

Mapping the Thematic Network of Humanities Research Output Evaluation: A Mixed Method

Abstract

Purpose: Evaluating the outputs of the humanities has always been one of the fundamental challenges in scientific systems and research policy-making; unlike empirical sciences that often use quantitative and tangible criteria and indicators to evaluate their outputs, the humanities, due to their qualitative, interdisciplinary, and context-focused nature, require a different evaluation model. Therefore, it is important and necessary to manage the evaluation of scientific outputs of the humanities in a way that is appropriate to their characteristics. Although research has pointed out the characteristics of the humanities and the need for evaluation that is appropriate to them, as well as suggestions in this regard, how to plan for it has received less attention. The present study aims to provide a pattern for a better understanding of the evaluation process and its application in policymaking by systematically analyzing the interactions between the components of humanities output evaluation.

Methodology: The implementation process of this research follows a combined approach, incorporating both quantitative and qualitative methods simultaneously. Initially, we conducted a scoping review to identify and screen sources related to the evaluation of scientific outputs in the humanities. To identify potentially relevant documents, Scopus and Web of Science databases were searched without time limits. These sources were then coded using thematic analysis. To arrive at the themes, theoretical, inductive, and descriptive coding methods were used to develop a hierarchical structure of themes (category, domain, and taxonomy). Concurrently, we developed the conceptual structure of the data using bibliometric methods. Conceptual structure in Bibliometrix software (from the R programming language libraries) is a tool for drawing the word co-occurrence network (clustering), thematic map (matrix), and Thematic evolution over time (time course), and as a quantitative method complements qualitative methods. Finally, by integrating the results from these two stages, we created a network of themes related to the subject at hand. The theme network as an illustration tool is a network that summarizes the main themes of a text and presents its underlying pattern.

نویسنده اول (لاتین)^۱

نویسنده دوم (لاتین)^۲

نویسنده سوم (لاتین)^{*۳}

^۱ PhD in Knowledge and Information Science, Department of Knowledge and Information Science, University of Isfahan, Isfahan, Iran; e.brahim@edu.ui.ac.ir

^۲ Associate Professor, Department of Knowledge and Information Science, University of Isfahan, Isfahan, Iran. (Corresponding Author); a.mansouri@edu.ui.ac.ir

Receive:

.././....

Acceptance:

.././....

The thematic network is presented graphically to make any hierarchical concepts tangible, to give fluidity to the themes, and to emphasize the interrelationships throughout the network.

Findings: In the first stage, three taxonomies of evaluation approaches, evaluation methods, and evaluation tools were identified with different domains and categories. Evaluation approaches include five domains of evaluation fairness, conceptual consensus, impact citation, formative evaluation, and evaluation policy. Evaluation methods have three domains beyond the article, evaluation criteria, and peer review arm. Two domains of strong infrastructure and technological evaluation can also be included in the category of evaluation tools. In the second stage, the motor, niche, basic, and emerging themes and their changes during the years 2001 to 2025 were identified. Topics related to the evaluation of scientific outputs in the humanities were among the motor themes between 2001 and 2018, well-developed, and highly interconnected. However, these topics tend to become basic and niche topics. This can be seen as an indication of the topics' embeddedness in research in this area and, consequently, their wider attention in the field of humanities evaluation. The clusters have changed between 2019 and 2025; the largest shift has been from driving, fundamental-driving, and specialized-driving themes to other quadrants. In the third stage, bibliometric results were used to draw relationships between the themes of the first stage of identification and the network of themes for evaluating scientific outputs in the humanities. In all three taxonomies of approaches, methods, and evaluation tools, there is an internal relationship between the domains. In the taxonomy of approaches, the domain of evaluation policy has the most relationships with the other domains, which indicates its underlying nature. In the taxonomy of methods, the domain of evaluation criteria is related to the other two domains, and this may be due to the importance of how evaluation is done.

