مقابل، مقاله‌ای با ارجاعات زیاد نیز لزوماً، بااهمیت تلقی نمی‌شود. به‌خصوص هنگامی که استنادها توسط حلقه‌های بزرگی از دانشمندان در حوزه‌های تقریباً غیرمهم به‌وجود آمده باشد.

خلاصه این که شاخص‌های علم‌سنجی می‌توانند مفید باشند، اگر بر مزایا و معایب آن‌ها اشراف داشته باشیم و در کنار آن، ارزیابی انسانی به‌کمک دانشمندان حرفه‌ای نیز مورد استفاده قرار گیرد. ارزیابی دقیق فعالیت‌ها و آثار علمی بدون هریک، یعنی شاخص‌های علم‌سنجی و ارزیابی انسانی، ممکن نیست، و این دو بدون دیگری مثل کالسکه بدون چرخ است.