

درباره کیفیت و تأثیر تحقیقات: آنچه پنج دهه حضور در محیط دانشگاهی به من آموخت

بهروز پرهامی

دانشکده مهندسی برق و کامپیوتر

دانشگاه کالیفرنیا - سانتا باربارا

پست الکترونیکی: parhami@ece.ucsb.edu

ترجمه: علیرضا خلیلیان

ORCID: 0000-0002-4079-703X

چکیده

کیفیت و تأثیر تحقیقات برای اهداف مختلفی همچون ترفیع یا افزایش حقوق در دانشگاه، پیشرفت در آزمایشگاه‌های تحقیقاتی، پرداخت کمک هزینه توسط سازمان‌های تامین‌کننده بودجه‌های تحقیقاتی و اعطای افتخارات علمی و فنی باید تعیین شود. تحقیقات باید در نهایت در خدمت مردمی باشد که از آن حمایت می‌کنند اما کارهای بیشتری برای ارزیابی تحقیقات لازم است تا بتوانیم تأثیر اجتماعی مستقیم آن‌ها را بسنجیم. برنامه‌های تحقیقی درون رشته‌ها و بین رشته‌های مختلف به هم مرتبط هستند، به طوری که اغلب اوقات حلقه‌ای از بخش خاصی از تحقیقات قبل از این که به کاربرد عملی برسد، در طی چندین پروژه (و گاهی نسل بشری) تکمیل می‌شود. در زمینه‌های علمی، انتظار از تحقیقات این است که نظریه‌های پیشنهادی چگونه عملکرد دنیای ما را اعتبارسنجی کنند یا آن‌ها را به چالش بکشند. در زمینه‌های فنی، تحقیقات با بهره‌گیری از طراحی‌های خلاقانه یا مدل‌ها و روش‌های ارزیابی دقیق‌تر در خدمت ساخت دستگاه‌ها، سیستم‌ها و فرایندهای بهتر است. در این مقاله من برخی از نظرات را راجع به معیارهای مورد استفاده برای سنجش کارهای تحقیقاتی علمی و فنی و تأثیرات آن‌ها ارائه می‌کنم.

واژگان کلیدی: سنجش، تأثیر، بازدهی تحقیقات، انتشار تحقیقات،

۱. مقدمه

کار در چندین هیئت ویراستاری برای آی‌تریپل‌ای و حامیان مالی کنفرانس‌ها و ناشران نشریه‌ها. در اینجا من برخی از ملاحظات، ایده‌ها و تجربه‌های خودم و خطراتی که در کمین همکاران جوان‌ترم در محیط‌های دانشگاهی و دیگر موقعیت‌های نیازمند ارزیابی کارها و خروجی‌های تحقیقاتی قرار دارد را با شما در میان می‌گذارم.

تا امروز من تقریباً ۵۰ سال درگیر سنجش کیفیت و تأثیر تحقیقات بوده‌ام. تجربه‌های من عبارتند از عضویت و سرپرستی کمیته ترفیحات دانشگاه (در ایران و آمریکا)، سه سال معاون آموزشی امور هیئت علمی در کالج مهندسی دانشگاه کالیفرنیا در سانتا باربارا، داوری‌های فراوان مقالات تحقیقی و پیشنهاد‌های کمک هزینه پژوهشی و

الف) استواری نظری/مفهومی

- آیا نظریه پشت تحقیق در صورت وجود، منطقاً قابل اجرا و به‌طور کامل توجیه شده است؟
- آیا نویسندگان کارهای مرتبط پیشین را به‌درستی تفسیر و متناسباً تلفیق کرده‌اند؟
- آیا فرضیه‌ها در صورت وجود، که از نظریه قابل آزمایش گرفته شده باشد، به‌وضوح بیان شده‌اند؟

ب) استواری روشگانی

- آیا طراحی تحقیق (نمونه، روش، اندازه‌گیری) برای مسئله مورد نظر مناسب است؟
- آیا فنون تحلیلی مناسب روی داده‌های جمع‌آوری شده به‌کار رفته است؟
- آیا نتایج مطالعه به‌درستی تفسیر شده‌اند؟
- آیا نتیجه‌گیری‌ها و/یا استنباط‌ها به‌درستی از یافته‌های تحقیق استخراج شده‌اند؟

پ) نوآوری و سهم علمی

- آیا مقاله باعث پیشرفت دانش داخلی یا پیرامون رشته مرتبطش می‌شود؟
- آیا یافته‌ها و استنباط‌های آن‌ها جالب توجه هستند یا یک کار عادی بی‌حاصلند؟
- آیا مقاله مورد علاقه افراد بسیار یا دست‌کم گروه یا دسته خاصی در آن رشته هست؟

ت) ارتباط

- آیا مقاله به‌طور واضح نوشته شده است (ساده، زبان مستقیم، بدون رنگ و لعاب)؟
- آیا مقاله در قالب منطقی که ایده‌های مرتبط بجا گروه‌بندی شده باشند، طرح‌بندی شده است؟
- آیا نکات اصلی به‌سادگی قابل درک است که به‌وضوح گفته شده باشد که چه چیزی جدید است؟

شکل ۱: راهنمای ارزیابی مقاله

حتی وقتی داوران داخلی یا خارجی باصلاحیت هم پیدا شود، پدیده‌ای مشابه تورم نمره^۲ برای دانشجویان ممکن است رخ دهد. محقق الف از محقق ب بیش از حد تعریف و تمجید می‌کند و انتظار دارد محقق ب یا استاد راهنمایش در آینده به‌طور متقابل این کار را برای او انجام دهد. این چیزی است که من مستقیماً در سمت‌های رسمی و کارهای مشاوره‌ای غیررسمی گوناگون مشاهده کرده‌ام که جزئیاتش را ذیلاً بیان می‌کنم.