Conclusion: In policymaking, evaluation should act like a puzzle, meaning that approaches, methods, and evaluation tools should change simultaneously and proportionately. It is necessary to analyze the strengths and weaknesses, as well as the opportunities and threats, of the current situation and plan for the desired outcome. The result of this change in approach can be self-sufficiency in humanities evaluation at the policy-making level and reduced evaluation anxiety at the researcher level.

Keywords: Humanities, Research evaluation, Scientific output, Research output, Thematic analysis, Bibliometrics, Thematic network

ترسیم شبکه مضامین ارزیابی بروندادهای علمی علوم انسانی با

رویکرد ترکیبی

نویسنده اول^۱

نویسنده دوم^۲

نویسنده سوم^{*۳}

چکیده

هدف: هدف مطالعه حاضر تحلیل نظام مند تعاملات مضامین ارزیابی بروندادهای علوم انسانی است تا الگویی برای درک بهتر فرایند ارزیابی و به کارگیری آن در سیاست گذاری ها ارائه دهد.

روش شناسی: پژوهش از نظر فرایند اجرا، ترکیبی (کمی و کیفی) از نوع هم زمان است. شناسایی و غربال منابع مرتبط با ارزیابی بروندادهای علمی علوم انسانی از طریق مرور دامنه انجام شد. سپس منابع با استفاده از روش تحلیل مضمون، کدگذاری گردید. به طور هم زمان، ساختار مفهومی داده ها با استفاده از روش کتابسنجی ترسیم شد. در نهایت، با ترکیب خروجی این دو مرحله، شبکه مضامین موضوع مورد نظر ترسیم گردید.

یافته ها: در مرحله اول سه طبقه رویکردهای ارزیابی، روش های ارزیابی و ابزارهای ارزیابی با حوزه ها و مقوله های مختلف شناسایی شد. رویکردهای ارزیابی دربرگیرنده پنج حوزه عدالت در ارزیابی، اجماع مفهومی، از استناد به تأثیر، ارزیابی تکوینی و سیاست ارزیابی است. روش های ارزیابی شامل سه حوزه فراتر از مقاله، معیارهای ارزیابی و بازوی هم تراز خوانی بوده و ابزارهای ارزیابی نیز با دو حوزه زیرساخت قوی و ارزیابی فناوریانه شناخته شدند. در مرحله دوم مضامین محرک، بسیار تخصصی، اساسی، و پیرامونی و تغییرات آن ها طی سال های ۲۰۰۱ تا ۲۰۲۵ مشخص شد. در مرحله سوم از نتایج کتابسنجی استفاده گردید تا روابط بین مضامین مرحله اول شناسایی و شبکه مضامین ارزیابی بروندادهای علمی علوم انسانی ترسیم شود. بیشترین ارتباط مربوط به حوزه معیارهای ارزیابی بود.

نتیجه گیری: در سیاست گذاری ارزیابی باید مانند یک پازل عمل کرد بدین معنا که رویکردها، روش ها و ابزارهای ارزیابی به طور هم زمان و متناسب با هم تغییر کنند. توجه به تفاوت علوم امکان طراحی ابزارها و روش های کارآمدتر را فراهم می آورد. لازم است نقاط قوت و ضعف، فرصت ها و تهدیدهای وضعیت موجود را واکاوی و برای وضعیت مطلوب برنامه ریزی نمود. نتیجه این تغییر می تواند خود کفایی ارزیابی علوم انسانی در سطح سیاست گذاری و کاهش اضطراب ارزیابی در سطح پژوهشگران باشد.

واژگان کلیدی: علوم انسانی، ارزیابی پژوهش، برونداد علمی، برونداد پژوهشی، تحلیل مضمون، کتابسنجی، شبکه مضامین

مقدمه و بیان مسئله

علوم انسانی در عصر حاضر به عنوان یکی از ارکان بنیادین توسعه، نقشی بی‌بدیل در تبیین جنبه‌های مختلف و حل مسائل پیچیده اجتماعی، فرهنگی، سیاسی و اقتصادی ایفا می‌کند. با توجه به گستردگی علوم انسانی و تلقی نامناسب از وضعیت آن در بعضی جوامع، ارزیابی برونادهای این حوزه با هدف شناسایی نقاط قوت و ضعف و زمینه‌سازی ارتقای کیفیت و جایگاه آن برای پیشبرد کارکردهای مورد انتظار، به ضرورتی انکارناپذیر تبدیل شده‌است. اما آنچه مهم است، چگونگی ارزیابی و لحاظ نمودن تفاوت علوم انسانی با سایر علوم است تا از به کارگیری الگوهای واحد و یکسان‌سازی جلوگیری شود.

