



Identifying and Categorizing the Indicators Used to Evaluate Scientific Outputs in the Scholarly Publication Ecosystem


Afroz Hamrahi ^{1*}

Roya Pournaghi ²

Dariush Matlabi ³

 1. Ph.D. in Knowledge and Information Science, Iranian Research Institution for Information Science and Technology (IranDoc), Tehran, Iran (Corresponding Author)

 2. Ph.D. in Knowledge and Information Science; Assistant Professor; Iranian Research Institution for Information Science and Technology (IranDoc), Tehran, Iran.
Email: pournaghi@irandoc.ac.ir

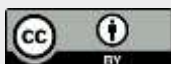
 3. Associate professor. department of educational science, Yadegare Imam Khomeini (RAH) Shahre rey branch, Islamic Azad University, Tehran, Iran.
Email: dariushmatlabi@iau.ac.ir

Email: Hamrahi@students.irandoc.ac.ir

Abstract

Date of Reception:
06/12/2023

Date of Acceptation:
08/04/2024



Purpose: Science and technology are the most critical infrastructures for the country's progress and essential tools for competition in various fields. Evaluation is at the core of all scientific efforts, which has become increasingly important with the proliferation of scientific publications. Evaluation is not a simple and transparent process; it is considered a delicate activity. The presence of multiple evaluation indicators to assess the value of scientific outputs in texts, databases, and scientific centers or publications has prompted a study of these three sources (texts, scientific networks, and experts) to establish integrated criteria for evaluating outputs in the scholarly publication ecosystem. Some scientific outputs, such as lectures, workshops, and scientific meetings, are often overlooked due to the absence of an integrated framework in the scholarly publication ecosystem. Additionally, only a few specific quantitative aspects, such as the impact factor, the number of citations, or the number of uses, are typically considered in the evaluation of scientific works. Moreover, evaluations are usually confined to a short timeframe. Identifying and categorizing these indicators within a framework can positively impact addressing these issues and establishing a continuous evaluation process for both pre and post-publication of scientific works. Therefore, the current research aims to identify comprehensive evaluation criteria in the scholarly publication ecosystem by taking into account texts, scholarly publication networks, and the perspectives of scientific publication experts.

Methodology: Results from all three methods indicate that there are three key indicators and nine sub-indicators crucial for evaluating research outcomes within the Scientific Publishing System. These indicators are predominantly categorized based on form, type, and format. The triangulation method has introduced a conceptual framework of evaluation indicators in the scholarly publication ecosystem. Initially, a systematic review was conducted to extract evaluation criteria from 331 sources. Subsequently, to validate the extracted criteria and finalize the initial framework, the identified criteria were scrutinized across 12 scientific databases, and ultimately, it was endorsed by 30 domestic and in-

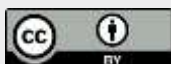
Afroz Hamrahi ^{1*}

Roya Pournaghi ²

Dariush Matlabi ³

Date of Reception:
06/12/2023

Date of Acceptation:
08/04/2024



ternational scholarly publication experts. The purposive sampling method was employed in all three studies.


Findings: Research shows that the ecosystem of scientific publication consists of various components, including experts, scientific centers, information media, subject areas, information, and knowledge systems, which require different indicators and methods. The data extracted from the systematic review in the field of evaluation were classified into three groups: form, type, and format. Evaluation forms include content, open, altmetric, and bibliographic evaluations (creative and source evaluation). However, some experts distinguish between bibliometric evaluation indicators and scientometric and informatics evaluation indicators, and most experts in different subject areas define all three categories as bibliographic evaluations. In the evaluation form, creators include individuals and scientific organizations such as universities. Open evaluation can refer to judging an output not just by a jury of experts but rather by a jury of anyone interested in the output. In other words, open evaluations are an ongoing post-publication process of transparent peer evaluation. Multiple paper evaluation functions freely defined by individuals or groups provide various perspectives on the scientific literature. Multiple paper evaluation functions alongside more diverse research evaluation criteria beyond traditional methods are emerging, and with these come a range of practical, ethical, and social factors to consider. Altmetric evaluations are a set of methods based on the social web that measure and monitor the reach and impact of scholarly output through online interactions. Simply, altmetrics are metrics beyond traditional citations. This evaluating form measures cite, like, view, reuse, discussion, bookmark, etc. The types of evaluations include quantitative, qualitative, and mixed evaluations. The form of evaluation also includes technical evaluation and non-technical evaluation (researcher-made evaluation, discussion-based evaluation). Technical evaluations are indicators that follow predefined procedures or repetitive processes to reach the result, while experts define non-technical evaluations according to specific situations and conditions.

Conclusion: The results of all three methods indicate that there are three key indicators and nine sub-indicators crucial for evaluating scientific output within the scholarly publication ecosystem. These indicators are predominantly grouped based on form, type, and format. The findings demonstrate alignment among the three studies (systematic review, observation of scholarly publication networks, and survey of experts). However, each study has highlighted specific evaluation indicators. According to the systematic review, observation of scholarly publication networks, and expert opinions, the primary focus in evaluating the scholarly publication ecosystem is on the form and type of evaluation. There is a greater emphasis on common and well-established formats in the evaluation process. Furthermore, individual and organizational needs and objectives significantly influence the selection of evaluation indicators. Categorizing evaluation indicators will assist stakeholders in understanding the evaluation procedures of scholarly publication ecosystems and in choosing appropriate evaluation methods.


Keywords: Scholarly Evaluation, Evaluation Indicators, Scholarly Publication Ecosystem, Scientific Publication, Triangulation.

شناسایی و دسته‌بندی معیارهای ارزیابی برون‌داده‌های علمی در زیست‌بوم نشر علمی


افروز همراهی^{*۱}

۱.  دکتری علم اطلاعات و دانش‌شناسی، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، پژوهشگاه علوم اطلاعات، تهران، ایران (نویسنده مسئول).

رویا پورنقی^۲

۲.  دکتری علم اطلاعات و دانش‌شناسی، استادیار پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، تهران، ایران.

داریوش مطلبی^۳

۳.  دانشیار علوم تربیتی واحد یادگار امام خمینی (ره) شهرری، دانشگاه آزاد اسلامی، تهران، ایران.

Email: pournaghi@irandoc.ac.ir

Email: dariushmatlabi@iausr.ac.ir

Email: Hamrahi@students.irandoc.ac.ir

چکیده

هدف: وجود یک سامانه ارزیابی علمی قدرتمند و قابل اطمینان، برای دریافت بازخورد دقیق و پیوسته از نتایج، پیامدها و آثار برون‌داده‌های علمی، ضروری است. چنین نظامی از گوناگون، ارزش یک برون‌داد علمی را به تصویر می‌کشد. از این رو هدف اصلی پژوهش حاضر شناسایی و دسته‌بندی معیارهای ارزیابی برون‌داده‌های علمی در حوزه زیست‌بوم نشر علمی با در نظر گرفتن متون، شبکه‌های نشر علمی و دیدگاه خبرگان نشر علمی است تا به طراحی نظام ارزیابی برون‌داده‌های علمی کمک نماید.

روش‌شناسی: برای ارائه چارچوب مفهومی از شاخص‌های ارزیابی در زیست‌بوم نشر علمی از روش آمیخته و رویکرد سه‌سویه‌سازی استفاده شده‌است. ابتدا معیارهای استفاده شده در حوزه نشر علمی با به‌کارگیری مرور نظام‌مند از ۳۳۱ منبع استخراج شد. سپس برای تعیین اعتبار معیارهای استخراج‌شده و تکمیل چارچوب اولیه، معیارهای شناسایی‌شده در ۱۲ پایگاه علمی مورد بررسی قرار گرفت و در پایان به تأیید ۳۰ نفر از خبرگان داخلی و خارجی رسید.

یافته‌ها: روی هم‌رفته ۳۳۱ مدرک برای مطالعات مرور نظام‌مند معیارهای ارزیابی در زیست‌بوم نشر علمی مورد تأیید قرار گرفت. داده‌های استخراج‌شده از مرور نظام‌مند در حوزه ارزیابی در ۳ گروه دسته‌بندی شد و شامل شکل، نوع و قالب ارزیابی است. براساس دو مطالعه دیگر، شاخص جدیدی به معیارهای شناسایی‌شده اضافه نشد؛ اما یافته‌ها نشان از تأکید و توجه متخصصان به نوع خاصی از شاخص‌های شناسایی‌شده در حوزه ارزیابی برون‌داده‌های علمی داشت.

نتیجه‌گیری: در مجموع ۳ شاخص اصلی و ۹ شاخص فرعی برای ارزیابی زیست‌بوم نشر علمی شناسایی شد. در شکل ارزیابی، ۴ شاخص فرعی (ارزیابی‌های محتوایی، باز، دگرسنجی و کتاب‌شناختی) در نوع ارزیابی، ۳ شاخص فرعی (ارزیابی کمی، ارزیابی کیفی، ارزیابی آمیخته) و در قالب ارزیابی، ۲ شاخص فرعی (ارزیابی فنی و غیر فنی) شناسایی شد. همچنین، نتایج نشان از هم‌راستایی هر سه مطالعه (مرور نظام‌مند، مشاهده شبکه‌های نشر علمی و نظرسنجی از خبرگان) دارد. هرچند در هر یک از سه مطالعه، بر شاخص‌های خاصی از مؤلفه ارزیابی تأکید دارند. خروجی هر سه مطالعه نشان می‌دهد، آنچه در ارزیابی زیست‌بوم نشر علمی در اولویت است شکل و نوع ارزیابی است. در قالب ارزیابی نیز بیشتر بر روش‌های معمول و شناخته‌شده تأکید شده است. پژوهش حاضر می‌تواند الگوی مناسبی برای طراحی نظام ارزیابی برون‌داده‌های علمی قرار گیرد.

واژگان کلیدی: ارزیابی علمی، برون‌داد علمی، شاخص‌های ارزیابی، زیست‌بوم نشر علمی، نشر علمی، سه‌سویه‌سازی.

صفحه ۲۶-۱

دریافت: ۱۴۰۲/۰۹/۱۵

پذیرش: ۱۴۰۳/۰۱/۲۰

مقدمه و بیان مسئله

نظام علم و فناوری کشور نیازمند طراحی، توسعه، ارتقا و تقویت سامانه یکپارچه و جامع نظارت علمی و فناوری است، به طوری که دولت‌ها بتوانند از قابلیت‌های آن در جهت بررسی پیشرفت، رشد و حرکت در راه شکوفایی علم و فناوری استفاده کنند (آزادی احمدآبادی، ۱۳۹۹). هر یک از مراکز یا واحدهای علمی برای رسیدن به اهداف خویش باید پیوسته وضعیت گذشته و حال خود را پایش نمایند تا با شناخت نقاط قوت و ضعف آن، بتوانند مسیر آینده خود را به درستی ترسیم نمایند (کرامت‌فر و رفیعی خشنود، ۱۳۹۵؛ Seong Leem & Kim, 2004). تجربه جوامع توسعه‌یافته نشان می‌دهد که توسعه متوازن و پایدار در تمامی ابعاد کشور، مستلزم توسعه علم و پژوهش است (قنادی‌نژاد و حیدری، ۱۳۹۹) و یکی از ملزومات پایه در توسعه پژوهش، سنجش و ارزیابی برون‌دادهای علمی است (کرامت‌فر و رفیعی خشنود، ۱۳۹۵). در ایران نیز علم و فناوری از مهم‌ترین زیرساخت‌های پیشرفت کشور و ابزار ضروری برای رقابت در عرصه‌های مختلف به‌شمار می‌آید (جنوی و همکاران، ۱۳۹۹) و به‌کارگیری مقیاس‌های ارزیابی مناسب از اجزای لاینفک آن به حساب می‌آید (خیاطیان یزدی و همکاران، ۱۳۹۹).

ارزیابی را می‌توان فرایندی توصیف کرد که به وسیله آن مردم، ارزشمندی و اعتبار چیزی را قضاوت می‌کنند (Oliver, 2000). به دلیل اینکه ارزیابی محتوای علمی فرایندی ساده و روشن نیست، بدیهی است این فرآیند، کاری ظریف و حساس تلقی شود. ارزیابی به شکل رسمی یا غیر رسمی با مجموعه‌ای از سؤالات و مسائل همانند: چه چیزی باید ارزیابی شود؟ چه زمانی و چرا ارزیابی باید انجام شود؟ چگونه ارزیابی شود؟ چگونه فرآیند ارزیابی، منصفانه و عینی تعریف شود؟ همراه است (Lamanauskas, 2011). از این روی این پژوهش در نظر دارد دسته‌بندی جامعی از معیارهای ارزیابی در زیست‌بوم نشر علمی ارائه دهد تا تصمیم‌گیری در باره فرایند ارزیابی برای متخصصان شفاف‌تر شود. زیرا، جوامع علمی بر پایه داوری و ارزیابی کیفیت آثار علمی شکل می‌گیرد. زیست‌بوم نشر علمی باید فرایندی را برای ارزیابی آثار علمی فراهم نماید که با مکانیسم شفاف کار کند و امکان داوری هر یک از برون‌دادهای علمی را میسر نماید (Gradinarov, 2000).