اگر شما از تعریف و تمجید و اعتباربخشی که به افراد به‌خاطر نوآوری خاصی داده می‌شود مطلع باشید، احساس می‌کنید که هر کدام به‌تنهایی باعث پیشرفت شده‌اند. حدود چهار دهه پیش من به موسسه‌ای در دانشگاه شریف راهنمایی کردم که به‌ازای مطالبی که هر فرد تحت ارزیابی برای هر مقاله ارسال می‌کند باید درصدی مشخص شود که سهم او را مشخص کند و جمع درصدها ۱۰۰ شوند. این کار باعث گفت‌وگویی بین نویسندگان شد که هر کدام چقدر سهم دارند. همچنین این کار باعث حذف یا کاهش این عادت شد که اسم افرادی که سهمی در مقاله نداشتند را به‌خاطر لطف در مقاله درج می‌کردند. این سیستم اگرچه کامل نبود و منجر به کشمکش‌هایی شد، اما بهترین کاری بود که آن موقع می‌توانستیم انجام دهیم.

برای جلوگیری از این موقعیت شرم‌آور که هر یک از نویسندگان همکار ادعا کنند همه ایده‌های مقاله متعلق به اوست (محرمانگی داوری و سنجش مانع از افشای چنین تقلب‌هایی می‌شود)، اکنون بیشتر نشریه‌ها از نویسندگان می‌خواهند اعلامیه‌ای درون مقاله درج کنند که در آن سهم و نقش هر نویسنده مشخص شده باشد. یک نمونه اعلامیه این است: «الف ایده‌های پروژه را در ذهن پرورش داد. ب و ج آزمایش‌ها را طراحی کردند. الف و د آزمایش‌ها را انجام دادند.

۲- این اصطلاح به دو چیز ممکن است اشاره کند: یکی دادن نمراتی به دانشجویان بیشتر از استحقاقشان که معدل نمرات داده شده به کل دانشجویان را بالا می‌برد. دوم به‌طور افزایشی نمرات بالاتر به کارهایی داده شود که گذشته همان کارها نمرات کمتری می‌گرفتند. م.

هر رشته‌ای معیارهای خودش را برای ارزیابی کیفیت و تاثیر کارها دارد. در مقام قیاس، صنعت فیلم جوایز و سنجه‌های عددی گوناگونی همچون فروش/درآمد حاصل از بلیط دارد. صنعت نشر هم جوایز و افتخارات و نیز قرار گرفتن در فهرست پرفروش‌ها و آمار و ارقام کل فروش جلد گالینگور، جلد کاغذی، نسخه صوتی و دیگر قالب‌ها دارد. به‌همین شکل جوایز معتبر کمک می‌کنند کیفیت تحقیقات را در بالاترین سطح در نظر بگیریم ولی ما به سنجه‌هایی نیاز داریم که در طیف وسیع‌تری قابل استفاده باشند.

این مقاله کوتاه جامع نیست. شاید آن‌را بعداً به‌شکل رساله جامعی حاوی جزئیات بیشتر در مورد عناوین بحث شده و گردآوری بهترین تجارب توسعه دهم. در حال حاضر ملاحظات مختصری در مورد پنج عنوان ارائه می‌دهم: سنجش کیفیت تحقیقات، سنجه‌های استناد، کیفیت کنفرانس‌ها، ضریب تاثیر نشریه و نشریه‌های جعلی.

۲. سنجش کیفیت تحقیقات

تنها راه مطمئن برای سنجش کیفیت و تاثیر یک مقاله، پروژه یا برنامه تحقیقی اینست که بدهیم دیگر متخصصان دقیقاً در همان رشته آن‌را ارزیابی کنند. متأسفانه این کار به‌شدت غیرعملی است. اغلب موسسه‌ای که سنجش در آنجا باید برای ترفیع یا ارتقاء انجام شود، فاقد افراد دیگری است که دقیقاً در همان رشته تحقیقی باشند. حتی وقتی چنین همکاری هم وجود داشته باشد، ارزیابی درون‌سازمانی اغلب کار ناپسندی محسوب می‌شود. پیدا کردن داوران خارجی که حاضر به ارزیابی باشند هم به‌دلیل برنامه پرمشغله محققان برجسته بسیار دشوار است. نوشتن پیشنهادها در یافت کمک هزینه تحقیقاتی و کشمکش در محیط‌هایی که سیاست انتشار اعتباری^۱ دارند، وقت کمی برای خدمات اجتماعی باقی می‌گذارد.

در شرایطی که آن قدر محققان باسابقه گرفتارند که نمی‌توانند دعوت به داوری یا سنجش سایر تحقیقات را بپذیرند، دانشگاه‌ها و نشریه‌ها مجبور می‌شوند سراغ ارزیاب‌های کم‌رتبه‌تر یا کم‌صلاحیت‌تر بروند. حاصل این کار گاهی فاجعه‌بار یا شرم‌آور است. گاهی من وقتی مشغول خواندن یک داوری یا ارزیابی هستم که پر از خطاهای زبانی و استدلال‌های کاذب منطقی است، از خجالت خودم را جمع می‌کنم. اثر داوری بی‌کیفیت در محیط‌های رقابتی تشدید هم می‌شود و به خروجی‌های ضعیف منتج می‌گردد.

به این دلیل است که به محققان جوان باید شیوه ارزیابی صحیح یک مقاله تحقیقی را آموزش دهیم (شکل ۱ را ببینید). توصیه‌نامه‌ها در تصمیم‌های کلیدی استخدامی و ترفیع در محیط‌های دانشگاهی نقش برجسته‌ای دارند و به‌همین دلیل دانشجویان تحصیلات تکمیلی باید در بخشی از آموزششان با چنین فرایندی از طریق مثال آشنا شوند.

۱- منظور اینست که از استاد دانشگاه انتظار دارند تعداد زیادی مقالات علمی و کتاب منتشر کند و در غیر این صورت از اعتبار و شأن استاد کم می‌شود و تأثیر بدی بر آینده حرفه‌ای او خواهد گذاشت. م.