ارزیابی برونادهای علوم انسانی یکی از چالش‌های اساسی نظام‌های علمی و سیاست‌گذاری پژوهشی بوده و هست. برخلاف علوم تجربی که غالباً از معیارها و شاخص‌های کمی و ملموس برای ارزیابی استفاده می‌کنند، علوم انسانی به دلیل ماهیت کیفی، بین‌رشته‌ای و وابسته به بافت، نیازمند الگوی ارزیابی متفاوتی است (Donovan, ۲۰۰۷; Hug et al., ۲۰۱۳). این ماهیت پیچیده موجب شده چارچوب ارزیابی در علوم انسانی با عدم انسجام مواجه گردد و سنجش کیفیت و تأثیر پژوهش‌های آن مستلزم معیارها و شاخص‌های متنوع باشد. هرچند پژوهش‌ها بر ویژگی‌های خاص علوم انسانی و ضرورت ارزیابی متناسب با آن‌ها تأکید کرده و پیشنهادهایی نیز ارائه داده‌اند (Si vert sen, ۲۰۱۶; Faggi ol ani & Sol i mi ne, ۲۰۱۸; Kul czycki, ۲۰۲۲; Lauronen, ۲۰۲۰; et al.), اما برنامه‌ریزی و ساماندهی دقیق برای این امر کمتر مورد توجه قرار گرفته‌است. از مرتبط‌ترین نمونه‌های پژوهش مرتبط با ویژگی‌های علوم انسانی مطالعه آلموس پنوئلا و همکاران (Ol mos-Peñuel a et al., ۲۰۱۴) با عنوان "آیا STEM از مریخ و SSH از زهره است؟" است که تنوع رشته‌ای چالش‌برانگیز برای ارزش اجتماعی پژوهش در آن آزمون شده‌است. بر اساس یافته‌ها، پژوهش‌های علوم انسانی و اجتماعی با پژوهش‌های علم، فناوری، مهندسی و ریاضیات متفاوت است اما نه به طوری که باعث شود به صورت نظام‌مند برای جامعه کمتر مفید باشند بلکه نحوه تعامل با جامعه و تأثیرگذاری متفاوتی دارند.

همان‌طور که رویکردهای مختلف ارزیابی تأثیر پژوهش بر مبنای مفروضات متفاوتی در باب ماهیت دانش، هدف و کیفیت پژوهش شکل می‌گیرند (Greenhal gh et al., ۲۰۱۶)، تفاوت میان علوم مختلف نیز بر شیوه ارزیابی اثرگذار است (Wóbl ewska, ۲۰۲۱). این اختلاف‌ها رفته‌رفته در تعریف «تأثیر» و شیوه ارزیابی آن در علوم انسانی متجلی شده‌اند (Pedersen et al., ۲۰۲۰; Real e et al., ۲۰۱۸) و از آنجا که علوم انسانی به عنوان علوم نرم در برابر علوم سخت قرار می‌گیرد (Bonaccorsi et al., ۲۰۱۷)، ارزیابی پژوهش‌های آن بحث‌برانگیز شده و مباحث متعددی را در سطح جهانی برانگیخته‌است (Tol edo, ۲۰۱۶).