مسئله ذاتی در نشر علمی این است، دیگر صاحب‌نظران باید بتوانند به ادعاهای منتشر شده پدیدآوران تکیه کنند و در آثار خود از آن استفاده نمایند (Smit & Gruttmeier, 2011). زیرا، تصمیم‌گیری بر پایه اطلاعات علمی نادرست می‌تواند به طیف گسترده‌ای از نتایج منفی منجر شود و نیز بر زندگی مردم، محیط‌زیست، حکمرانی یا سایر حوزه‌های توسعه تأثیر گذارد و خطرات بیشتری را برای محیط پیرامون ایجاد نماید (آزادی احمدآبادی، ۱۳۹۹). تمام تلاش‌های علمی، ارزیابی است که با انفجار نشر علمی به‌ویژه در بحث انتخاب، اهمیت بیشتری پیدا کرده است. بنابراین یک چالش بزرگ جامعه علمی، طراحی نظامی است که با استفاده از آن به تصمیم‌گیری و انتخاب بهتر منابعی که سزاوار توجه گسترده و خواندن عمیق هست، ختم شود (Kriegeskorte et al., 2012). علاوه بر منابع، شبکه‌ها و محمل‌هایی که این برون‌دادهای را عرضه می‌کنند نیز نیازمند ارزیابی هستند. از این رو لازم است در ارزیابی برون‌دادهای علمی زیست‌بوم نشر علمی، به مفهوم پیچیده و گسترده ارزیابی با نگاهی سیستمی نگرینده شود (همراهی، ۱۴۰۱). وجود نظام جامع و یکپارچه ارزیابی، علاوه بر تعیین ارزش واقعی اجزا؛ مزایایی نظیر دسترسی یکپارچه به منابع، توجه به نیاز تمامی ذی‌نفعان، امکان همکاری علاقه‌مندان (Alterman et al., 1984)، سنجش عملکرد گذشته و ارائه خدمات بهتر به مشتریان، اصلاح شیوه‌ها و افزایش شانس موفقیت، سنجش میزان تحقق اهداف، سنجش میزان تأثیر ارزیابی‌های انجام‌شده در رشد زیست‌بوم، پیش‌بینی بودجه، تجزیه و تحلیل الگوی تقاضا

در زیست‌بوم نشر علمی، شناسایی شکاف‌ها و راه‌های ممکن برای پر کردن آن (Neely et al., 2005) را برای جامعه و ذی‌نفعان در پی خواهد داشت که اهمیت نگرش سیستمی به مسئله ارزیابی را بیشتر می‌کند. نلی (Nelly 2005) معتقد است؛ مفهوم ارزیابی به این دلیل اهمیت دارد که اگر نتوان مسئله‌ای را ارزیابی کرد، پس نمی‌توان آن را درک کرد و اگر نتوان درک کرد، نمی‌توان آن را کنترل کرد و بهبود داد. بنابراین، برای درک بهتر مسئله ارزیابی در زیست‌بوم نشر علمی، لازم است ابعاد مختلف ارزیابی در این نظام شناسایی و دسته‌بندی شود. این در حالی است، در نظام‌های ارزیابی کنونی، دریافت بازخورد دقیق و پیوسته از زیست‌بوم نشر علمی میسر نیست و بازخوردهای تقریبی، بیشتر مبتنی بر شاخص‌های عددی است. در حالی که هر برون‌داد علمی از جنبه‌های مختلفی قابل سنجش و ارزیابی است.

افزون بر این، در ارزیابی برون‌دادهای علمی، سنجش میزان تأثیر برون‌دادهای علمی بر تصمیم‌گیری‌ها و برنامه‌ها یک رابطه ساده نیست. زیرا گاهی اوقات مطالعاتی که به نظر حیاتی می‌رسند، نمی‌توانند بر تصمیم‌گیری‌های کوتاه مدت تأثیر بگذارند و مطالعاتی که در ابتدا به نظر می‌رسد هیچ تأثیری ندارند، می‌توانند در صورت بروز شرایط مناسب تر، تأثیر تأخیری ایجاد کنند (Trochim & Donnelly, 2001). بنابراین، شناسایی شاخص‌ها، به شکل‌گیری نظام ارزیابی قدرتمند و قابل اطمینان می‌انجامد که می‌تواند به دریافت بازخورد دقیق و مستمر از نتایج، پیامدها و آثار برون‌دادهای علمی کمک کند و جنبه‌های مختلف، ارزش یک برون‌داد علمی را به تصویر بکشد (شجاعی و همکاران، ۱۳۹۳).

با توجه به مطالب مطرح‌شده، به نظر می‌رسد اگر مؤلفه‌ها و شاخص‌های مختلف ارزیابی در زیست‌بوم نشر علمی شناسایی نشود، درک و شناخت روشنی از مفهوم ارزیابی در این حوزه شکل نخواهد گرفت. نبود چهارچوب یکپارچه از مفهوم ارزیابی در زیست‌بوم نشر علمی، مانع سنجش، کنترل و بهبود درست ارزش برون‌دادهای علمی در این زیست‌بوم خواهد شد. همچنین، یکپارچه نبودن شاخص‌ها ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی و نبود امکان ارزیابی مستمر (پیش از انتشار و پس از انتشار) می‌تواند فرصت سازگاری و اصلاح برون‌داد علمی در طی تکمیل فرآیند نشر را از بین ببرد و احتمال شکل‌گیری خروجی موفق و تأثیرگذار را کاهش دهد.

بنابراین، در این مطالعه مسئله ارزیابی به‌عنوان یک فرآیند مستمر و یکپارچه و بخش مهمی از زیست‌بوم نشر علمی در نظر گرفته شده است تا بتواند امکان سنجش برون‌دادهای علمی، براساس شاخص‌های انبوه در زمان‌های مختلف (پیش از انتشار و پس از انتشار) را میسر نماید. از این‌رو، برای کمک به این مهم و حل این مسئله به شناسایی معیارهای جامع ارزیابی در زیست‌بوم نشر علمی با در نظر گرفتن متون، شبکه‌های نشر علمی و دیدگاه خبرگان نشر علمی نیاز است به‌طوری که با ارائه چارچوب مفهومی از شاخص‌های ارزیابی در زیست‌بوم نشر علمی به‌روش سه سویه‌سازی و با آماده کردن فرایندی برای ارزیابی آثار علمی، بتواند مکانیزمی شفاف که امکان داوری هر یک از برون‌دادهای علمی را آماده کند، در اختیار قرار دهد. به‌عبارت دیگر، پژوهش حاضر در پی پاسخ به این پرسش است که براساس روش سه‌سویه‌سازی، معیارهای جامع ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی کدام است و دسته‌بندی آنها به چه صورت است؟

پرسش‌های پژوهش

۱. براساس مرور نظام‌مند متون، شاخص‌های اصلی ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی کدام است؟
۲. براساس مشاهده شبکه‌های علمی، شاخص‌های اصلی ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی کدام است؟

۳. براساس نظرسنجی از خبرگان نشر علمی، شاخص‌های اصلی ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی کدام است؟

چارچوب نظری

تعاریف مختلفی از زیست‌بوم ارائه شده‌است. برخی زیست‌بوم را تعامل بین موجودیت‌ها و محیط آن‌ها به حساب می‌آورند (Tansley, 1935) و برخی بافتی را در نظر دارند، مجموع جانداران و محیط طبیعی آن‌ها با یکدیگر در تعامل هستند (معین، ۱۳۸۹). اغلب تعاریف زیست‌بوم، دربرگیرنده شبکه‌ای پیچیده و طبیعی از اهداف، نگرش‌ها، محصولات، فرآیندها، بازیگران و روابط آن‌هاست که مهم‌ترین ویژگی آن هماهنگی است؛ به این معنی که هر رفتار و تصمیم منطقی اعضا بر کل بوم یا بافت تأثیر می‌گذارد. میزان هماهنگی، تعیین‌کننده میزان پایداری هر زیست‌بوم است (Tsujiimoto et al., 2018).

نشر علمی نیز از بافتی مجزا برخوردار است و سال‌ها در تعامل بین بازیگران و عوامل دیگر شکل گرفته است و به سبب پویایی و انسجام با عنوان زیست‌بوم نشر علمی شناخته شده‌است (Mangiafico & Smith, 2014). اجزای این زیست‌بوم، طیف گسترده‌ای از متخصصان نشر علمی، سازمان‌ها و مراکز علمی، شبکه‌ها و پایگاه‌های علمی، برون‌دادهای علمی، قالب‌ها و ساختارها، روابط و پیوندهای بین آن‌ها را در برمی‌گیرد (McGuire et al., 2018). برای ارزیابی این عناصر، معیارهای ارزیابی متفاوتی شکل گرفته‌اند که به‌طور کلی به دو دسته، معیارهای ذهنی، نظیر بررسی عمیق محتوا توسط چند متخصص موضوعی و معیارهای عینی، نظیر شمارش استنادها تقسیم می‌شوند (Xu et al., 2013). این پیچیدگی در ساختار و روابط زیست‌بوم نشر علمی و تأثیر آن بر زیست‌بوم‌های دیگر؛ ارزیابی حوزه نشر با رویکردی وسیع‌تر را می‌طلبد. نگرش کلان به مسائل حوزه نشر علمی و به‌ویژه مفهوم ارزیابی، می‌تواند مبنایی برای درک و فهم ساختار و مفاهیم این حوزه فراهم نماید که در سطوح گسترده اثرات مطلوبی بر سیاست‌گذاری‌ها و برنامه‌های کلان خواهد داشت.

مفهوم ارزیابی نیز از جمله مفاهیمی است که همواره در مطالعات گوناگون حائز اهمیت بوده است (شهلانی، ۱۳۹۵؛ نامداریان و کلانتری، ۱۳۹۹؛ وثوق و همکاران، ۱۳۹۹)، و یزدانی و همکاران (Yazdani et al., 2015). ارزیابی علمی به راستی به‌دست آوردن اطلاعات مناسب و سنجش نظام‌مند اطلاعات برای ارائه بازخورد مفید در مورد یک شیء به مخاطبان مختلف از جمله حامیان مالی، اهداکنندگان، گروه‌های مشتری، مدیران، کارکنان و سایر حوزه‌های مرتبط است. منظور از مفید، بازخوردی است که بتواند به تصمیم‌گیری درست کمک کند. همچنین منظور از شیء می‌تواند برنامه، خط‌مشی، فناوری، شخص، نیاز، فعالیت و غیره را در برگیرد (Trochim & Donnelly, 2001).

با در نظر گرفتن اینکه ارزیابی می‌تواند در رسیدن به درک درست از امکانات، فرصت‌ها، نقاط قوت و ضعف هر نظامی کمک کند (Saxena et al., 2013) و از طریق شناخت نارسائی‌ها و ارائه اطلاعات از چگونگی و میزان تحقق اهداف و برنامه‌ها، به ابزار مناسب و اساسی برای تصمیم‌گیرندگان، برنامه‌ریزان و سیاست‌گذاران تبدیل شود (آزادی احمدآبادی، ۱۳۹۹)؛ بنابراین، نخستین گام برای سامان بخشیدن به زیست‌بوم نشر علمی، شناسایی ابعاد مختلف ارزیابی در این زیست‌بوم است. زیرا، یک ارزیابی زمانی نتایج قابل اتکایی فراهم می‌سازد که با استفاده از منابع اطلاعاتی مناسب (Bondi & Bondi, 2003)، به‌طور صحیح و منطقی (Jackson, 2005) و براساس شاخص‌های مناسب، طراحی و اجرا گردد (نیکخواه و همکاران، ۱۳۹۰). تکیه صرف بر یک شاخص سنجش نمی‌تواند ابزار چندان

کارآمدی برای ارزیابی برون‌داد علمی قرار گیرد (داورپناه، ۱۳۸۹؛ عمرانی، ۱۳۸۶). در سال‌های اخیر، پژوهشگران و صاحب‌نظران علم‌سنجی به دنبال شاخص‌هایی بوده‌اند که وضعیت پیشرفت دانش را دقیق‌تر و سنجیده‌تر ارزیابی نماید. این تلاش‌ها به ارائه شاخص‌های مختلف و دسته‌بندی‌های انبوه ختم شده‌است (Stufflebeam, 2000؛ بازگان، ۱۳۸۰؛ داورپناه، ۱۳۸۹). در هر یک از الگوی‌های ارائه شده در زمینه ارزیابی، ارزش یا شایستگی برخی از عناصر، برتر از دیگر عناصر تلقی شده‌است (Trochim, 2006). بنابراین، انتخاب و مبنا نهادن یک الگوی مناسب، بر اساس آن بتوان اطلاعات جامعی برای اطمینان از ارزش و اعتبار برون‌دادهای علمی کسب نمود، از جمله مسائلی است که ارزشیابی از برون‌دادهای نشر علمی را مشکل نموده‌است (نیکخواه و همکاران، ۱۳۹۰).

پیشینه پژوهش

به‌منظور مواجهه با چالش‌های اساسی ارزیابی علمی در ایران و استقرار نظام جامع ارزیابی علمی، فراست خواه (۱۳۸۷) در یک نگاه جامع به شناسایی موانع ارزیابی علمی در ایران پرداخته‌است و راه‌حل‌های ممکن برای کاهش این معضلات و رسیدن به یک نظام جامع را در چهار سطح: ۱. نظام اطلاعات دانشگاهی ۲. نظام ارزیابی در دانشگاه ۳. ارزیابی بیرونی و اعتبار سنجی و ۴. پایش ملی نظام علمی دسته‌بندی نموده‌است، شرط اساسی برای اجرایی شدن بهینه این سطوح را حمایت و بسترسازی دولت و ابتکارات درونزای دانشگاه‌ها و نهادهای علمی بیان کرده‌است. با توجه به تفاوت حوزه علوم انسانی با دیگر حوزه‌های علمی، تأکید بیشتر بر استفاده از کتب و منابع تک‌نگاشتی و نبود شاخص‌ها و معیارهای معین برای سنجش و ارزیابی برون‌دادهای علمی علوم انسانی و اجتماعی است.

داورپناه (۱۳۸۹) در پژوهشی به شناسایی الگویی ترکیبی (کمی و کیفی)، برای سنجش توان علمی رشته‌ها پرداخته‌است. این فرمول ترکیبی از چند شاخص علم‌سنجی است که می‌تواند از طریق محاسبه تعداد انتشارات، تعداد استناد، تعداد استناد نشده‌ها، حجم انتشارات و استنادهای مجموعه توان علمی هر رشته را تعیین نماید و سازوکار مقایسه‌ای رشته‌های درون یک حوزه را فراهم نماید. در این شاخص، داده‌ها برحسب نوع مدرک، دوره زمانی، موضوع و پنجره استنادی محاسبه و نرمال‌سازی و بر این اساس، امکان مقایسه رشته‌های درون حوزه میسر می‌گردد. براساس این یافته‌ها، می‌توان به سنجش قدرت و ضعف رشته‌های علمی در سطوح ملی و جهانی پرداخت.