اسکالر^۵ (شیمی)، سایک اینفو^۶ و پاب مد^۷ هم وجود دارند. علاوه بر استفاده در سنجش کیفیت یک مقاله یا بخشی از تحقیقات، از شمار استنادها می‌توان برای سنجش تاثیر یک محقق، گروه تحقیقاتی یا یک سازمان مثل دانشگاه یا مرکز تحقیقاتی هم استفاده کرد. برای یک محقق بخصوص می‌توان آثار منتشره او را برحسب شمار استنادها به ترتیب نزولی مرتب کرد. فرض کنید فهرست زیر آثار محقق X باشد، به طوری که یک حرف اسم یک مقاله باشد و پس از آن تعداد استنادهای آن مقاله آمده باشد:

X: A 150; B 122; C 70; D 56; E 26; F 8; G 6; H 3 ...

این محقق بخصوص چند مقاله پراستناد دارد (آنهایی که بیش از ۱۰۰ استناد دارند) ولی تعداد استنادها بعد از چند مورد اول، افت می‌کند. اکنون محقق Y را در نظر بگیرید که سابقه استنادهای او چنین باشد:

Y: P 45; Q 41; R 36; S 33; T 30; U 28; V 25; W 23 ...

اگرچه بیشینه تعداد استنادهای محقق دوم از محقق اول کمتر است، تاثیر کلی آن به طور شهودی بالاتر است چون بین تعداد بیشتری از مقاله‌ها پخش شده است.

این تاثیر بیشتر را معمولاً با شاخص h اندازه می‌گیرند که تعریف آن چنین است: h اثر منتشره پراستنادتر که هر کدام دست کم h استناد داشته باشد (شکل ۲). برای محقق X که در مثال اشاره کردیم، شاخص h ۶ است چون ۶ اثر منتشره اول او ۶ استناد یا بیشتر دارند (و ۷ نیست چون ۷ اثر منتشره اول او هر کدام ۷ استناد یا بیشتر ندارند). شاخص h برای محقق Y دست کم ۸ است، ما مطمئن نیستیم چون ۸ اثر منتشره پراستناد او هر کدام دست کم ۲۳ استناد دارند و اگر این الگو با استنادهایی مثل ۲۰، ۱۸ و غیره ادامه یابد، شاخص h او می‌تواند بیشتر هم باشد.

گوگل اسکالر هم شاخص h عمومی محاسبه می‌کند و هم شاخص h «اخیر» برای آثار منتشره در ۵ سال گذشته. عدد دوم معمولاً خیلی کوچک‌تر است چون هم بازه زمانی کوتاه است و هم زمان سپری شده از لحظه انتشار کم است، با این حال این عدد شاخص تاثیر کارهای اخیر محقق است. در مورد یک محقق با سابقه شاخص h بالا ممکن است استنادهای کارهای اولیه او که دهه‌ها قبل منتشر شده‌اند را نشان دهد. شاخص h اخیر بالا نشانی از رشد و تعالی مستمر است از این نظر که کارهای محقق پیوسته خیلی مورد استناد قرار گرفته‌اند. سنجه سومی که گوگل اسکالر محاسبه می‌کند شاخص آی‌تن^۸ محقق است، تعداد آثار منتشره‌ای که دست کم ۱۰ استناد داشته باشند. عدد ۱۰ آستانه اختیاری برای تمایز بخشی بین آثار منتشره کم‌تاثیر و پرتاثیر است. اکثریت مقاله‌های منتشر شده شمار استنادهای تجمعی تک‌رقمی دارند بنابراین اگر محقق شاخص آی‌تن ۵۰ داشته باشد،

ب و ج داده‌ها را تحلیل کردند. الف پیش‌نویس اولیه مقاله را نوشت و بعد از گرفتن نظرات دیگران آن را اصلاح کرد.»

این طرح جزئیات بیشتری دارد اما بیان عددی آن دشوارتر است. درخواست درصد مشارکت، سنجش را ساده‌تر می‌کند اما مانند سایر شاخص‌های عددی مشخص کردن سهم علمی تنها با یک شاخص عددی معایب خاص خودش را دارد.

مشخص کردن سهم علمی هر نویسنده خواه کیفی یا کمی جوابی به سوال کیفیت تحقیقات نمی‌دهد. به‌عنوان جایگزینی برای سنجش مستقیم کیفیت بخشی از تحقیق، می‌توان از سنجه‌های غیرمستقیم استفاده کرد. استناد و میزان شهرت محل انتشار، کمیته‌های جانشین ممکن هستند. در سه بخش بعدی در مورد استفاده از استناد، کیفیت کنفرانس‌ها و شهرت نشریه‌ها صحبت می‌کنم. به‌خاطر تنوع گسترده در کیفیت داوری هر نوع کمیته تجمعی برای کیفیت یا تاثیر نادقیق خواهد بود. با این وجود، در غیاب گزینه‌های بهتر از این کمیته‌های تجمعی به‌طور گسترده‌ای استفاده می‌شود.

۳. سنجه‌های استناد

یکی از شاخص‌های تاثیر یک مقاله تعداد دفعاتی است توسط دیگر محققان استناد می‌شود. می‌گوییم «توسط دیگر محققان» چون استنادهای به کار خود اغلب در تحلیل استناد حذف می‌شوند. تعداد استنادها باید به‌دقت و با احتیاط در ارزیابی تاثیر تحقیق مورد استفاده قرار گیرد. بعضی از موارد احتیاطی عبارتند از:

- فراوانی استنادها بین رشته‌ها/گرایش‌ها متغیر است
- کتاب‌ها معمولاً بیشتر از مقاله‌ها مورد استناد قرار می‌گیرند
- مقاله‌های مروری و آموزشی اغلب بیشتر استناد می‌شوند
- مقالات تخصصی اغلب استنادهای کمتری می‌گیرند
- موضوعات داغ/مد روز ممکن است شمار استنادهایی داشته باشند که مستحق آن نیستند
- به محققان مشهور ممکن است بیشتر استناد شود
- آثار منتشره اخیر استنادهای کمتری دارند

به فهرست ملاحظات بالا باید امکان تقلب و استفاده نابجا را هم اضافه کنید. محققان ممکن است شمار استنادهای خود را با روش استنادهای متقابل یا چرخشی بالا ببرند.