از سوی دیگر، ارزیابی متناسب با برونادهای علمی علوم انسانی پدیده‌ای نسبتاً جدید است که هم نقدهای جدی به آن وارد شده و هم پیشنهادات متعددی برای اصلاح و تکامل آن مطرح گردیده‌است (De Ri j cke et al., ۲۰۲۳). تحول و تکامل این حوزه در آینده نیازمند برنامه‌ریزی است. به طور نمونه، پژوهش علی‌نژاد چمازکتی و میرحق جو لنگرودی (۱۴۰۲) نشان می‌دهد که در افق ۱۴۲۰ مهم‌ترین نیروی پیشران تولیدات علمی علوم انسانی، ساختار سیاست‌گذاری و جهت‌دهی به پژوهش‌هاست؛ عاملی

که بر توانمندی، مهارت، خلاقیت پژوهشگران، همگامی آن‌ها با تحولات نوظهور فناورانه، و توسعه رویکردهای میان‌رشته‌ای تأثیر مستقیم دارد. همچنین، مرزبان و همکاران (۱۴۰۲) در سناریوهای آینده آیین‌نامه ارتقای مرتبه علمی اعضای هیئت علمی ایران بسترهایی را برای اتخاذ راهبردها، سیاست‌ها و برنامه‌ریزی‌های نو فراهم کرده‌اند. این شواهد نشان می‌دهد که برنامه‌ریزی برای ارزیابی باید با توجه به پیشران‌ها، روندها و کلان‌روندهای مؤثر انجام گیرد.

یکی از ایرادات اساسی در مدیریت ارزیابی، فقدان رویکردی جامع و نظام‌مند در تحلیل و ارزیابی بروندهای علمی است. مطالعات گذشته بیشتر بر شناسایی و دسته‌بندی مؤلفه‌ها به صورت جداگانه تمرکز داشته و از بررسی ارتباط و تعامل میان آن‌ها غفلت کرده‌اند (Darai o et al., ۲۰۱۱). این فقدان شبکه تعاملات منجر به پراکندگی و ناهماهنگی در ارزیابی‌ها و در نتیجه تصمیم‌گیری‌های کم‌اثر در علوم انسانی شده است. پیامدهای آن نیز در تخصیص منابع پژوهشی، سیاست‌گذاری علمی و حتی در تعیین ارزش پژوهش‌های علوم انسانی قابل مشاهده است (Wl sdon, ۲۰۱۵). حال آنکه ارزیابی دقیق مستلزم توجه به ابعاد گوناگونی همچون کیفیت پژوهش، تأثیرات اجتماعی، ارتقای مهارت‌های تفکر انتقادی و توانمندسازی‌های علمی است. شبکه‌ای از مؤلفه‌ها می‌تواند به پژوهشگران و سیاست‌گذاران کمک کند تا فرایندهای ارزیابی را منسجم‌تر و بهینه‌تر طراحی کنند.

با چنین رویکردی، پژوهش از طریق داشتن چشم‌انداز روشن می‌تواند جایگاه فعلی را بسنجد و برای آینده حساب‌شده‌تر برنامه‌ریزی کند؛ به عبارتی "پژوهش باید به داد پژوهش برسد". مطالعات گذشته اگرچه به نقد وضعیت گذشته، بررسی وضعیت موجود و ترسیم آینده پرداخته‌اند، اما برای مدیریت بهتر، نیاز است نوع و سطح عوامل مؤثر بر ارزیابی به‌طور دقیق شناسایی گردد. داشتن پشتوانه نظری در برنامه‌ریزی، سازماندهی، رهبری و کنترل ارزیابی کمک می‌کند تا نقاط قوت و ضعف بهتر مشخص شوند، منابع (زمان، نیرو، بودجه) به عوامل اثرگذارتر اختصاص یابند، و پایش عملکرد با رویکرد «مدیریت مبتنی بر شواهد» تحقق یابد. از این منظر، چشم‌انداز پژوهش حاضر ترسیم "شبکه مضامین ارزیابی بروندهای علمی علوم انسانی" با رویکرد ترکیبی است؛ به این معنا که با تلفیق نتایج دو مرور کیفی (تحلیل مضمون) و کمی (کتابسنجی)، مضامین و روابط آن‌ها استخراج و در قالب یک شبکه تبیین خواهد شد. هدف این است که با تحلیل نظام‌مند تعاملات مؤلفه‌های ارزیابی، الگویی نو برای درک بهتر فرایند ارزیابی و به‌کارگیری آن در تصمیم‌گیری‌ها ارائه شود. بنابراین، مسئله اساسی پژوهش آن است که: مضامین ارزیابی بروندهای علمی علوم انسانی کدام‌اند و چه روابطی میان آن‌ها برقرار است؟ پاسخ به این پرسش می‌تواند گامی در تحول رویکردهای ارزیابی، ارتقای کارآمدی سیاست‌گذاری‌ها و بهبود مدیریت بروندهای علمی علوم انسانی باشد.