شهلایی (۱۳۹۵) در پژوهشی موردی - تاریخی به شناسایی ابعاد و شاخص‌های ارزیابی قابلیت علم و فناوری در نظام جمهوری اسلامی ایران پرداخته‌است. یافته‌های این مطالعه بر مبنای ۹ بُعد اصلی دسته‌بندی شده‌است که یکی از این ابعاد اصلی، شاخص‌های ارزیابی انتشارات علمی است که شامل ۷ بُعد فرعی است. شاخص‌های شناسایی شده به منظور ارزیابی انتشارات علمی بر کمیت برون‌دادهای علمی نظیر تعداد فصلنامه‌های تدوین شده، تعداد مقاله‌های علمی - پژوهشی منتشر شده، تعداد مقاله علمی - ترویجی منتشر شده، تعداد مجموعه مقاله‌های کامل همایش‌های معتبر علمی، تعداد مقاله‌های منتشر شده در مجموعه مقاله‌های همایش‌های معتبر، تعداد کتاب‌های علمی تألیف یا ترجمه شده انتشار یافته و تعداد کتاب‌های مرجع انتشار یافته تأکید می‌کند. مزیت این مطالعه در این است که در این پژوهش طیف گسترده‌ای از قالب‌های اطلاعاتی مطرح شده‌است و محدود بر محمل اطلاعاتی خاص نظیر مقاله نیست. هر چند بسیاری از قالب‌های دیگر علمی مانند جلسه‌های بحث، سمینارها، طرح‌های تحقیقاتی، پایگاه‌های علمی تفکیک شده و زیرمجموعه، ابعادی مانند ابعاد پژوهشی و فناوری در نظر گرفته شده‌است.

کرامت‌فر و همکاران (۱۳۹۴) در یک مطالعه تطبیقی به ارزیابی کمیت و کیفیت برون‌دادهای علمی در ایران، ترکیه

و مالزی در طی سال‌های ۱۹۹۶ تا ۲۰۱۳ م. پرداخته‌اند. معیارهای استفاده شده در این مطالعه بیشتر کمی و از نوع کتاب‌سنجی است و شاخص‌هایی نظیر تعداد مدارک، تعداد استنادات، میانگین استنادات و شاخص هرش^۱ را پوشش می‌دهد. نتایج پژوهش نشان داده است گرچه وضعیت برون‌دادهای علمی ایران نسبت به لحاظ کمیت و سرعت نشر قابل توجه است، اما این مقادیر با در نظر داشتن جمعیت پژوهشگران، چندان مطلوب نیست.

قنادی‌نژاد و حیدری (۱۳۹۹) در مطالعه‌ای نظام‌مند به بررسی روش‌ها و شاخص‌های ارزیابی برون‌دادهای علمی در حوزه علوم انسانی و اجتماعی پرداخته‌اند. نتایج پژوهش نشان داد: معیارهای کمی کتاب‌سنجی و علم‌سنجی به‌تنهایی نمی‌تواند به‌عنوان مقیاس ارزیابی استفاده شود و لازم است طیف گسترده‌ای از معیارهای کیفی برای ارزیابی مطالعات این حوزه استفاده شود و ارزیابی‌های کمی نیز به‌عنوان روش تکمیلی مورد توجه قرار گیرد.

جنوی و همکاران (۱۳۹۹) نیز به ارزیابی وضعیت انتشارات علمی ایران بر پایه نقشه جامع علمی کشور پرداخته‌اند. براساس این نقشه، هشت شاخص برای انتشارات در نظر گرفته شده است که این شاخص‌ها، ترکیبی از معیارهای کمی و کیفی در ارزیابی برون‌دادهای کشور را پوشش می‌دهند. نتایج این پژوهش نشان می‌دهد در شاخص‌های کمی مانند «تعداد مقالات در هر میلیون نفر از جمعیت»، «نسبت فارغ‌التحصیلان دانشگاهی و حوزوی به مقالات نمایه‌سازی شده در نمایه‌های بین‌المللی» و «نسبت مقالات نمایه‌سازی شده در سطح بین‌المللی به تعداد اعضای هیئت علمی» سیر روبه‌رشد قابل قبولی دارد و امکان تحقق این شاخص‌ها تا سال ۱۴۰۴ ش. دور از انتظار نیست. اما در شاخص‌های کیفی مانند «میزان استنادات در واحد انتشارات»، با وجود سیر روبه‌رشد در طی سال‌های اخیر، مقدار به‌دست آمده با میزان پیش‌بینی شده تا سال ۱۴۰۴ ش. فاصله زیادی دارد.

خیاطیان یزدی و همکاران (۱۳۹۹) با مطالعه اسناد بالادستی و گزارش‌های بین‌المللی و تحلیل مضمون آن‌ها به ارائه چارچوبی برای پایش و ارزیابی نظام ملی علم، فناوری و نوآوری ایران پرداخته‌اند. این مطالعه شامل ۱۰ معیار اصلی است که معیار چهارم آن درباره خلق برون‌دادهای علمی اعم از کتاب، ثبت اختراع و انواع مختلف مقالات است. در این بعد، به مقیاس‌های کمی و کیفی در پایش و ارزیابی نظام ملی علم، فناوری و نوآوری تأکید شده است. همچنین به روش‌های مختلف ارزیابی از جمله روش‌های کتاب‌شناختی، علم‌سنجی، شهرت و اعتبار پدیدآور و تولیدکننده اشاره شده است.

آزادی احمدآبادی (۱۴۰۱) در مطالعه‌ای به بررسی وضعیت کمی و کیفی برون‌دادهای علمی جمهوری اسلامی ایران در طی سال‌های ۲۰۱۰ تا ۲۰۲۱ م. پرداخته است. شاخص‌های ارزیابی در این پژوهش شامل مقیاس‌های کتاب‌سنجی نظیر تعداد مجلات، تعداد مقالات، استنادات و اسنپ^۲ است.

ابراهیمی درچه و همکاران (۱۴۰۲) در مطالعه‌ای به بررسی سیاست‌ها و معیارهای ارزیابی برون‌دادهای علمی در علوم انسانی پرداخته‌اند. در این مطالعه قالب‌های مختلف اطلاعاتی اعم از مقاله نشریه علمی، مقاله همایش، کتاب، طرح پژوهشی و پایان‌نامه‌های مورد بررسی قرار گرفتند. معیارهای ارزیابی در این پژوهش به ۸ دسته کلی تقسیم شده است که غالب معیارهای استفاده شده در این مطالعه از نوع معیارهای کمی است و شامل اثرگذاری در محیط آنلاین، اثرگذاری علمی، اثرگذاری اجتماعی، اثرگذاری اقتصادی و اثرگذاری فرهنگی است. معیارهای دیگر در این مطالعه شامل بستر ایجاد، ارائه و نشر T ساختار نگارشی و محتوای برون‌دادهای علمی است.

1 . Hirsch.
2 . SNIP.

کریمان و مائونسِل (Kreiman & Maunsell, 2011) در مطالعه‌ای به بررسی معیارهای کمی و کیفی اندازه‌گیری بازده علمی پرداختند. آن‌ها معتقدند، یکی از راه‌های سنجش اعتبار مقیاس‌های ارزیابی فراهم نمودن امکان مقایسه ارزیابی‌ها با معیارهای مختلف است زیرا هیچ مقیاس و معیار واحدی در ارزیابی‌های علمی نمی‌تواند پیچیدگی‌های موجود در خروجی‌های علمی را نشان دهد.

برای ارزیابی تمامی برون‌دادها در زیست‌بوم نشر علمی، برخی از پژوهشگران (Öchsner, 2013) به طبقه‌بندی محمل‌های نشر علمی مانند پایندها؛ کتب علمی، تک‌نگاشت‌ها، راهنماها و دایره‌المعارف‌ها پرداخته‌اند و متناسب با هر یک از محمل‌ها معیارهای معینی را برای ارزیابی در نظر گرفته‌اند. برخی دیگر از پژوهشگران (Hartgerink & Zelst, 2018) نیز اقدام به ایجاد معیارهای جدید ارزیابی نظیر «ماژول‌های سؤال‌محور» نموده‌اند که انعطاف‌پذیر است و در موقعیت‌های مختلف (قبل از انتشار و پس از انتشار)، برای ارزیابی برون‌دادهای متعدد می‌تواند به کار می‌رود. وُتر و همکاران (Wouter et al., 2019) افزودن معیارهای جدید ارزیابی از جمله؛ در نظر گرفتن مقیاس‌های رسانه‌های اجتماعی و دگرسنجی به سایر معیارهای ارزیابی در زیست‌بوم نشر علمی را ضروری دانست.

در مطالعه‌ای دیگر، سیورتنسن (Sivertsen, 2019) به بررسی درک و ارزیابی تحقیقات و انتشارات علمی در علوم اجتماعی و انسانی پرداخت. در این پژوهش با توجه به ماهیت مطالعات اجتماعی و انسانی تأکید می‌شود که در ارزیابی برون‌دادهای این حوزه صرفاً به استفاده از معیارهای وب‌آوساینس و اسکوپوس اکتفا نشود. زیرا متخصصان بخشی از استفاده‌کنندگان اطلاعات علمی در این حوزه هستند و بخش بیشتری از استفاده‌کنندگان، افراد جامعه را دربرمی‌گیرند. بنابراین، برخلاف سایر مطالعات بهتر است شاخص‌هایی برای ارزیابی برون‌دادهای علمی حوزه علوم اجتماعی و انسانی تعریف شود که جامعه بتواند فارغ از محدودیت‌های زبانی و قالب ویژه اطلاعاتی به ارزیابی محتوا بپردازد.

ژائو و وانگ (Zhao & Wang, 2020) در مطالعه‌ای با ترکیب ارزیابی‌های استنادی و ارزیابی‌های التمریکس به بررسی تأثیر مجلات دانشگاهی پرداختند. نمونه این مطالعه از میان مجلات بین‌المللی Library Hi Tech انتخاب شد. در این مطالعه ۱۴ شاخص استنادی با ۱۴ شاخص دگرسنجی برای ایجاد یک مدل ارزیابی جامع از تأثیر مجلات دانشگاهی (تأثیر علمی و تأثیر اجتماعی) در نظر گرفته شدند. نتایج این مطالعه نشان می‌دهد معیارهای دگرسنجی نقش مهمی را در نشان دادن میزان تأثیر مجلات دارند. با در نظر گرفتن این قابلیت، ترکیب هر دو ارزیابی و استفاده از چند شاخص می‌تواند به بهینه‌سازی وضعیت ارزیابی تأثیر مجلات کمک کند. هوانگ و همکاران (Huang et al., 2021) در مطالعه‌ای به بررسی معیارهای ارزیابی نشریات در چین پرداختند. در این مطالعه مراکز مختلف ارزیابی علمی و معیارهای مورد توجه آن‌ها^۱ مورد بررسی قرار گرفت. معیارهای ارزیابی، شامل شاخص‌های ارزیابی پس از انتشار می‌شد بر شاخص‌هایی نظیر دامنه مجله (ملی یا بین‌المللی)، پوشش موضوعی و معیارهای استنادی (ضریب تأثیر، شاخص عامل ویژه،^۲ امتیاز تأثیر مقاله،^۳ شاخص داغی،^۴ شاخص نفوذ،^۵ استناد داخلی و نرخ بهره‌برداری) تأکید داشت. در حالی که تانگ و همکاران (Tang et al., 2022) بر نبود استفاده از یک معیار ارزیابی مانند میزان استناد در بررسی خروجی‌های

1. The Chinese Science Citation Database (CSCD); the journal partition table (JPT); the AMI Comprehensive Evaluation Report (AMI); the Chinese STM Citation Report (CJCR); the "A Guide to the Core Journals of China" (China) (GCJC); the Chinese Social Sciences Citation Index (CSSCI); and the World Academic Journal Clout Index (WAJCI).

2. Eigenfactor score.

3. Article Influence score.

4. Hot index.

5. Diffusion index.

شناسایی و دسته‌بندی معیارهای ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی

علمی تأکید دارد و دلیل آن را نادیده گرفتن ارزش واقعی اثر بیان می‌کند.

ابراهیم و همکاران (Ibrahim et al., 2022) در یک مطالعه تطبیقی به طبقه‌بندی شیوه‌های مختلف ارزیابی برون‌دادهای علمی پرداخته‌اند. این پژوهش با دو رویکرد بیرونی و درون‌گرایانه انجام شده‌است. مقیاس‌های استفاده شده در این رویکرد نیز از نوع شاخص‌های کمی است و شاخص‌های تأثیر، تولید و ترکیب را پوشش می‌دهد که در سطوح خرد و کلان دسته‌بندی شده‌اند.

آنجلووا-ستانیمیرووا (Angelova-Stanimirova, 2023) در پژوهشی به بررسی شاخص‌های ارزیابی تحقیقات علمی پرداخته‌است. نتایج مطالعه نشان داده است، کتاب‌سنجی، ضریب تأثیر و معیارهای کمی و کیفی در سنجش برون‌دادهای علمی از شاخص‌های اصلی ارزیابی به حساب می‌آیند.

مصطفی و همکاران (Mustafa et al., 2023) در پژوهشی به ارزیابی معیارهای انتشار و استناد برای تعیین کمیت تأثیر علمی برون‌دادهای علمی پرداخته‌اند. در این پژوهش شاخص اچ و ۳۲ نوع به‌منظور ارزیابی نویسندگان مورد تجزیه و تحلیل قرار گرفت. یافته‌های مطالعه نشان از همبستگی بالای میان شاخص اچ و بسیاری از شاخص‌های مشتق شده از آن دارد. هرچند در برخی از شاخص‌ها نظیر A-index, E-index, H-dash index, h2 lower index, K-index, M-index میزان همبستگی کاهش می‌یابد که بیانگر تفاوت میان شاخص‌ها در رتبه‌بندی میزان تأثیر آثار علمی است. در این مطالعه، تفاوت‌های مشاهده شده بر اهمیت انتخاب دقیق معیارهای مناسب برای ارزیابی بهره‌وری و تأثیر نویسنده تأکید دارد، زیرا شاخص‌های مختلف ممکن است نتایج متمایزی را به بار آورند.

با توجه به مطالعات انجام شده در داخل و بیرون از کشور می‌توان به اهمیت شاخص‌های ارزیابی در زیست‌بوم نشر علمی پی‌برد. پژوهش‌های انجام شده در حوزه ارزیابی، بیشتر نتیجه‌محور است (شجاعی و همکاران، ۱۳۹۳؛ آزادی احمدآبادی، ۱۳۹۹؛ Jayasundara, 2020؛ Huang et al., 2021؛ Sivertsen, 2019)، و به‌دنبال بررسی و ارزیابی برون‌دادهای علمی با هدف شناسایی کمیت و کیفیت آثار علمی منتشر شده در یک برهه زمانی و سنجش عملکرد مراکز علم و فناوری است. در حالی که، پژوهش حاضر با ترکیب نتایج مختلف سعی در ارائه چهارچوب مفهومی یکپارچه دارد و مسئله ارزیابی را به‌عنوان مفهومی مهم در فرآیند قضاوت عادلانه در زیست‌بوم نشر علمی در نظر گرفته است که در نهایت بتواند به توسعه شاخص‌های ارزیابی در شبکه‌ها و پایگاه‌های نشر علمی، کمک کند.