باوجود معایب اجتناب‌ناپذیر، شمار استنادها سنجه خوبی در سطح کلان است. مقاله‌ای که ۱۰۰ استناد دارد، تقریباً می‌توان مطمئن بود که تاثیر بیشتری از مقاله ۵ استنادی دارد. مشکل وقتی به‌وجود می‌آید که بخواهیم بین دو مقاله مثلاً با ۱۰ و ۱۵ استناد تمایز قائل شویم. برای گرفتن داده‌های استنادی، گوگل اسکالر منبع جامع و در دسترس شده است. تارنمای علم^۳ و اسکوپوس^۴ دو پایگاه داده گسترده دیگر هستند. پایگاه‌های مختص هر رشته مثل سای‌فایندر

5- SciFinder Scholar

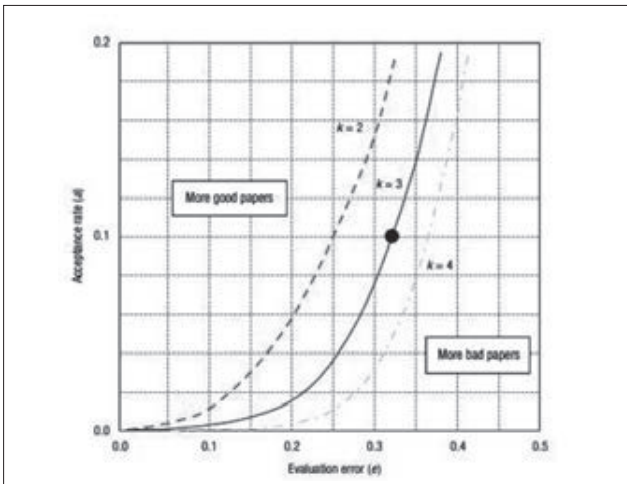
6- PsychInfo

7- PubMed

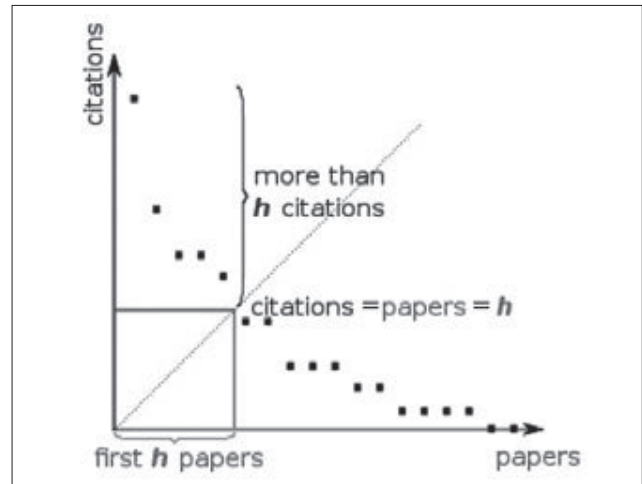
8- i10-index

3- Web of Science

4- SciVerse Scopus



شکل ۳: تصویر ترسیمی آثار زیان‌بخش احتمالی نرخ پذیرش پایین



شکل ۴: تصویر ترسیمی شاخص اچ (ویکی‌پدیا)

مقالات چنین کنفرانس‌هایی منتشر می‌شود، هرگز در نشریه‌ای چاپ نشود. مجموعه مقالات این کنفرانس‌ها همچون نشریه‌های عالی‌رتبه نمایه‌گذاری شده و در بسیاری جاها در دسترسند.

من جای دیگری در مورد خطرات انکاء بر آثار منتشره کنفرانسی به‌عنوان سوابق بایگانی‌شده تحقیقات نوشته بودم. اول این که نرخ پذیرش کم مترادف با کیفیت بالای مقاله‌های پذیرش شده نیست. به‌خاطر وجود مهلت‌های زمانی غیرعملی، داوران به‌جای کیفیت آن مقاله نوشته شده، به شهرت نویسنده بیش از حد اهمیت می‌دهند. با گذشت زمان این کنفرانس‌ها حلقه بسته‌ای از نویسندگان و مشارکت‌کنندگان را جمع می‌کنند و برای تازه‌واردان نوآوری و شرکت در مدیریت فنی کنفرانس دشوار می‌گردد.

دیگر پیامد جنبی مهلت زمانی‌های سخت و سخت کنفرانس‌ها اینست که نویسندگان فرصتی پیدا نمی‌کنند که در مقابل ارزیابی‌های داوران پاسخ‌های متقابل دهند و کارهایشان را اصلاح کنند. بجز معدودی استثناء، کنفرانس‌ها به‌طور یک‌طرفه و بدون مهلت اصلاح، مقاله را رد/قبول می‌کنند. در مورد نشریه‌ها چندین دور مباحثه بین نویسندگان و داوران گمنام وجود دارد که باعث بهبود کیفیت می‌شود.