مسئله پژوهشگران نبود شاخص‌های ارزیابی در زیست‌بوم نشر علمی نیست، بلکه مسئله، نبود دسته‌بندی جامعی از شاخص‌های ارزیابی است که فرآیند ارزیابی در موقعیت‌های مختلف را برای آنان تسهیل نماید. به نظر می‌رسد یکی از راه‌های دستیابی به این شاخص‌ها، مطالعه متون مختلف در حوزه ارزیابی نشر علمی است، حال آن‌که در شبکه‌های مختلف نشر علمی نیز معیارهای مختلفی برای ارزیابی برون‌دادهای علمی استفاده شده است. همچنین، در میان محققان کسانی هستند که ممکن است شیوه‌های جدیدی را به‌منظور ارزیابی آثار علمی پیشنهاد نمایند. از این‌رو، پژوهش حاضر قصد دارد با در نظر گرفتن هر سه رویکرد، چهارچوب یکپارچه‌ای از شاخص‌های ارزیابی در زیست‌بوم نشر علمی ارائه نماید.

روش شناسی پژوهش

مطالعه حاضر، پژوهشی آمیخته با رویکرد اکتشافی متوالی است و از نظر هدف از نوع مطالعات بنیادی است. برای گردآوری داده‌ها از راهبرد ترکیبی زنجیره استفاده شده است. در این راهبرد، ابتدا اطلاعات کیفی فراهم می‌شود و

سپس با یافته‌های روش‌های دیگر شرح و بسط داده می‌شود (Creswell, 2009). به منظور افزایش اعتبار نتایج از راهبرد سه‌سویه‌سازی استفاده شده است؛ دلیل استفاده از سه‌سویه‌سازی در این پژوهش، ریشه در نظر متخصصانی دارد که معتقدند یک فرضیه زمانی می‌تواند دوام و بقاء یابد که در برابر روش‌های مختلف و مکمل به تأیید برسد (Webb et al., 1966). استفاده از رویکرد سه‌سویه‌سازی در این مطالعه با هدف جلوگیری از تبیین سطحی معیارهای ارزیابی است (مدنی بروجنی و نصر، ۱۳۸۸). به همین منظور، برای افزایش میزان اطمینان از معیارهای ارزیابی شناسایی شده، علاوه بر «مرور نظام‌مند متون» به مشاهده و بررسی معیارهای ارائه شده در پایگاه‌های اطلاعاتی و نظرسنجی از متخصصان نشر علمی پرداخته شده است. علاوه بر متون، در بسیاری از پایگاه‌های اطلاعاتی نظیر Scopus و Web of Science معیارهای مختلفی برای ارزیابی برون‌دادهای علمی تعریف شده است. این در حالی است که در برخی از شاخص‌های ارزیابی، بین پایگاه‌های مختلف هم‌پوشانی و در برخی شاخص‌ها تفاوت‌هایی مشاهده می‌شود. این تفاوت‌ها می‌توانست در تکمیل شاخص‌های احتمالی ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی کمک نماید. با فرض اینکه متخصصان حوزه نشر علمی در مطالعات اخیر به معیارهای جدیدی برای ارزیابی برون‌دادهای علمی دست‌یافته‌اند در رویکرد سوم به نظرسنجی از متخصصان پرداخته شد. جامعه آماری پژوهش در فاز کیفی، متون مرتبط با حوزه نشر علمی است و از دو پایگاه استنادی وب‌آوساینس و اسکوپوس، براساس شیوه‌نامه کاکرین استخراج شده است. دلیل استفاده از شیوه‌نامه کاکرین، حجم زیاد پیشنهادهای تولید شده در راستای ارزیابی برون‌دادهای علمی است که هر یک از مطالعات تنها به یک یا چند جنبه از این موضوع پرداخته‌اند. استفاده از این شیوه‌نامه در گزینش اطلاعات با کیفیت برای تصمیم‌گیری و تولید شواهد ترکیبی قابل اعتماد کمک می‌کند و یکی از معتبرترین شیوه‌نامه‌های انجام مطالعات مرور نظام‌مند در جهان است. همچنین، مرورهای نظام‌مند با استفاده از شیوه‌نامه کاکرین، فرصت جمع‌آوری شواهد از پیش مشخص شده که با معیارهای واجد شرایط بودن مطابقت دارد - را فراهم می‌نماید تا بتوان به یک سؤال تحقیقی خاص پاسخ داد. این ویژگی شیوه‌نامه که از قبل یک پروتکل مستند شده را فراهم می‌نماید باعث به حداقل رساندن سوگیری در مرور نظام‌مند می‌شود (Chandler et al., 2019).

بر مبنای این شیوه‌نامه، ۸ گام اصلی برای استخراج یافته‌ها تعریف شده است و شامل، انتخاب موضوع، طراحی عبارت جستجو، تعیین معیارهای ورود و خروج، غربالگری مطالعات و حذف منابع تکراری، مرور چکیده‌ها و حذف منابع نامرتب، گزینش منابع براساس معیارها ورود و خروج برای مطالعه تمام متن، استخراج داده‌هاست (Higgins et al., 2023). مطالعات مورد بررسی در این پژوهش، منابع منتشر شده از سال ۱۹۵۷ م. تا اواسط سال ۲۰۲۱ م. است که در پایگاه‌های استنادی اسکوپوس و وب‌آوساینس یافت شده است. معیارهای گزینش مطالعات برای «مرور نظام‌مند مفهوم ارزیابی در زیست‌بوم نشر علمی» در جدول ۱، بیان شده است.

جدول ۱. معیارهای انتخاب منابع تحلیل محتوا

نوع معیار	معیارهای انتخاب شده
محتوا	حوزه نشر علمی، نشر دانشگاهی و نشر تخصصی
نوع محتوا	علمی - تخصصی و علمی-پژوهشی
زمینه	علوم، علوم اجتماعی، علوم انسانی، تکنولوژی و پزشکی
تاریخ انتشار	مقاله‌های چاپ شده از ۱۹۵۷ م. تا اواسط سال ۲۰۲۱ م.
نوع مدرک	منابع منتشر شده در قالب مقالات (علمی - پژوهشی، علمی - تخصصی)، پایان‌نامه، کتاب، گزارش (طرح‌های پژوهشی، گزارشات سازمان‌ها)، و مقالات همایش‌ها

علاوه بر معیارهای ورود و خروج، برای اطمینان از کیفیت منابع مورد مطالعه در مرور نظام‌مند، دو ارزیاب برای سنجش میزان «کیفیت اسناد گزینش‌شده» همکاری کردند. برای نظرسنجی و تعیین دیدگاه متخصصان از طیف لیکرت ۵ امتیازی استفاده شد.

پس از استخراج داده‌های اولیه بر اساس مرور نظام‌مند، به‌منظور اعتبارسنجی و تکمیل شاخص‌های احتمالی «مفهوم ارزیابی»، به بررسی ۱۲ شبکه نشر علمی پرداخته شد. این شبکه‌ها در فرآیند مرور نظام‌مند متون شناسایی شدند که روی هم‌رفته شامل ۷۲ شبکه نشر علمی بودند. این شبکه‌ها شامل پایگاه‌های اطلاعاتی، پایگاه‌های استنادی، مراکز خرید و فروش منابع علمی، پایگاه‌های دسترسی آزاد، کنسرسیوم محتوای علمی و نظایر آن را پوشش می‌دادند. مبنای گزینش این ۱۲ شبکه از میان ۷۲ شبکه علمی، فراوانی آن‌ها در مطالعات مرور نظام‌مند است. این دوازده شبکه شامل Amazon، Web of science Core Collection، Scopus، PubMed، Springer، Elsevier، Nature، SPARC، Google scholar، Arxive، DOAJ و NIH را در برمی‌گیرد. در فاز کمی نیز معیارهای ارزیابی نشر علمی توسط ۳۰ نفر از خبرگان نشر علمی داخلی و بین‌المللی مورد بررسی قرار گرفت. خبرگان نشر علمی از میان نویسندگان مقالات مرور نظام‌مند انتخاب شد و اطلاعات مورد نیاز در بازه یک ماهه گردآوری شد. به دلیل محدودیت زمانی، تغییر نشانی نویسندگان، مسدود بودن یا فوت برخی از آن‌ها نمونه پژوهش به ۳۰ نفر محدود شد. اغلب مشارکت‌کنندگان در بخش نظرسنجی از میان متخصصان بلغاری، نیجری، انگلیسی، هندی، اسپانیایی، استرالیایی، آلمانی، آمریکایی و ایرانی بودند.

جزئیات روش‌شناسی هر سه مطالعه در جدول ۲ به‌طور کامل بیان شده است.

جدول ۲. اطلاعات روش‌شناسی معیارهای ارزیابی در زیست‌بوم نشر علمی

ردیف داده‌ها	روش گردآوری	جامعه نمونه	نمونه	روش نمونه‌گیری	معیار انتخاب ابزار گردآوری اطلاعات	روایی پایایی	ابزار تحلیل یافته‌ها
۱	مرور نظام‌مند متون	پایگاه استنادی وب آوساینس و اسکوپوس	۳۳۱ متن علمی و تخصصی	نمونه‌گیری هدفمند	استفاده از شیوه‌نامه کارکن	گام بندی مطالعات در چارچوب پرزما	استخراج ضریب کاپا و اطمینان از اکسل کیفیت مطالعات
۲	مشاهده شبکه‌های نشر علمی	۷۲ پایگاه‌های اطلاعاتی، پایگاه‌های استنادی، شبکه‌های نشر دانشگاهی	۱۲ شبکه نشر علمی	نمونه‌گیری هدفمند	بر اساس بیشترین فراوانی انتخاب‌شد.	روایی منطقی (۶ نفر از خبرگان نشر علمی (متخصص علمی و عمومی)	نرم‌افزار اکسل
۳	نظرسنجی از خبرگان	۵۰۰ نفر از نویسندگان حوزه نشر علمی که از مطالعات مرور نظام‌مند شناسایی شد.	۳۰ نفر از خبرگان داخلی و خارجی	نمونه‌گیری هدفمند	تخصص افراد (نویسندگان پرسش‌نامه حوزه نشر علمی)	روایی منطقی (۶ نفر از خبرگان نشر علمی (متخصص علمی و عمومی)	نرم‌افزار اکسل

براساس جدول ۲، دو گام کیفی و یک گام کمی برای استخراج یافته‌ها در نظر گرفته شده است. در این مطالعه

نخست گام‌های کیفی به منظور استخراج شاخص‌های مؤلفه ارزیابی زیست‌بوم نشر علمی تدوین شده‌است و در نهایت با نظرسنجی از خبرگان یافته‌ها گزارش شده است. نمونه دو گام بعدی پژوهش، براساس گام نخست - مطالعات مرور نظام‌مند - استخراج شده‌است. روش نمونه‌گیری در تمامی مطالعات از نوع نمونه‌گیری هدفمند است. از مزایای استفاده از این سبک نمونه‌گیری این است که نمونه به گونه‌ای انتخاب می‌شود که تا حد امکان ویژگی‌های جامعه واقعی را نشان دهد. بنابراین داده‌هایی که از این نوع نمونه‌گیری فراهم می‌شود، اطلاعات و درک عمیقی در زمینه مورد بررسی ارائه می‌دهد. میزان ضریب توافق در مرور نظام‌مند با هدف بررسی توافق دو ارزیاب در باره کیفیت متون مورد بررسی، نشان می‌دهد که توافق دو داور از نظر آماری معنی‌دار شده‌است ($p < 0.05$) و شدت ضریب کاپا برابر با ۰.۵۶۴ است که مقداری متوسط و پذیرفتنی است. چک لیست بررسی شبکه‌های نشر علمی و پرسشنامه نظرسنجی از نظر روایی ترجمه‌ای و منطقی توسط ۶ نفر از متخصصان حوزه نشر علمی که یا سابقه فعالیت در حوزه نشر علم را داشتند و یا در حوزه نشر علمی صاحب آثار بودند مورد تأیید قرار گرفت.

یافته‌های پژوهش

یافته‌های پژوهش براساس مرور نظام‌مند، مشاهده شبکه‌های نشر علمی و نظرسنجی از خبرگان حوزه نشر علمی به تفکیک در جدول ۳، ۴ و ۵ نشان داده شده‌است.

پاسخ به پرسش نخست پژوهش: براساس مرور نظام‌مند متون، شاخص‌های اصلی ارزیابی برون داده‌های علمی در زیست بوم نشر علمی کدام است؟

بر اساس مطالعات انجام شده سه دسته‌بندی کلی از شاخص‌های ارزیابی در زیست‌بوم نشر علمی ارائه شده است که در شاخص شکلی ۴ مصداق، در نوع ارزیابی ۳ مصداق و در قالب ارزیابی ۲ مصداق کلی برای آن‌ها در نظر گرفته شده‌است. جزئیات یافته‌های مرور نظام‌مند در جدول شماره ۳ گزارش شده‌است.