در شکل ۳ تاثیر زیان‌بخش نرخ پذیرش پایین را در مواجهه با ارزیابی‌های خطاخیز داوران ملاحظه می‌کنیم که با وجود مهلت زمانی سفت و سخت کنفرانس‌ها اجتناب‌ناپذیر است. وقتی $k=3$ داور هر مقاله را ارزیابی می‌کند و هر کدام مثلاً ۳۲ درصد نرخ خطا داشته باشد، از نقطه سیاه درشت در شکل ۳ متوجه می‌شویم که با نرخ پذیرش ۱۰ درصد، تعداد مقاله‌های خوب و بدی که پذیرفته می‌شوند برابر خواهد بود. افزایش نرخ پذیرش به مثلاً ۲۰ درصد نسبت مقاله‌های خوب پذیرفته شده را افزایش می‌دهد! برای تعریف دقیق مقاله‌های خوب و بد به مقاله پرهامی در ۲۰۱۶ مراجعه کنید. حرکت علم و مهندسی کامپیوتر به سمت مدل انتشار کنفرانس‌محور به ماهیت حرکت سریع این رشته برمی‌گردد که خواهان انتشار

نشانگر تاثیر سلامت در شمار عمده‌ای از آثار منتشره است. به‌طور خلاصه، شاخص اچ مثل ۲۵ که گوگل اسکالر محاسبه کرده نشان‌دهنده تاثیر قابل توجه است چون دلالت دارد بر این که محقق ۲۵ اثر منتشره دارد که ۲۵ استناد یا بیشتر کسب کرده است. شاخص‌هایی که گفتیم را می‌توان برای گروهی از محققان در دانشکده، دانشگاه یا غیره نیز تعریف کرد هر چند چندان رایج نیست.

۴. کیفیت کنفرانس‌ها

در بسیاری از رشته‌های علمی و فنی دو محل اصلی برای انتشار نتایج تحقیقی وجود دارد: کنفرانس‌ها (واژه‌ای فراگیر که کنگره، مجمع، سمپوزیوم و کارگاه را نیز پوشش می‌دهد) و نشریه‌ها. تمایزبخشی و وزن‌دهی مناسب آثار منتشره یک محقق در کنفرانس‌ها و نشریه‌ها در بعضی از رشته‌ها منشأ بسیاری از کشمکش‌ها و مشکلات بوده است. در رشته‌های علمی باسابقه، این دو نوع محل انتشار مکمل هم هستند و رقیب هم نیستند. کنفرانس‌ها را به‌شکل محلی برای گزارش تحقیقات در جریان یا خیلی جدید در نظر می‌گیرند، که به‌محضی که آن کار با تحقیقات بیشتر پالایش شد و تحت داوری هم‌تا قرار گرفت، در نشریه مناسبی چاپ می‌شود. در چنین رشته‌هایی کنفرانس‌ها هیچ نوع داوری هم‌تا ندارند و آن‌را به‌عنوان محلی برای برقراری ارتباط و معاشرت با بزرگان یک رشته در نظر می‌گیرند. غیرعادی نیست که صدها ارائه و/یا پوستر در چندین نشست موازی وجود داشته باشد و هر شرکت‌کننده در حداکثر دو سخنرانی که با علاقه‌های او بیشترین انطباق را دارد شرکت کند.

علم و مهندسی کامپیوتر یکی از معدود استثنائات در برنامه‌هایی است که بالا توضیح دادیم. چندین کنفرانس کامپیوتر هست که فرایندهای سخت‌گیرانه داوری هم‌تا (دست‌کم از جنبه نظری) و نرخ پذیرش خیلی کم دارند. خیلی معمول است که مقاله‌ای که مجموعه

از محاسبه ضریب تاثیر سال بعدی، استناد داشته باشند. من به عنوان سردبیر چندین نشریه فنی در چند دهه گذشته، متوجه شده‌ام که بعضی از داوران اصرار دارند نویسندگان به مقاله‌های آن‌ها استناد کنند و این را شرط پذیرش مقاله قلمداد می‌کنند، حتی اگر کارشان در حد خیلی کمی به مقاله تحت داوری مرتبط باشد. معلوم شده که بعضی از سردبیران نشریه‌ها از نویسندگان می‌خواهند به مقاله‌هایی در شماره‌های اخیر نشریه آن‌ها استناد کنند تا ضریب تاثیرش را بالا ببرند.

مثل هر سنجه عددی منفرد، ضریب تاثیر نشریه هم اشکال‌هایی دارد اما اگر با احتیاط مورد استفاده قرار بگیرد، ارزشمند است. در تدریس محاسبات موازی من اغلب به دانشجویان یادآوری می‌کنم که کمیت‌های عملکردی عددی (مثل حداکثر فلاپس^{۱۱}): عملیات ممیز شناور در ثانیه یا فلاپس پایدار) برای مقایسه سریع آبر کامپیوترها قابل قبولند، اما باید آن‌ها را با کمیت‌های دیگری نیز تقویت کنیم تا قضاوت سطحی نداشته باشیم. در مورد نشریه‌ها، ضریب تاثیر را باید با معیارهای کیفی همچون قرار گرفتن در ردیف نشریه‌های گوناگون در یک رشته خاص علمی یا فنی تقویت کنیم.

گونه تقریبی از این رویکرد تقسیم نشریه‌ها به دو دسته خوش‌نام و نامعلوم است. قرار گرفتن در پایگاه داده نمایه‌گذاری مثل موسسه اطلاعات علمی اغلب نشانه اعتبار نشریه است. اتکاء بر قرارگیری در موسسه اطلاعات علمی به عنوان معیار دودویی کیفیت نشریه در بسیاری از کشورهای جهان سوم رواج دارد که هیچ روش و رویه بایات تحقیقی برای خودشان ندارند.