جدول ۳. یافته‌های مرور نظام‌مند متون در حوزه معیارهای ارزیابی در زیست‌بوم نشر علمی

منبع	شاخص فرعی ارزیابی	شاخص‌های اصلی ارزیابی
Halliday, 2001; Bohlin, 2004; Harley, 2013; Higgs, 2018; Kindelan, 2009; Col- lins, 2005	ارزیابی‌های محتوایی ارزیابی‌های باز ارزیابی‌های دگرسنجی ارزیابی‌های کتاب‌شناختی {ارزیابی‌های پدیدآور، ارزیابی‌های منبع}	شکل ارزیابی
Sutherland, 1977; Johannisson, 2015; Bullini Orlandi, et al, 2018	ارزیابی کمی ارزیابی کیفی ارزیابی آمیخته	نوع ارزیابی
Correia and Teixeira, 2005; Krishnan, 2013	ارزیابی فنی ارزیابی غیر فنی (ارزیابی محقق‌ساخته، ارزیابی بحث محور، ارزیابی الکترونیکی)	قالب ارزیابی

ارزیابی

بر اساس اطلاعات جدول ۳، اشکال کلی ارزیابی را می‌توان به «ارزیابی محتوایی»^۱، «ارزیابی باز»^۲، «ارزیابی‌های دگرسنجی»^۳ و «ارزیابی‌های کتاب‌شناختی»^۴ تقسیم نمود (Kriegeskorte et al., 2012; Assante et al., 2015; Harley, 2013; Visagie, 2013; Xia, 2017; Saxena et al., 2013; Higgs, 2018). هر یک از اشکال ارزیابی می‌تواند به صورت «ارزیابی کمی»، «ارزیابی کیفی» و «ارزیابی آمیخته» (Sutherland, 1977; Johannisson, 2015; Orlandi et al., 2019) تفکیک شود و در قالب «ارزیابی فنی» و «ارزیابی غیر فنی» طبقه‌بندی شود. ارزیابی‌های غیر فنی ارزیابی‌هایی هستند که به صورت استاندارد در دسترس عموم ذی‌نفعان نشر علمی نیست و در قالب «ارزیابی محقق‌ساخته»، «ارزیابی بحث‌محور»، و «ارزیابی الکترونیکی» تبیین می‌شود (Correia & Teixeira, 2005; Hartgerink & Zelst, 2018). کلیه شاخص‌های شناسایی‌شده در ارزیابی زیست‌بوم نشر علمی می‌تواند طیف گسترده‌ای از فرآیندها و روش‌های ارزیابی نظیر ارزیابی افراد (نویسنده، داور، خواننده، کتابدار)، ارزیابی سازمان‌ها (جوامع علمی، رتبه‌بندی مراکز علمی، کتابخانه‌ها و ناشران)، ارزیابی محمل‌های اطلاعاتی (کتاب، مجله، پایان‌نامه، طرح پژوهشی، گزارش‌های علمی) و ارزیابی‌های بر پایه شبکه را در برگیرد. هرچند برخی از متخصصان شاخص‌های ارزیابی‌های کتاب‌شناختی را از شاخص‌های ارزیابی‌های علم‌سنجی و اطلاع‌رسانی تفکیک می‌نمایند؛ اما بیشتر متخصصان در حوزه‌های مختلف موضوعی هر سه دسته را با عنوان ارزیابی‌های کتاب‌شناختی تعریف می‌کنند (Mejia et al., 2021).

پاسخ به پرسش دوم پژوهش: براساس مشاهده پایگاه‌های علمی، شاخص‌های اصلی ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی کدام است؟

به منظور سنجش و تکمیل شاخص‌های شناسایی‌شده در ارزیابی برون‌دادهای علمی، دوازده شبکه علمی اعم از پایگاه‌های اطلاعاتی (Elsevier, Nature, Springer, PubMed)، پایگاه‌های استنادی (Scopus و Web of science Core Collection)، فروشگاه الکترونیکی (Amazon)، کنسرسیوم علمی (SPARC)^۵، موتور کاوش علمی (Google scholar)، پایگاه‌های دسترسی آزاد علمی (DOAJ, ArXive) و مراکز علمی (NIH) مورد بررسی قرار گرفت.

1. Content Evaluation.

ارزیابی محتوا براساس زیرشاخص‌هایی نظیر هدفمند بودن، صحیح بودن، مرتبط بودن، مستند بودن و روزآمد بودن محتوا انجام می‌شود.

2. Open Evaluation.

آنچه در ارزیابی باز مورد توجه متخصصان است شامل: باز بودن فرآیند ارزیابی محتوا در پیش و پس از انتشار، آشکار بودن فرآیندهای داوری، آشکار بودن هویت داوران و آزاد بودن علاقه‌مندان به مشارکت در فرآیند داوری (پس از انتشار)، حفظ و گسترش اخلاقیات، افزایش سطح رفاه اجتماعی است.

3. Altmetrix Evaluations.

مقیاس‌های ارزیابی دگرسنجی بر اساس میزان پیوندهای دریافتی، میزان تگ‌زدن، به اشتراک گذاشتن، مشاهده کردن، توییت کردن، دانلود کردن، هایلایت کردن، پسندیدن و نظایر آن است.

4. Bibliographic Evaluations.

معیار ارزیابی در ارزیابی‌های کتاب‌شناختی، براساس میزان پیوند برون‌دادهای بازیگران در چرخه نشر علمی (پس از انتشار) است.

۵. SPARC (Scholarly Publishing and Academic Resources Coalition) یا «ائتلاف انتشارات علمی و منابع دانشگاهی»، سازمانی است غیر انتفاعی که برای حمایت از سیستم‌های آموزشی و پژوهشی ائتلافی از کتابخانه‌ها را تشکیل داده است. این کنسرسیوم ابتدا شامل کتابخانه‌های آمریکای شمالی بود اما امروزه شاهد عضویت کتابخانه‌های مختلف از کشورهای مختلف نظیر ژاپن، استرالیا، عربستان، کشورهای اروپایی و آفریقایی نیز هستیم. لازم به ذکر است دسترسی به منابع این کنسرسیوم به صورت دسترسی باز انجام می‌شود.

جدول ۴. وضعیت شاخص‌های شناسایی شده در ارزیابی شبکه‌های نشر علمی

مؤلفه	شاخص اصلی	شاخص فرعی	Arxiv	Elsevier	Springer	Google Scholar	PubMed	DOAJ	Nature	WOSCC	NIH	SPARC	Scopus	Amazon	
ارزیابی	شکل ارزیابی	ارزیابی محتوایی	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
		ارزیابی‌های باز	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
		ارزیابی‌های دگرسنجی	*	✓	✓	✓	✓	*	✓	✓	✓	✓	*	✓	
		ارزیابی کتابشناختی	*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
		{ارزیابی پدیدآور	*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
		{ارزیابی منبع}	*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
		ارزیابی کمی	*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		ارزیابی کیفی	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		ارزیابی آمیخته	*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
		ارزیابی فنی	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ارزیابی	نوع ارزیابی	ارزیابی محقق ساخته	*	✓	✓	✓	*	✓	✓	✓	✓	✓	*	✓	
		ارزیابی بحث محور	*	✓	✓	✓	✓	*	*	✓	✓	*	*	*	
		ارزیابی الکترونیکی	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

با مشاهده شبکه‌های علمی، شاخص جدیدی بر شاخص‌های شناسایی شده در ارزیابی علمی اضافه نشد. اما یافته‌های بخش نخست پژوهش در این گام، تا حد بسیار زیادی مورد تأیید قرار گرفت. براساس جدول ۴، علی‌رغم این که برای هر شبکه علمی اهمیت دارد، محتوایی که در صفحات خود بارگذاری می‌کند به لحاظ شاخص‌های ارزیابی اعتبار داشته باشد، اما این ارزیابی در برخی شبکه‌ها نظیر اسکوپوس، مجموعه هسته وب‌گاه علم^۱، پابند، گوگل اسکالر، آمازون، مؤسسه ملی سلامت و اسپارک توسط ناشرانی اعمال می‌شود که منابع موجود در شبکه، از آن‌ها خریداری یا تهیه شده است. به‌طور کلی هر یک از شبکه‌ها علاوه بر ارزیابی صاحبان محتوا، برای کاربران امکانات ارزیابی منابع و محتوای گردآوری شده را در نظر گرفته‌اند، از جمله ارزیابی‌های دگرسنجی که در اغلب شبکه‌ها مشاهده می‌شود. برخی

1. Web of Science Core Collection (WOSCC).

از این شبکه‌ها که به‌عنوان ناشر فعالیت دارند اغلب شاخص‌های ارزیابی را پیش از چاپ منابع اعمال می‌کنند. در آرشیو با توجه به سیاست‌ها و اهداف شبکه، به ارزیابی محتوایی، باز، کیفی و فنی در قالب ارزیابی‌های الکترونیکی توجه شده است. از میان قالب‌های ارزیابی، بر ارزیابی‌های بحث محور تأکید چندانی نشده است؛ اما ارزیابی الکترونیکی و فنی، مورد توجه کلیه شبکه‌ها واقع شده است. برخی شبکه‌ها از قالب ارزیابی‌های خودساخته برای ارزیابی محتوا استفاده می‌کنند. به‌عنوان مثال: در الزویر از «اختراعی برای بهبود ارزش»^۱ استفاده شده است. در شبکه آمازون نیز بعد از هر جستجو می‌توان «نظر ویراستار»، «درجه»، «بهترین فروشنده‌گان»، «نرخ اشتراک‌گذاری» و «نظرات بینندگان»^۲ را مشاهده کرد. همچنین در مؤسسه اطلاعات سلامت برای صاحبان محتوا لازم است برای ارزیابی ضریب تأثیر محتوایی که قرار است بر روی سایت بارگذاری کنند؛ در clinicaltrials.gov ثبت‌نام نمایند تا توسط گروه مرتبط ارزیابی شود.

پاسخ به پرسش سوم پژوهش:

بر اساس نظرسنجی از خبرگان نشر علمی، شاخص‌های اصلی ارزیابی برون‌دادهای علمی در زیست بوم نشر علمی کدام است؟

مشارکت کنندگان در بخش نظرسنجی از کشورهای بلغارستان،^۳ آمریکا،^۴ اسپانیا،^۵ هند،^۶ نیجریه،^۷ استرالیا،^۸ آلمان،^۹ انگلستان^{۱۰} و ایران بودند که مقالات آن‌ها در مرور نظام‌مند نیز استفاده شد. این خبره‌ها بیشتر استادان دانشگاه یا مدیران بخش‌های مرتبط با نشر علمی بودند و حداقل چند مقاله متناسب با مؤلفه‌ها و شاخص‌های نشر علمی منتشر کرده بودند. یافته‌های این بخش از مطالعات در جدول ۴ بیان شده است.

جدول ۵. یافته‌های نظرسنجی از خبرگان در ارزیابی از شبکه‌های نشر علمی

ضریب پایایی درونی مؤلفه‌های نشر علمی			
مؤلفه	شاخص	ضریب آلفای کرونباخ	تعداد آیتم
	شکل ارزیابی	۰.۷۰۶	۴
ارزیابی	نوع ارزیابی	۰.۷۰۶	۳
	قالب ارزیابی	۰.۶۰۴	۲
پایایی کل	ارزیابی	۰.۷۵۲	۹

1. Innovation to improve the value.
2. Editorial review, grade level, best level, best seller rank, share rating, customer review.
3. University Of Food Technology.
4. Ithaka.
5. Fundación Universitaria.
6. Pipalrawan Dewas (Madhya Pradesh).
7. University of Nigeria.
8. Charles Sturt University.
9. Georg-August-Universität Göttingen.
10. University of Northampton.

در این بخش از پژوهش نیز خبرگان علمی به تأیید شاخص‌های شناسایی شده در حوزه ارزیابی در زیست‌بوم نشر علمی پرداختند که در دو مطالعه قبلی مورد تأیید قرار گرفته بود. در این نظرسنجی هیچ‌یک از نخبگان شاخص جدیدی بر مؤلفه ارزیابی اضافه نمودند. براساس جدول ۵، روی هم‌رفته شاخص‌های شناسایی شده در ارزیابی مورد تأیید (۷۵۲). خبرگان نشر علمی است. و از نظر آماری، دارای اعتبار قابل قبولی است (محمدبیگی و همکاران، ۱۳۹۳). کمتر شدن ضریب پایایی در «قالب ارزیابی» به نظر می‌رسد بیشتر تحت تأثیر تعداد آیتم‌های مورد بررسی در این پژوهش است.

بحث و نتیجه‌گیری

یافته‌های پژوهش نشان می‌دهد مصادیق مختلف ارزیابی برون‌دادهای علمی در زیست‌بوم نشر علمی در ۳ شاخص اصلی و ۹ شاخص فرعی طبقه‌بندی می‌شوند. شاخص‌های اصلی شامل شکل، نوع و قالب ارزیابی است. در شکل ارزیابی، ۴ شاخص فرعی شناسایی شد که شامل ارزیابی محتوایی، ارزیابی‌های باز، ارزیابی‌های دگرسنجی و ارزیابی‌های کتاب‌شناختی (ارزیابی پدیدآور و ارزیابی منبع) است. نوع ارزیابی، نیز دارای ۳ شاخص فرعی است که ارزیابی کمی، ارزیابی کیفی و ارزیابی آمیخته را پوشش می‌دهد. قالب‌های ارزیابی، همچنین ۲ شاخص فرعی را در بردارند که شامل قالب ارزیابی فنی و قالب ارزیابی غیرفنی (محقق‌ساخته) می‌شوند.

نتایج هر سه مطالعه اعم از مرور نظام‌مند، مشاهده شبکه‌های نشر علمی و نظرسنجی از خبرگان حاکی از آن است که تمامی شاخص‌های اصلی و فرعی مورد تأیید است هرچند در برخی شاخص‌ها مانند قالب ارزیابی بین مطالعات مرور نظام‌مند با بررسی شبکه‌های نشر علمی و نظرسنجی از خبرگان تفاوت ناچیزی مشاهده گردید که می‌تواند برآمده از افزایش تمرکز متخصصان به یکی از قالب‌ها که همان قالب‌های ارزیابی استاندارد و شناخته‌شده در جوامع علمی است، باشد.