نشریه‌های باکیفیت نه تنها در پذیرش مقاله‌ها گزینشی عمل می‌کنند، در پایش واکنش‌ها به کارهای منتشرشده‌شان نیز بسیار حساس‌اند تا خطاهای رفتارهای تحقیقی غیراخلاقی را پیدا کنند. برای خطاهای پیدا شده چاپ نسخه‌های تصحیح شده در شماره بعدی ضرورت پیدا می‌کند. رفتارهای علمی غیراخلاقی وقتی بعد از شکایت اولیه و از طریق بازبینی جامع و منصفانه به ثبوت می‌رسند، منجر به استرداد کار منتشر شده و اقدامات انضباطی برای نویسندگان خواهد شد. شاخص استرداد یک نشریه به این صورت تعریف می‌شود: تعداد مقاله‌های استرداد شده ضرب در هزار، تقسیم بر مقاله‌های منتشرشده. استرداد مقاله جریمه نهایی نویسنده است، بنابراین چنین اقداماتی بازدارنده قوی برای داده‌سازی یا دیگر انواع تقلب در گزارش تحقیقات فراهم می‌آورد. از طرف دیگر، همبستگی قوی کیفیت نشریه، که با ضریب تاثیر مشخص می‌شود، با شاخص استرداد آن (شکل ۴) نشان می‌دهد نویسندگان ممکن است مایل باشند بیشتر ریسک کنند تا مقاله‌هایشان را در نشریه‌های پرتاثير چاپ کنند.

استدلال جالبی برای استفاده از داوری پس از انتشار به جای داوری پیش از انتشار وجود دارد. فرض کنید محققان هر تعداد مقاله می‌خواهند منتشر کنند، بخصوص حالا که خودانتشاری برخط

به‌موقع و سریع است. طرفداران استفاده از کنفرانس به‌عنوان محل اصلی انتشار گفته بودند که زمان پاسخ نشریه‌ها در محدوده سال است تا این که هفته‌ها یا ماه‌ها باشد و این موضوع باعث می‌شود آثار منتشره بایگانی شده کهنه شده یا اعتبارشان را به‌خاطر نتایج جدید از دست بدهند.

با نگاه به گذشته، استفاده از کنفرانس به‌عنوان محل اصلی انتشار در علم و مهندسی کامپیوتر راه‌حل اشتباهی برای مشکل زمان پاسخ طولانی انتشار تحقیقات بود. راه بهتر به نظر این بود که کاری کنیم نشریه‌ها در مقابل درخواست‌های این رشته پرشتاب بهتر جواب‌گو باشند. جالب توجه است که بعضی از ممتازترین نشریه‌های علمی مثل ساینس و نیچر زمان پاسخی در حدود چند روز تا چند هفته دارند. خوشبختانه در نشریه‌های علم و مهندسی کامپیوتر حرکت به سمت کاهش زمان پاسخ در جریان است. محرک‌های بیشتری برای داوران لازم است تا داوری‌هایشان را به‌موقع بفرستند.

۵. ضریب تاثیر نشریه

سنجه ضریب تاثیر توسط موسسه اطلاعات علمی^۹ گارفیلد پیشنهاد شد تا کیفیت نشریه‌های علمی و تحقیقاتی که در آن منتشر می‌شود را اندازه‌گیری کند. در سال ۱۹۶۴ موسسه اطلاعات علمی گارفیلد اولین شاخص استناد علم^{۱۰} را با پوشش ۲۲۰۰ نشریه منتشر کرد. هشت سال بعد گارفیلد اولین مجموعه از ضریب تاثیر نشریه را در مقاله دیگری منتشر نمود.

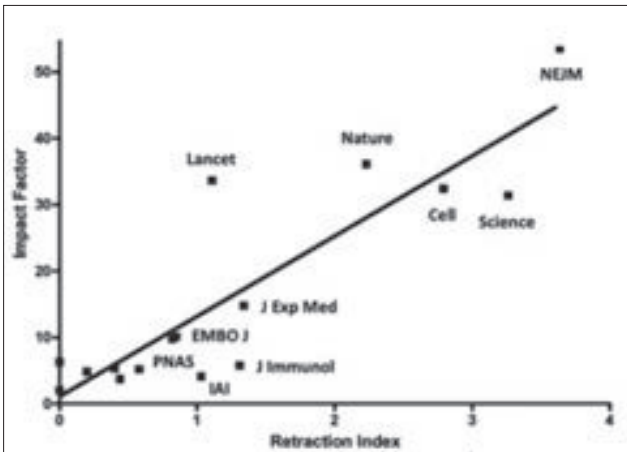
ضریب تاثیر نشریه تعداد کل استنادهایی است که در دو سال گذشته به مقاله‌های آن نشریه می‌شود، تقسیم بر کل تعداد مقاله‌هایی که آن دو سال، نشریه منتشر کرده است. ضریب تاثیر نشریه ممتاز نیچر برای سال ۲۰۱۷ و براساس داده‌های سال ۲۰۱۶ و ۲۰۱۵ عبارتست از: $(41,701 + 32,389) / (880 + 902)$ که مقدار ۴۱/۶ می‌شود. این مقدار عدد تاثیرگذاری است چون ضریب تاثیر نشریه‌ها تخصصی خیلی کم و گاهی حتی زیر ۱ است.

تعریف و استفاده از ضریب تاثیر نشریه ایده درخشانی است، اما مثل هر چیزی در این دنیا در معرض سوءاستفاده و دست‌کاری توسط کلاه‌برداران است (بهینه‌سازی وب را در نظر بگیرید که به‌طور مصنوعی می‌خواهند رتبه بالایی در جست‌وجوهای گوگل به‌دست آورند)، به‌ویژه چون پیشرفت و ترفیع دانشگاهی محققان و حقوقشان به آن وابسته است. حتی اگر این مشکل که مقاله‌ای به‌خاطر انتقادهای گسترده (و نه ارزش‌گذاری) خیلی مورد استناد قرار می‌گیرد، استناد کردن به یک مقاله به‌خاطر لطف به نویسندگان یا محل انتشار آن چیز غیرمتعارفی نیست. بعضی از سردبیران نشریه‌ها مقاله‌هایی چاپ می‌کنند که به احتمال زیاد خیلی زود هنگام استنادهایی در یک دوره/سال می‌گیرند تا آن مقاله‌ها فرصت بیشتری داشته باشند تا قبل

11- Floating Point Operation per Second (FLOPS)

9- Institute for Scientific Information (ISI)

10- Science Citation Index (SCI)



شکل ۴: همبستگی بین ضریب تاثیر نشریه و شاخص استرداد

دسترسی باز گذار مهم و مثبتی است، اگرچه هنوز مشکلاتی هست که باید مرتفع شوند. وجه منفی آن نیز اهمیت دارد. اگر هر نویسنده مثلاً ۱۰۰۰ دلار برای انتشار مقاله‌اش بپردازد، ناشر دیگر خیلی انگیزه‌ای ندارد که مقاله‌ها را سخت‌گیرانه ارزیابی کند و فقط ۱۵ درصد از مقاله‌ها را منتشر کند. از این رو نرخ پذیرش نشریه به تدریج بالا می‌رود تا درآمد را افزایش دهد.