آنچه از نتایج کسب شده از مقاله حاضر استنتاج می‌شود این است که زیست‌بوم نشر علمی از اجزای مختلفی شامل متخصصان (احمدی ۱۳۹۸؛ دباغ و عزیز پور، ۱۳۹۹؛ خراسانی و همکاران، ۱۴۰۰؛ پورعباس و همکاران، ۱۴۰۱؛ Salim, 2013؛ Schimanski & Alperin, 2018؛ Bittner, 2019)، مراکز علمی (سلطانی و همکاران، ۱۳۹۲؛ شریعت و همکاران، ۱۳۹۹؛ شفیع، ۱۴۰۰؛ Krishna, 2002؛ Brankovic, 2023؛ Udupi et al., 2023)، محمل‌های اطلاعاتی (نوروزی‌چاکلی و صمدی، ۱۳۸۹؛ آرمند و همکاران، ۱۴۰۰؛ غائبی و همکاران، ۱۴۰۱؛ Kim & Oh, 2019؛ Hook et al., 2018)، حوزه‌های موضوعی (بهاری‌فر و دانایی‌فرد، ۱۳۸۵؛ نوروزی‌چاکلی، ۱۳۹۹؛ نوروزی‌چاکلی و راه‌جو، ۱۴۰۱؛ Tang et al., 2022؛ Nefedov 2019)، نظام‌های اطلاعاتی و دانشی (حکیمی، ۱۴۰۰؛ یعقوبی و شکوهی، ۱۴۰۱؛ Martín-Martín et al., 2018؛ Halevi et al., 2017) تشکیل شده است که نیازمند شاخص‌ها و روش‌های مختلفی برای ارزیابی است. در انتخاب شاخص‌های ارزیابی، علاوه بر اجزا، نیازها و اهداف؛ افراد و سازمان‌ها نیز نقش تعیین‌کننده‌ای دارند. دسته‌بندی شاخص‌های ارزیابی در شفافیت فرایندهای نشر علمی و انتخاب روش‌های مختلف ارزیابی به صاحب‌نظران کمک خواهد کرد. ضرورت وجود معیارهای مختلف در مطالعات دیگر (فراستخواه ۱۳۸۷؛ جنوی و همکاران، ۱۳۹۹؛ فنادی‌نژاد و حیدری ۱۳۹۹؛ Kreiman & Maunsell, 2011؛ Wouters, et al., 2019)، نیز مورد تأکید قرار گرفته است که نشان می‌دهد شرایط و تفاوت‌ها در زیست‌بوم نشر علمی، استفاده از شاخص‌های مختلف برای مقاصد گوناگون را توجیه می‌کند. به‌عنوان مثال، تعیین کمیت دقیق

بازده علمی در کوتاه مدت کار ساده‌ای نیست (Kreiman & Maunsell, 2011)، اما در حال حاضر، در جوامع علمی برای ارزیابی برون‌دادهای علمی دانشمندان، آزمایشگاه‌ها، مراکز علمی و پژوهشی، استفاده از این معیارها ضرورت دارد. در این مطالعه هر یک از شاخص‌های اصلی و فرعی ارزیابی به‌نوعی مکمل یکدیگر به حساب می‌آیند و استفاده از چند شاخص اصلی یا فرعی در ارزیابی برون‌دادهای علمی، شاید درک بیشتر و قضاوت عادلانه‌تر کمک خواهد کرد.

پیشنهاد‌های اجرایی پژوهش

- در راستای ارائه پیشنهاد‌های برگرفته از این مقاله، متخصصان حوزه نشر علمی می‌توانند از شاخص‌های شناسایی شده در پژوهش حاضر به‌منظور توسعه و ارتقا سامانه‌های موجود در نشر علمی استفاده نمایند.
- سیاست‌گذاران با استفاده از طیف گسترده‌ای از شاخص‌های ارزیابی که در این مطالعه ارائه شده‌است، می‌توانند سنجش‌های عادلانه‌تری برای ارزشیابی و ارتقا افراد برگزینند. محدود نبودن شاخص‌های ارزیابی در این چارچوب می‌تواند از تأکید مدیران و سیاست‌گذاران حوزه علم و فناوری بر قالب خاصی از محمل‌های اطلاعاتی نظیر مقاله کم کند.
- این مطالعه می‌تواند به‌عنوان الگوی مناسبی برای طراحی نظام جامع ارزیابی علمی قرار گیرد.
- توسعه ارزیابی‌های باز می‌تواند به رشد ارزیابی‌های سازنده و کاهش قضاوت‌های سوء‌گیرانه در فرآیندهای نشر علمی کمک کند.
- توسعه ارزیابی‌های پس از انتشار نیز می‌تواند در توسعه برون‌دادهای علمی در راستای ارتقا سطح رفاه و سلامت و اقتصاد جامعه نقش مهمی ایفاء کند.

پیشنهاد برای پژوهش‌های آتی

- بر اساس چارچوب ارائه‌شده، ظرفیت لازم برای توسعه پارامترها و روش‌ها و قالب‌های ارزیابی وجود دارد از این رو برای پژوهش‌های آتی پیشنهاد می‌شود:
- در شکل ارزیابی، ارزیابی محتوایی و ارزیابی باز با پژوهش بیشتر تعریف و شناسایی شاخص‌های سنجش پذیرتر انجام شود تا یکپارچه‌سازی نظام ارزیابی در زیست‌بوم نشر علمی میسر شود.
 - در حوزه قالب‌های ارزیابی به‌ویژه قالب‌های محقق‌ساخته که توسط هر سازمان به‌منظور ارتقا و رتبه‌بندی اعضا استفاده می‌شود، پژوهش‌های بیشتری انجام شود.
 - ارزیابی‌های بحث محور کاربرد وسیعی در حوزه‌های پزشکی دارد که به نظر می‌رسد بیشتر در عمل دیده می‌شود و پژوهش‌های انجام‌شده در راستای استانداردسازی و تعریف مقیاس‌های معتبر در بعد ارزیابی محدود است و نیازمند مطالعات بیشتری است.

تقدیر و تشکر

این مقاله برگرفته از رساله دکتری افروز همراهی به راهنمایی رویا پورنقی و مشاوره داریوش مطلبی است با عنوان «طراحی مدل مفهومی نظام ملی نشر علمی» در سال ۱۴۰۱ ش. در پژوهشگاه علوم و فناوری اطلاعات ایران دفاع شده است.

فهرست منابع

- آرمند، م.، همت، م.، رحمانی، م.، فهیمی فر، س.، بهروزی، ط.، فرزانه، حقیقی، ع.، و زرقانی، م. (۱۴۰۰). شناسایی و اولویت‌بندی ملاک‌های ارزیابی کتب چاپی در حوزه علوم پزشکی. پژوهش و نگارش کتب دانشگاهی، ۲۵(۴۹)، ۲۷۱-۲۳۷.
<https://doi.org/10.30487/rwab.2021.533853.1462>
- آزادی احمدآبادی، ق. (۱۳۹۹). واکاوی نقش نهادهای ارزیابی علم، فناوری نوآوری کشور و پیشنهاد نظام جامع پایش و ارزیابی. ره یافت. ۳۰(۷۷)، ۱۲۵-۱۴۰.
<https://doi.org/10.22034/rahyaft.2020.13823>
- آزادی احمدآبادی، ق. (۱۴۰۱). تحلیل و ارزیابی رشد کمی و کیفی برون‌دادهای علمی جمهوری اسلامی ایران. پژوهش‌نامه علم‌سنجی، ۲۸(۲)، ۲۶۵-۲۸۶.
<https://doi.org/10.22070/rsci.2022.15445.1546>
- ابراهیمی درجه، الف، منصوری، ع.، پشوتنی زاده، م.، میرباقری فرد، ع.، و شعبانی، الف. (۱۴۰۲). سیاست‌ها و معیارهای ارزیابی برون‌دادهای علمی علوم انسانی از دیدگاه متخصصان: مطالعه موردی رشته‌های زبان و ادبیات. پژوهش‌نامه پردازش و مدیریت اطلاعات، ۳۹(۱)، ۶۳-۱۰۰.
https://jipm.irandoc.ac.ir/article_705519.html?lang=en
- احمدی، ع. (۱۳۹۸). بررسی رفتار ارزیابان در نمره دهی مهارت گفتاری در زبان دوم: استفاده از بحث گروهی به عنوان ابزار پرورش ارزیاب. مسائل آموزش زبان انگلیسی، ۸(۱)، ۱۹۵-۲۲۴.
<https://doi.org/10.22054/ilt.2020.49511.461>
- بازرگان، ع. (۱۳۸۰). ارزشیابی آموزشی: مفاهیم، الگوها و فرایند عملیاتی. تهران: سمت.
- بهاری فر، ع.، و دانایی فرد ح. (۱۳۸۵). توسعه نظریه‌ها در علم الاجتماع: ضرورت تدوین معیارهای ارزیابی تولید علم نظری. مدرس علوم انسانی، ۱۰(پیاپی ۴۶) ویژه نامه مدیریت، ۸۳-۱۰۲.
<https://sid.ir/paper/7019/fa.۱۰۲-۸۳>
- پورعباس، الف، سیدنظری، ن.، نادرفام، ش.، حضرتی، ح.، تقی زاده، ع.، و فخاری، ع. (۱۴۰۱). شناسایی معیارهای انتخاب استاد نمونه در آموزش پزشکی: یک مطالعه کیفی. تصویر سلامت، ۱۳(۳)، ۲۸۶۲۵۹-۲۸۶۲۸۶.
<https://doi.org/10.34172/doh.2022.34.286259>
- جنوی، الف.، مرادی، ش.، و پاکزاد، م. (۱۳۹۹). ارزیابی وضعیت انتشارات علمی ایران بر مبنای نقشه جامع علمی کشور. پژوهش‌نامه علم‌سنجی، ۶(۱۱)، ۲۱۳-۲۳۶.
<https://doi.org/10.22070/rsci.2019.4529.1300>
- حکیمی، ه. (۱۴۰۰). بررسی وضعیت درگاه‌های وب مراکز تحقیقات و آموزش کشاورزی کشور از منظر معیارهای وب‌سنجی فائو. مطالعات کتابداری و علم اطلاعات، ۱۳(۲)، ۲۷-۴۵.
<https://doi.org/10.22055/sliss.2019.29173.1586>
- خراسانی، م.، سلامی، م.، سهیلی، ف.، و دلغندی، ف. (۱۴۰۰). معیارهای شایستگی شغلی متخصصان علم‌سنجی: مصاحبه با اعضای هیئت علمی و صاحبان شغل. پژوهش‌نامه علم‌سنجی، ۹(۱)، (بهار و تابستان)، ۱۸۹-۲۱۰.
<https://doi.org/10.22070/rsci.2021.13890.1478>
- خیاطیان یزدی، م.، فرتاش، ک.، و پورعسگری، پ. (۱۳۹۹). ارائه چارچوبی برای پایش و ارزیابی نظام ملی علم، فناوری و نوآوری ایران. راهبرد فرهنگ، ۱۳(۴۹)، ۱۱۹-۱۵۴.
<https://doi.org/10.22034/jsfc.2020.109868>

داورپناه، م. (۱۳۸۹). شاخص توان علمی: الگویی برای سنجش و مقایسه باروری علمی رشته‌ها. *کتابداری و*

اطلاع‌رسانی، ۱۳(۳)، ۲۹-۹. https://lis.aqr-libjournal.ir/article_43509.html

دباغ، ر. و عزیز پور، پ. (۱۳۹۹). ارزیابی و رتبه‌بندی معیارهای انتخاب استاد راهنما با روش‌های تصمیم‌گیری چندمعیاره؛ مطالعه موردی دانشگاه صنعتی ارومیه. *فصلنامه آموزش مهندسی ایران*، ۲۲(۸۶)، ۳۹-۵۳.

<https://doi.org/10.22047/ijee.2020.207488.1695>

سلطانی، م.، الهی، م.، فضلی، ص.، و رشیدی، م. (۱۳۹۲). ارائه چارچوب جامع ارزیابی عملکرد موسسه پژوهشی با تأکید بر مدیریت منابع انسانی. *مطالعات راهبردی در صنعت نفت و انرژی*، ۱۸(۵)، ۱۶۵-۱۹۴.

<http://iieshrm.ir/article-1-419-fa.html>

شجاعی، م.، فاتح‌راد، م.، و طباطبائیان، ح. (۱۳۹۳). طراحی برای نظام ارزیابی سیاست‌های علم، فناوری و نوآوری.

رهیافت، ۲۴(۵۷)، ۱-۱۵. https://rahyaft.nrisp.ac.ir/article_13556.html

شریعت، الف.، قنبری، ر.، سیادتی، م.، و صدیقی، س. (۱۳۹۹). استخراج و تدوین معیارهای ارزیابی شرکت‌های کوچک و متوسط مستقر در مراکز رشد فناوری براساس مدل BSC. *رشد فناوری*، ۱۷(۶۵)، ۳۵-۴۵.

<https://doi.org/10.52547/jstpi.20894.17.65.35>

شفیعی، م. (۱۴۰۰). ارزیابی عملکرد پارک‌های علم و فناوری با رویکرد مدل شایستگی سازمانی. *فصلنامه زیست بوم*

نوآوری، ۲۱(۲)، ۲۲۱-۱۲۲. <https://doi.org/10.22111/innoeco.2020.35396.1005>

شهلائی، ن. (۱۳۹۵). ابعاد و شاخص‌های ارزیابی قابلیت‌های علم و فناوری در سازمان‌های نظامی ج. ا. ایران. *راهبرد*

دفاعی، ۱۴(۲)، ۲۷۱-۱۲۷. https://ds.sndu.ac.ir/article_233.html

عمرانی، الف. (۱۳۸۶). مروری بر شاخص‌های ارزشیابی برون‌دادهای پژوهشی. *فصلنامه کتاب*، ۷۱(۱۸)، ۱۵۷-۱۷۶.

<https://sid.ir/paper/500021/fa>

غائبی، الف.، تاجداران، م.، و فهیمی‌فر، س. (۱۴۰۱). معیارهای ارزیابی کتاب‌های الکترونیکی: سیاهه پیشنهادی.

پژوهشنامه پردازش و مدیریت اطلاعات، ۲۵(۴)، ۷۲۱-۷۴۳.

https://jipm.irandoc.ac.ir/article_699033.html

فراستخواه، م. (۱۳۸۷). بررسی وضع موجود و مطلوب ارزیابی نظام علمی در ایران با تأکید بر بخش آموزش عالی.

سیاست علم و فناوری، ۲۱(۲)، ۴۳-۵۸. https://jstsp.nrisp.ac.ir/article_12749.html

قنادی نژاد، ف.، و حیدری، غ. (۱۳۹۹). روش‌ها و شاخص‌های ارزیابی تولیدات علمی در علوم انسانی و اجتماعی:

مرور نظام‌مند. *پژوهش‌نامه علم‌سنجی*، ۶(۲) (پیاپی ۱۲)، ۲۰۳-۲۳۰.

<https://doi.org/10.22070/rsci.2020.4998.1341>

کرامت فر، ع.، نوروزی چاکلی، ع.، و اسپرایی، ف. (۱۳۹۴). کمیت یا کیفیت؟ ارزیابی تطبیقی تولید علم ایران، ترکیه و

مالزی طی سال‌های ۱۹۹۶-۲۰۱۳. *دو فصلنامه علم‌سنجی کاسپین*، ۲(۱)، ۳۳-۳۸.

<https://doi.org/10.22088/acadpub.BUMS.2.1.33>

کرامت فر، ع. و رفیعی خشنود، م. (۱۳۹۵). ارزیابی بروندهای علمی پژوهشگاه رویان. نشریه علم سنجی کاسپین،

<https://doi.org/10.22088/acadpub.BUMS.3.1.36>. ۳۶-۴۴، (۱)۳

محمدبیگی، الف، محمدصالحی، ن. و علی گل، م. (۱۳۹۳). روایی و پایایی ابزارها و روش‌های مختلف اندازه‌گیری

آنها در پژوهش‌های کاربردی در سلامت. مجله دانشگاه علوم پزشکی رفسنجان. ۱۳ (۱۲)، ۱۱۵۳-۱۱۷۰.

<http://journal.rums.ac.ir/article-1-2274-fa.html>

مدنی بروجنی، الف. و نصر، الف. (۱۳۸۸). سه‌سوسازی: راهبردی برای نوآوری در پژوهش‌های آموزشی.

نوآوری‌های آموزشی، ۸ (۳۰)، ۵۳-۷۳. https://noavaryedu.oerp.ir/article_135687.html

معین، م. (۱۳۸۹). فرهنگ معین. تهران: زرین. ۱۸۱۶ ص.

نامداریان، ل. و کلاتری، ن. (۱۳۹۹). تحلیل تماتیک شاخص‌های علم، فناوری و نوآوری کشورهای منطقه سند

چشم‌انداز ۱۴۰۴. ریافت، ۳۰ (۷۷)، ۵۴-۷۰. <https://doi.org/10.22034/rahyaft.2020.13819>

نوروزی چاکلی، ع. (۱۳۹۹). سخن سردبیر: تنوع حوزه‌های موضوعی و مسئله دیرینه ارزیابی و ارتقا اعضای هیئت

علمی علوم انسانی و هنر. پژوهش‌نامه علم‌سنجی، ۶ (۱۲)، ۱-۳. <https://doi.org/10.22070/rsci.2020.1237>

نوروزی چاکلی، ع. و راهجو، الف. (۱۴۰۱). شناسایی و اعتبارسنجی شاخص‌های ارزیابی کیفیت نمایه‌های تخصصی

در حوزه‌های موضوعی علوم پایه، مهندسی، کشاورزی، علوم انسانی، علوم پزشکی و هنر. پژوهشنامه پردازش و

مدیریت اطلاعات، ۲۹ (۴)، ۱۰۳۷-۱۰۶۸. <https://doi.org/10.35050/JIPM010.2014.017>

نوروزی چاکلی، ع. و صمدی، ل. (۱۳۸۹). ارزیابی تطبیقی معیارهای ارزیابی مجله‌ها در پایگاه گزارش استنادی

مجله‌های مرکز منطقه منطقه‌ای اطلاع‌رسانی علوم و فناوری و پایگاه اطلاعات علمی جهاد دانشگاهی. کتابداری و

اطلاع‌رسانی، ۵۰ (۱)، ۳۳-۳۶. https://lis.aqr-libjournal.ir/article_43527.html

نیکخواه، م. شریف، م. و طالبی، ه. (۱۳۹۰). بررسی وضعیت موجود و ممکن کاربرد شاخص‌های ارزشیابی از

دروندهای برنامه درسی دوره تحصیلات تکمیلی دانشگاه‌های اصفهان، صنعتی اصفهان و شهرکرد. رویکردهای

نوین آموزشی، ۶ (۱)، ۲۷۵-۲۸۰. https://nea.ui.ac.ir/article_19062.html

وثوق، ف. احمدی، غ. حسن مرادی، ن. و محمدداوودی، الف. (۱۴۰۱). شناسایی ابعاد، مؤلفه‌ها و شاخص‌های بومی

ارزیابی عملکرد پژوهش در دانشگاه‌های جامع) مطالعه موردی: دانشگاه‌های آزاد اسلامی شهر تهران. مدیریت و

برنامه‌ریزی در نظام‌های آموزشی. ۱۴ (۲۶)، ۱۳-۴۲. <https://doi.org/10.52547/MPES.14.1.13>

همراهی، الف. (۱۴۰۱). طراحی مدل مفهومی نظام ملی نشر علمی [رساله دکتری منتشر نشده]. پژوهشگاه علوم و

فناوری اطلاعات ایران. (ایراندک).

<https://ganj.irandoc.ac.ir/#/articles/3f20c2850dc7969ca5540428501b8e8d>

یعقوبی، ن. و شکوهی، ج. (۱۴۰۱). ارائه چارچوبی به‌منظور ارزیابی مستمر اثربخشی سیستم‌های اطلاعاتی در

سازمان‌های دفاعی با رویکرد فازی. پژوهشنامه پردازش و مدیریت اطلاعات، ۳۰ (۴)، ۱۱۷۳-۱۱۹۶.

<https://doi.org/10.35050/JIPM010.2015.024>

- Ahmadi, A. (2019). A Study of Raters' Behavior in Scoring L2 Speaking Performance: Using Rater Discussion as a Training Tool. *Issues in Language Teaching*, 8(1), 195-224. <https://doi.org/10.22054/ilt.2020.49511.461> [In Persian].
- Alterman, R., Carmon, N., & Hill, M. (1984). Integrated evaluation: A synthesis of approaches to the evaluation of broad-aim social programs. *Socio-Economic Planning Sciences*, 18(6), 381-389. [https://doi.org/10.1016/0038-0121\(84\)90047-8](https://doi.org/10.1016/0038-0121(84)90047-8)
- Angelova-Stanimirova, A. (2023). Critical criteria for evaluation of scientific research activity in Science Direct. *Knowledge – International Journal*, 58(1), 171-177. <https://www.researchgate.net/publication/371694977>
- Armand, M., Hemmat, M., Rahmani, M., Fahimifar, S., Behrozi, Talebhighi, F., & Zarghani, M. (2022). Identifying and prioritizing the criteria for evaluating printed books in the field of medical sciences. University Textbooks; *Research and Writing*, 25(49), 237-271. <https://doi.org/10.30487/rwab.2021.533853.1462> [In Persian].
- Assante, M., Candela, L., Castelli, D., Manghi, P., Pagano, P., & Nazionale, C. (2015). Science 2.0 repositories: time for a change in scholarly communication. *D-Lib Magazine*, 21(1/2), 1-14. <https://doi.org/10.1045/january2015-assante>
- Azadi Ahmadabadi, G. (2020). Analyze the role of the country's science, technology and innovation assessment bodies and propose a comprehensive monitoring and evaluation system. *Rahyaft*, 30(77), 125-140. <https://doi.org/10.22034/rahyaft.2020.13823> [In Persian].
- Azadi Ahmadabadi, G. (2021). Study and Evaluation of Quantitative and Qualitative Growth of Scientific Outputs of the Islamic Republic of Iran. *Scientometrics Research Journal*. 8(2), 265-286. <https://doi.org/10.22070/RSCI.2022.15445.1546> [In Persian].
- Baharifar, a., & Danaeifard, h. (2006). Theory development in social science: evaluation criteria for theoretical science. *Modarres human sciences*, 10((46) management), 83-102. <https://sid.ir/paper/7019/en>. [In Persian].
- Bazargan, A. (2001). *Educational evaluation: Educational evaluation: concepts, models and operational process*. Tehran: Samt. [In Persian].
- Bittner, H. (2019). *Evaluating the Evaluator*. New York: Routledge, <http://dx.doi.org/10.4324/9780367815882-6>
- Bondi, J., & Bondi, J. (2003). *Curriculum development: a guide to practice (3th Edition)*. Columbus, OH: Merrill Publishing Company.
- Brankovic, J. (2023). Worlds of Rankings Research. In *The Routledge International Handbook of Valuation and Society* Publisher: Routledge.
- Chandler, J., Cumpston, M., Thomas, J., Higgins, J. P. T., Deeks, J. J., & Clarke, M. J. (2019). Chapter I: introduction. *Cochrane handbook for systematic reviews of interventions version*, 5(0), 3-8.
- Correia, A.M.R., & Teixeira. J.C. (2005). Reforming scholarly publishing and knowledge communication: From the advent of the scholarly journal to the challenges of open access. *Online information review*. 29 (4), pp.349-364. <https://doi.org/10.1108/14684520510617802>
- Creswell, J. W. (2009) *Research design: Qualitative, quantitative, and mixed methods approaches*. London: Sage.

- Dabbagh, R., & azizpoor, p. (2020). Evaluating and ranking the criteria of choosing a supervisor with multi-criteria decision-making techniques "a case study of Urmia University of technology". *Iranian journal of engineering education*, 22(86), 39-53. <https://doi.org/10.22047/ijee.2020.207488.1695> [In Persian].
- Davarpanah, M. R. (2010). A scientometric model for the assessment of strength and weakness of scientific disciplines. *Library and Information Sciences*, 13(3), 9-29. https://lis.aqr-libjournal.ir/article_43509.html [In Persian].
- Ebrahimi Dorcheh, E., Mansouri, A., Pashootanzadeh, M., Mirbagheri fard, A., & Shabani, A. (2023). Evaluation policies and criteria of humanities scientific outputs from experts' point of view: a case study of language and literature fields. *Iranian Journal of Information Processing and Management*, 39(1), 63-100. https://jipm.irandoc.ac.ir/article_705519.html?lang=en [In Persian].
- Emrani, E. (2007). An overview of evaluation indicators of research outputs. *Book Quarterly*, 71(18), Issue 3, 157-176. <https://sid.ir/paper/500021/fa> [In Persian].
- Farasatkah, M. (2008). A study of the existing and the desired state of the "scientific evaluation system" in Iran; emphasizing the higher education sector. *Journal of science and technology policy*, 1(2), 43-58. <https://sid.ir/paper/133069/en> [In Persian].
- Ghaebi, A., Tajdaran, M., & Fahimifar, S. (2022). The evaluation criteria of Persian academic e-books: Master Studentsâ attitude at Alzahra University. *Iranian Journal of Information Processing and Management*, 27(1), 371-397. https://jipm.irandoc.ac.ir/article_699033.html [In Persian].
- Ghanadinezhad, F., & Heidari, G. (2021). Methods and Indicators for the Evaluation of Scientific Production in the Humanities and Social Sciences: A Systematic Review. *Scientometrics Research Journal*, 6(2 (12)), 203-230. <https://doi.org/10.22070/rsci.2020.4998.1341> [In Persian].
- Gradinarov, P. (2000). An emerging system for scholarly E-publishing: how to make the cake without breaking the E-community eggs. *Exploit Interactive*, 7. <https://thecraft.icaap.org/content/2000/gradinarov/emerger.html>
- Hakimi, H. (2021). An altmetric evaluation of the researchers' scientific activities in Agricultural Research, Education and Extension Organization, Iran. *Journal of Studies in Library and Information Science*, 13(2), 27-45. <https://doi.org/10.22055/slis.2021.37785.1845> [In Persian].
- Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the literature. *Journal of informetrics*, 11(3), 823-834. <https://doi.org/10.1016/j.joi.2017.06.005>
- Hamrahi, A. (2022). *Designing a Conceptual Model for the National Scholarly Publication System* [Unpublished PhD Thesis]. Iranian Research Institute for Information Science and Technology (IranDoc). <https://ganj.irandoc.ac.ir/#/articles/3f20c2850dc7969ca5540428501b8e8d> [In Persian].
- Harley, D. (2013). Scholarly communication: cultural contexts, evolving models. *Science*, 342(6154), 80-82. <https://doi.org/10.1126/science.1243622>
- Hartgerink, C. H., & Zelst, M. V. (2018). "As-You-Go" Instead of "After-the-Fact": A Network approach to scholarly communication and evaluation. *Publications*, 6(2), 21 <https://doi.org/10.3390/publications6020021>

- Higgins, J.P.T., Thomas J., Chandler J., Cumpston, M., Li, T., Page M.J., Welch V.A. (editors). (2023). *Cochrane Handbook for Systematic Reviews of Interventions version 6.4* Cochrane www.training.cochrane.org/handbook
- Hook, D. W., Porter, S. J., & Herzog, C. (2018). Dimensions: building context for search and evaluation. *Frontiers in Research Metrics and Analytics*, 3, 23. <https://doi.org/10.3389/frma.2018.00023>
- Huang, Y., Li, R., Zhang, L., & Sivertsen, G. (2021). A comprehensive analysis of the journal evaluation system in China. *Quantitative Science Studies*, 2(1), 300-326. https://doi.org/10.1162/qss_a_00103
- Ibrahim, N., Habacha Chaibi, A., & Ben Ghézala, H. (2022). Comparative study of science evaluation practices. *VINE Journal of Information and Knowledge Management Systems*. <https://doi.org/10.1108/VJKMS-12-2021-0293>
- Jackson, R. (2005). Curriculum access for students with low-incidence disabilities: The promise of universal design for learning. *Wakefield, MA: National Center on Accessing the General Curriculum. Retrieved*, 3(15), 09. <https://sde.ok.gov/sites/ok.gov.sde/files/LowIncidence.pdf>
- Janavi, E., Moradi, S., & Pakzad, M. (2020). Assessment of Iran's scientific publications based on National Master Plan for Science and Education. *Scientometrics Research Journal*, 6((Issue 1, spring & summer)), 213-236. <http://doi.org/10.22070/rsci.2019.4529.1300> **[In Persian]**.
- Jayasundara, C. C. (2020). Beware the Tyranny, Resulted by Impact Factor Mania among the Scientists: A Critical Review. *Journal of the University of Ruhuna*. 8 (1), 1-17. <http://doi.org/10.4038/jur.v8i1.7960>
- Johannisson, J. (2015). Open Access scholarly publishing on the competitive market: university management as obstacle and enabler. *Culture Unbound*, 7(4), 610-617. <http://www.cultureunbound.ep.liu.se>
- Keramatfar, A., Noroozi Chakoli, A., & Esparaein, F. (2015). Quantity or Quality? Comparative assessment of the science production of Iran, Turkey and Malaysia during 1996-2013. *Caspian Journal of Scientometrics*. 2(1), 33-38. <https://doi.org/10.22088/acadpub.BUMS.2.1.33> **[In Persian]**.
- Keramatfar, A., & Rafiee Khoshnood, M. (2016). Evaluation of scientific outputs of royan institute. *Caspian Journal of Scientometrics*, 3(1), 36-44. <https://doi.org/10.22088/acadpub.BUMS.3.1.36> **[In Persian]**.
- Khayyatian Yazdi, M. S., Fartash, K., & Pourasgari, P. (2020). Development of a Framework for Monitoring and Evaluation of Iran's National System of Science, Technology and Innovation. *Strategy for Culture*, 13(49), 119-154. <https://doi.org/10.22034/jsfc.2020.109868> **[In Persian]**.
- Khorasani, M., Salami, M., Soheili, F., & Delghandi, F. (2023). Job Competency Criteria for scientometrics specialists: Interviews with Faculty Members and Job Owners. *Scientometrics Research Journal*, 9(1), 189-210. <https://doi.org/10.22070/rsci.2021.13890.1478> **[In Persian]**.
- Kim, S., & Oh, S. G. (2019). Extracting and applying evaluation criteria for ontology quality assessment. *Library Hi Tech*, 37(3), 338-354. <https://doi.org/10.1108/LHT-01-2019-0012>