عیب دیگر انتشار با دسترسی باز اینست که به سمت محققانی متمایل می‌شود که کمک هزینه بیشتر دارند و موسساتی که وضع مالی خوبی دارند چون هزینه‌های ارسال و انتشار گاهی به‌طور سرسام‌آوری بالاست. بسیاری از نشریه‌های دسترسی باز تخفیف‌هایی برحسب نیاز می‌دهند اما این‌ها موردی است و به‌طور مداوم تمدید نمی‌شود.

در طول دهه گذشته چندین هزار نشریه مبادرت به درآمدزایی از طریق این روند انتشار دسترسی باز کرده‌اند. بسیاری از این نشریه‌های جعلی دفتری در کشورهای جهان سوم دارند که دستمزد کارها خیلی ارزان است. نشریه (اغلب بخشی از یک مجموعه که صدها عنوان دارد) معمولاً سردبیر بی‌نامی دارد که با ایجاد انگیزه‌های مختلف همچون چاپلوسی یا تخفیف از نویسندگان درخواست مقاله می‌کند. نشریه‌های جعلی در تعریف و تمجید افراطی‌شان از شما خیلی سخاوتمندانه عمل می‌کنند خواه شما شایسته آن باشید یا خیر. آن‌ها به‌طور دوره‌ای دعوت‌نامه‌هایی برای ارسال مقاله، سردبیر مهمان یک شماره ویژه و عضویت در هیئت ویراستاریشان (که خیلی بزرگ هم هست) می‌فرستند. آن‌ها گاهی قول انتشار مقاله در چندین روز می‌دهند، مثلاً درخواست ارسال مقاله را برای شماره ویژه ماه جون در اواخر می می‌فرستند.

دانشگاه‌یانی که فرصت عضویت در هیئت ویراستاری نشریه‌های ممتاز را پیدا نمی‌کنند، از این فرصت استفاده می‌کنند که در فهرست هیئت ویراستاری چندین نشریه قرار بگیرند و شرح سوابق علمی و کاریشان را با آثار منتشره و عضویت در چندین هیئت ویراستاری پر کنند. یک نشان نشریه‌های جعلی، عنوان‌های پر زرق و برق و خیلی

بی‌هزینه است. اگر چنین انتشاری داوری‌های همتای مثبت بعد از قرارگیری در رایاسپهر نگیرد، بی‌ارزش خواهد شد. این رویکرد از جنبه نظری استوار است، اما این که بفهمیم چگونه باید از «آلودگی توضیحی» یا هرز نظرات، از نوع نظراتی که در وب‌نوشت‌ها می‌بینیم، اجتناب کنیم کار ساده‌ای نیست. در اساس استنادها نوعی داوری دودویی بله/خیر پس از انتشار هستند.

۶. نشریه‌های جعلی

تامین مالی نشریه‌های علمی و فنی همیشه مشکل سخت و پیچیده‌ای بوده است. با تکیه بر فعالیت‌های داوطلبانه (بی‌پرداخت) داوران و سردبیران، جوامع علمی در طی سال‌ها ترتیبی داده‌اند که نشریه‌هایی با کیفیت نسبتاً بالا منتشر کنند. اما این کار رفته‌رفته دشوارتر شده است. هم‌اکنون یافتن چندین داور برای یک مقاله ارسالی کار بسیار پرچالشی است چون برنامه‌های زمانی محققان بسیار پر است و نشریه‌ها و کنفرانس‌ها نیز خیلی زیاد شده‌اند.

زمانی که نشریه‌ها در قالب کاغذی منتشر می‌شدند، کتابخانه‌های تحقیقی قفسه‌های متعددی از شماره‌های گذشته هر نشریه داشتند که به‌طور منظم به کتاب‌های چندجلدی اختصاص داشت. ازدیاد نشریه‌ها باعث شد مشترک شدن در همه نشریه‌های مرتبط غیرممکن شود، به‌همین دلیل کتابخانه‌ها خیلی گزینشی عمل می‌کردند و برای تامین نیازهای مشتریان‌شان به همکاری بین قرض دادن بین کتابخانه‌های متکی شده بودند.

امروزه که نشریه‌ها به‌سمت قالب‌های الکترونیکی و انتشار برخط روی آورده‌اند (و نسخه‌های کاغذی فقط برای آن‌هایی که این گذار برایشان دشوار است توزیع می‌شود)، مشکل دیگری پدید آمده است. هزینه نگهداری بایگانی نشریه‌ها را چه کسی می‌پردازد تا همیشه در دسترس بمانند؟ این هزینه‌ها بسیار زیادند و با رشد مجموعه مقالات، بایگانی سازمان‌یافته و قابل دسترس چالشی می‌شود. در نشریه‌هایی که توسط جوامع حرفه‌ای و نیز ناشران خصوصی منتشر می‌شود، نگرانی اینست که اگر احیاناً متولی که آن‌را نگهداری می‌کند غیرفعال یا ورشکست شود، بایگانی‌ها ممکن است ناپدید شوند.

انتشار با دسترسی باز برای این که تاحدودی مشکل فوق را حل کند، به‌وجود آمده است. مدل‌های قدیمی انتشار مقاله‌ها را بی‌هزینه برای نویسندگان چاپ می‌کنند اما دسترسی به نتایج تحقیقات نیاز به پرداخت دارد (از طریق مشترک شدن در نشریه یا عضویت در اجتماعات حرفه‌ای). انتشار با دسترسی باز این مدل را معکوس می‌کند. نویسنده یا موسسه او هزینه‌های انتشار و بایگانی را می‌پردازد ولی این کار به‌طور رایگان برای همگان در دسترس قرار می‌گیرد. مقاله‌های دسترسی باز گرایش به خوانده شدن توسط افراد بیشتر دارند و به این وسیله تعداد استناد به آن‌ها افزایش می‌یابد.

به‌خاطر دسترسی آزاد به نتایج تحقیقات، حرکت به سمت انتشار با

یعنی راه درازی در پیش داریم تا بتوانیم در اجرای ارزیابی‌های منصفانه، دقت کافی به خرج دهیم.

وجود خطا در سنجش کیفیت و تاثیر تحقیقات ایجاب می‌کند که انعطاف‌پذیر باشیم و کورکورانه از راهنماهای سفت و سخت پیروی نکنیم، فارغ از این‌که با چه دقتی تهیه شده‌اند و چه میزانی از جزئیات ارائه می‌کنند. لوبیاشماری هنوز همان لوبیاشمار است^{۱۲}، حتی اگر لوبیاها از نظر رنگ و اندازه مرتب شده باشد!

زمینه‌های تحقیقی چند رشته‌ای مشکلات ددرسازاری در سنجش کیفیت و تاثیر به‌وجود می‌آورند. ترکیب دانش و روش‌ها از چندین رشته ذاتاً دشوار است و ارزش پاداش‌دهی دارد. از طرف دیگر، من نمونه‌هایی از نتایج و روش‌های بدیهی در یک رشته دیده‌ام که برای افراد رشته دیگر که آماده ارزیابی آن ادعاها نیستند، شگفت‌آور و نوآورانه قلمداد شده است. به بیان دیگر، کارهای چندرشته‌ای باید در هر یک از رشته‌های دخیل نیز تازه و نوآورانه باشند.

من این مقاله را نقطه شروعی می‌بینم. آن‌را می‌توان از جهت‌های مختلف توسعه و بسط داد. اگر از این کار بخواهیم برای آموزش نسل آینده محققان و منبعی برای ترفیعات دانشگاهی و کمیته‌های ارتقاء استفاده کنیم، جزئیات بیشتری لازم می‌شود. من در این جهت مشغول کار هستم و از هر گونه نظر، توصیه مفید یا انتقادی که باعث بهبود عمق و پوشش این کار شود قدردانی می‌کنم.

منبع

1. On Research Quality and Impact: What Five Decades in Academia Has Taught Me
https://www.ece.ucsb.edu/~parhami/pubs_folder/f50-m-research-quality-and-impact-200609.pdf

۱۲- منظور از لوبیاشمار فردی است که مثل حسابداران فقط به جزئیات توجه دارد و کلیات را از یاد می‌برد.

عمومی آن‌هاست که تقریباً شامل هر چیزی می‌شود. من این عنوان را برای خنده درست کردم ولی عنوان‌های واقعی بهتر از این نیستند: «نشریه جهانی مهندسی، علم، مدیریت و تحقیقات اجتماعی خلاق و پیشرفته»

فهرستی از نشریه‌ها و ناشران شکارچی گردآوری شده‌اند. بسیاری از چنین فهرست‌هایی از واژه «بالقوه» استفاده می‌کنند تا هم خودشان را از مسایل قانونی حفاظت کنند و نیز در هر فهرست بلند و بالایی احتمال بی‌دقتی وجود دارد که در طی زمان درست می‌شود.

همچنان‌که تعداد نشریه‌های قانونی و شکارچی بیشتر می‌شود، تمیز دادن آن‌ها از هم نیز دشوارتر می‌گردد. اگرچه نشریه‌های جعلی برای پول درآوردن تقریباً هر چیزی را منتشر می‌کنند، بعضی از نشریه‌های پژوهشی قانونی خیلی بهتر عمل نمی‌کنند. اما انگیزه این دسته دوم مالی نیست، بلکه ترویج دیدگاه‌ها و طرز تفکرهاست. مثال‌هایی از مقاله‌های هجوآمیز یا تولیدشده توسط کامپیوتر هست که در نشریه‌های با داورى همتا منتشر شده‌اند شاید چون داوران به‌شدت گرفتار یا بی‌دقت متوجه آن نشده‌اند.

۷. نتیجه‌گیری

در این مقاله من برخی از نظرات و نکات مفید در مورد چالش‌ها و مخاطرات ارزیابی کیفیت و تاثیر تحقیقات را با شما در میان گذاشتم. دوپهلوی حرف زدن‌های من را نباید این‌طور برداشت کنید که چنین ارزیابی‌هایی بی‌بهره است. در عوض به آن‌ها به این چشم نگاه کنید که احتیاط بیشتری لازم است و متوجه باشید که این فرایند خطاخیزی است. ما نمی‌توانیم بگوییم نتایج ارزیابی‌های تحقیقی به‌کلی خطادارند، اما همین آگاهی از این‌که خروجی‌ها دقیق نیستند

"تراوش‌های ذهنی"

۲۵ شیوه نگرش به هوش مصنوعی

از انتشارات انجمن انفورماتیک ایران

ترجمه ابراهیم نقیب‌زاده مشایخ
به زودی منتشر می‌شود!

شرکت‌هایی که مایل به حمایت مالی از چاپ این کتاب هستند

با دفتر انجمن (۶۶۴۱۲۸۶۱) تماس بگیرند