- Kreiman, G., & Maunsell, J. H. (2011). Nine criteria for a measure of scientific output. *Frontiers in computational neuroscience*, 5, 48. <https://doi.org/10.3389/fncom.2011.00048>
- Kriegeskorte, N., Walther, A., & Deca D. (2012). An emerging consensus for open evaluation: 18 visions for the future of scientific publishing. *Frontiers in computational neuroscience*, 6, p.94. <https://doi.org/10.3389/fncom.2012.00094>
- Krishna, D., Mohan, S. R., Murthy, B. S. N., & Rao, A. R. (2002). Performance evaluation of public research institutes using principal component analysis. *Journal of Scientific & Industrial Research*, 61, 940-947. <http://nopr.niscpr.res.in/handle/123456789/17729>
- Lamanauskas, V. (2011). The evaluation strategies of integrated science teaching and learning. *Problems of Education in the 21st Century*, 31, 5-6. <http://www.scientiasocialis.lt/pec/node/556>
- Madani Boroujeni, S.A., & Nasr, A.R. (2009). Triangulation: a strategy for innovating in educational researches. *Journal of educational innovations*, 8(30), 53-73. https://noavaryedu.oerp.ir/article_135687.html [In Persian].
- Mangiafico, P., & Smith, K. L. (2014). Reason, risk, and reward: models for libraries and other stakeholders in an evolving scholarly publishing ecosystem. *Cultural anthropology*, 29(2), 216-235. <https://doi.org/10.14506/ca29.2.03>
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar E. D. (2018) Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of informetrics*, 12(4), 1160-1177. <https://doi.org/10.1016/j.joi.2018.09.002>
- McGuire, H., Anthony, B., Hyde, Z. W., Ashok, A., Bjarnason, B., & Mays, E. (2018). The Scholarly Publishing Ecosystem. An Open Approach to Scholarly Reading and Knowledge Management. *Rebus Community*. <https://press.rebus.community/scholarlyreading/chapter/the-scholarly-publishing-ecosystem/>
- Mejia, C., Wu, M., Zhang, Y., & Kajikawa, Y. (2021). Exploring topics in bibliometric research through citation networks and semantic analysis. *Frontiers in Research Metrics and Analytics*, 6. 742311 <https://doi.org/10.3389/frma.2021.742311>
- Mohammadbeigi, A., Mohammadsalehi, N., & Aligol, M. (2015). Validity and Reliability of the Instruments and Types of MeasurementS in Health Applied Researches. *JRUMS*, 13 (12), 1153-1170. <http://journal.rums.ac.ir/article-1-2274-fa.html> [In Persian].
- Moin, M. (2010). *A Persian dictionary*. Tehran: Zarin. [In Persian].
- Mustafa, G., Rauf, A., Ahmed, B., Afzal, M. T., Akhuzada, A., & Alharthi, S. Z. (2023). Comprehensive evaluation of publication and citation metrics for quantifying scholarly influence. *IEEE*, v. 11, 65759-65774. <https://doi.org/10.1109/ACCESS.2023.3290917>
- Namdarian, L., & Kalantari, N. (2020). Thematic Analysis of the Evaluation Indicators for Science, Technology, and Innovation of Countries in Iran's Vision 1404 Document. *Rahyaft*, 30(77), 45-70. 742311. <https://doi.org/10.22034/rahyaft.2020.13819> [In Persian].
- Neely, A., Gregory, M., & Platts, K. (2005). Performance measurement system design: A literature review and research agenda. *International journal of operations & production management*, 25(12), 1228-1263. <https://doi.org/10.1108/01443570510633639>
- Nefedov, S. T. (2019). Towards evaluation in scientific reviews (based on German linguistics).

- Nelly, A. (2005). The evolution of performance measurement research. *International Journal of Operation & Production Management*, 25(12), 1267-1277. <https://doi.org/10.17516/1997-1370-0494>
- Nikkhah, M., Sharif, M., & Talebi, H. (2011). A study of the present and possible situations of the application of evaluation indicators of curriculum inputs at the postgraduate level at Isfahan, Shahrekord and Isfahan Industrial universities. *New Educational Approaches*, 6(1), 75-102. https://nea.ui.ac.ir/article_19062.html [In Persian].
- Noroozi Chakoli, A., & Samadi, L. (2010). Comparative Evaluation of Journal Selection Criteria in Journal Citation Reports (PJCR) available at the Reginal Information center for Science and Technology (RiCeST) and SID. *Library and Information Science*, 13(2), 33-60. https://lis.aqr-libjournal.ir/article_43527.html [In Persian].
- Noroozi Chakoli, A. (2020). Note from the Editor-in-Chief: Diversity of subject areas and the longstanding issue of evaluation and promotion of faculty members in the humanities and arts. *Scientometrics Research Journal*, 6(2), 1-3. <https://doi.org/10.22070/rsci.2020.1237> [In Persian].
- Noroozi Chakoli, A., & Rahjoo, A. (2022). Scientometrics, International Special Indexes, scientific evaluation. *Iranian Journal of Information Processing and Management*, 29(4), 1037-1068. <https://doi.org/10.35050/JIPM010.2014.017> [In Persian].
- Öchsner, A. (2013). *Introduction to scientific publishing: backgrounds, concepts, strategies*. Heidelberg: Springer.
- Oliver, M. (2000). An introduction to the evaluation of learning technology. *Journal of Educational Technology & Society*, 3(4), 20-30. <https://www.jstor.org/stable/jeductechsoci.3.4.20>
- Orlandi, L.B., Ricciardi, F., Rossignoli, C., and De Marco, M. (2019). Scholarly work in the Internet age: Co-evolving technologies, institutions and workflows. *Journal of Innovation & Knowledge*, 4(1), 55-61. <https://doi.org/10.1016/j.jik.2017.11.001>
- Pourabbas, A., Seyednazari, N., Naderfam, S., Hazrati, H., Taghizadieh, A., & Ghafari, R. (2022). Identifying the selection criteria for elite lecturer in medical education: a qualitative study. *Depiction of health*, 13(3), 286-295. <https://doi.org/10.34172/doh.2022.34> [In Persian].
- Salim, M. T. H. (2013). Evaluating the Evaluator: A Reflective Approach. *Research on Humanities and Social Sciences*. 3(7), 16-22. <http://www.iiste.org/Journals/index.php/RHSS/article/view/5577/5690>. 3. 16-22
- Saxena, A., Thawani, V. Chakrabarty, M., & Gharpure, K. (2013). Scientific evaluation of the scholarly publications. *Journal of Pharmacology and Pharmacotherapeutics*, 4(2), 125-129. <https://doi.org/10.4103/0976-500X.110894>
- Schimanski, L. A., & Alperin, J. P. (2018). The evaluation of scholarship in academic promotion and tenure processes: Past, present, and future [version 1; peer review: 2 approved]. *F1000Research*, 7(1605). <https://doi.org/10.12688/f1000research.16493.1>
- Seong Leem, C., & Kim, I. (2004). An integrated evaluation system based on the continuous improvement model of IS performance. *Industrial Management & Data Systems*, 104(2), 115-128. <https://doi.org/10.1108/02635570410522080>

- Shafiee, M. (2021). Evaluating the Performance of Science and Technology Parks through Implementing Organizational Excellence Model. *Journal of Innovation Ecosystem*, 1 (2), 1-22. <https://doi.org/avduj10.22111/INNOECO.2020.35396.1005> [In Persian].
- Shahlaee, N. (2016). The Dimensions and Indicators of Assessment for Capability of Science and Technology in the I.R. Iran's R. Iran's Military Organizations. *Defence Studies*, 14(2), 1-27. https://ds.sndu.ac.ir/article_233.html [In Persian].
- Shariat, I., Ghanbari, R., Siadati, S. M., & Sedigh, S. (2021). Extracting the Performance Indicators Based on BSC Model for SMEs Residing in Technology Incubators. *Roshd -e- Fanavari*, 17(65), 35-45. <https://doi.org/10.52547/jstpi.20894.17.65.35> [In Persian].
- Shojaei, S. M., Fatehrad, M., & Tabatabaeian, S. H. (2014). A Plan for Evaluating Science, Technology and Innovation Policies. *Rahyaft*, 24(57), 1-15. https://rahyaft.nrisp.ac.ir/article_13556.html [In Persian].
- Sivertsen, G. (2019). Understanding and evaluating research and scholarly publishing in the Social Sciences and Humanities (SSH). *Data and Information Management*, 3(2), 61-71. https://doi.org/10.1162/qss_a_00103
- Smit, E., & Gruttemeier, H. (2011). Are scholarly publications ready for the data era? Suggestions for best practice guidelines and common standards for the integration of data and publications. *New Review of Information Networking*, 16(1), 54-70. <https://doi.org/10.1080/13614576.2011.574488>
- Soltani, M., Elahi, M., Fazli, S., & Rashidi, M.M. (2013). "Providing a comprehensive framework for evaluating the performance of a research institute with an emphasis on human resource management". *Strategic Studies in Petroleum and Energy Industry*, 5(18), 165 - 194. <http://iieshrm.ir/article-1-419-fa.html> [In Persian].
- Stufflebeam, D.L. (2000). Foundational Models for 21st Century Program Evaluation. In: Stufflebeam, D.L., Madaus, G.F., Kellaghan, T. (Eds) *Evaluation Models. Evaluation in Education and Human Services*, 49, Springer, Dordrecht. https://doi.org/10.1007/0-306-47559-6_3
- Sutherland, W.R. (1977). Impact of technology on scientific publishing. *IEEE Transactions on Professional Communication*, (2), 56-58. <https://doi.org/10.1109/TPC.1977.6592323>
- Tang, Y. M., Chau, K. Y., Kwok, A. P. K., Zhu T., & Ma X. (2022). A systematic review of immersive technology applications for medical practice and education-trends, application areas, recipients, teaching contents, evaluation methods, and performance. *Educational Research Review*, 35, 100429. <https://doi.org/10.1016/j.edurev.2021.100429>
- Tansley, A. G. (1935). The use and abuse of vegetational concepts and terms. *Ecology*, 16(3), 284-307. <https://doi.org/10.2307/1930070>
- Trochim, K. (2006). *Introduction to evaluation*. Research Methods Knowledge Base. *Scientific Research Publishing*. <http://www.socialresearchmethods.net/kb/qualmeth.php>
- Trochim, W. M., & Donnelly, J. P. (2001). *Research methods knowledge base* (Vol. 2). Macmillan Publishing Company, New York: Atomic Dog Pub.
- Tsujimoto, M., Kajikawa, Y., Tomita, J., & Matsumoto, Y. (2018). A review of the ecosystem concept—towards coherent ecosystem design. *Technological Forecasting and Social Change*, 136, 49-58. <https://doi.org/10.1016/j.techfore.2017.06.032>

- Udupi, P. K., Dattana, V., Netravathi P. S., & Pandey J. (2023). Predicting Global Ranking of Universities across the world using Machine learning regression technique. *Science and Technology Ethics and Human Future (STEHF)*. 156, 244-250. <https://doi.org/10.1051/shsconf/202315604001>
- Visagie, S. (2013). Review of the book Scholarly communication in science and engineering research in higher education by Wei Wei (Eds). *South African Journal of Libraries and Information Science (SAJLIS)*, 72(2). <https://doi.org/10.7553/72-2-1136>
- Vosough, F., Ahmadi, G., Hassanmoradi, N., & Mohammaddavoodi, A. (2021). Identifying the native dimensions, components and indicators of research performance evaluation in university (A case study: Islamic Azad University in Tehran). *Journal of Management and Planning In Educational System*, 14(1), 13-42. <https://doi.org/10.52547/MPES.14.1.13> [In Persian].
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966) *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago: Rand McNally <https://psycnet.apa.org/record/1966-06543-000>
- Wouters, P., Zahedi, Z., & Costas, R. (2019). Social media metrics for new research evaluation. *Springer handbook of science and technology indicators*, 687-713. https://doi.org/10.1007/978-3-030-02511-3_26
- Xia, J. (2017). *Scholarly Communication at the Crossroads in China*. Chandos Publishing. 1-170.
- Xu, J., Li, Z., Shen, W., & Lev, B. (2013). Multi-attribute comprehensive evaluation of individual research output based on published research papers. *Knowledge-Based Systems*, 43, 135-142. <https://doi.org/10.1016/j.knosys.2013.01.029>
- Yaghoubi, N. M., & Shukuhy, J. (2022). Provide a framework to assess the continuing effectiveness of information systems in defense agencies using fuzzy approach. *Iranian Journal of Information Processing and Management*, 30(4), 1173-1196. <https://doi.org/10.35050/JIPM010.2015.024> [In Persian].
- Yazdani, K., Rahimi-Movaghar, A., Nedjat, S., Ghalichi, L., & Khalili, M. (2015). A 5-year Scientometrics analysis of research centers affiliated to Tehran University of Medical Sciences. *Medical Journal of the Islamic Republic of Iran*, 29, 206. <https://pubmed.ncbi.nlm.nih.gov/26157724/>