پرسش‌های پژوهش

۱. در منابع مرتبط با ارزیابی بروندهای علمی علوم انسانی به چه مضامینی اشاره شده است؟
۲. ساختار مفهومی مضامین ارزیابی بروندهای علمی علوم انسانی به چه صورت است؟
۳. شبکه مضامین ارزیابی بروندهای علمی علوم انسانی چگونه است؟

چارچوب نظری

علم سنجی به عنوان شاخه‌ای از مطالعات علم، در پی سنجش و تبیین بروندهای علمی با هدف پشتیبانی از سیاست‌های علم و فناوری شکل گرفته است. منظور از برونداد (خروجی یا ستانده) علمی هر قالب، بستر و فعالیتی است که از طریق آن نتیجه مستقیم، ملموس و قابل اندازه‌گیری پژوهش‌ها و مطالعات علمی در راستای تخصص پژوهشگر به اشتراک و بهره‌برداری گذاشته می‌شود. برونداد علمی اعم از برونداد پژوهشی در نظر گرفته می‌شود (ابراهیمی و همکاران، ب-۱۴۰۲). چگونگی ارزیابی بروندهای علمی یکی از ارکان علم‌سنجی است.

هرچند علم بودن، وجه تشابه علوم انسانی با علوم دیگر است اما تفاوت ذاتی و زبان خاص هر کدام باید در ارزیابی لحاظ شود و از استفاده از یک الگوی واحد برای همه آن‌ها اجتناب گردد (Rovi ra-Este va & Orero ۲۰۱۲). به تعبیر تونی بچر^۱ (Becher, ۱۹۸۶) در کتاب "قلمروها و قبایل دانشگاهی" رشته‌ها مانند قبایلی با زبان، فرهنگ، آداب و رسوم، نمادها، اسطوره‌ها و نشانه‌های خاص هستند (نقل در: فاضلی، ۱۳۹۶)، پس هر کدام ویژگی‌ها و کارکردهای مربوط به خود را از نظر ماهیت و اهداف پژوهش، روش پژوهش، رفتار استنادی و پوشش پایگاه‌های استنادی بین‌المللی دارند. علوم انسانی نیز ارزیابی متناسب با خود را می‌طلبد.

ارزیابی بروندهای علمی علوم انسانی بر تنوع، کثرت‌گرایی، بافت‌محوری، و تأثیر اجتماعی تأکید دارد. این مبانی در تقابل با رویکردهای کمی‌گرایانه صرف قرار داشته و خواستار نظام ارزیابی چندبعدی و منعطف است تا ویژگی‌های منحصر به فرد علوم انسانی به رسمیت شناخته شود. در دهه‌های اخیر، مفهوم ارزیابی علم دچار تحول شده است؛ الگوی سنتی کمی که بر شاخص‌هایی چون استناد، ضریب تأثیر و شمارش تولیدات علمی تکیه داشت، به تدریج جای خود را به الگوهای **ترکیبی و شبکه‌ای** داده است (Hicks et al., ۲۰۱۵). با چنین دیدگاهی می‌توان در ارزیابی از نظریه شبکه-بازیگر بهره برد؛ ادعای این نظریه آن است که انسان‌ها، فناوری‌ها، اسناد و زیرساخت‌های فیزیکی همگی به عنوان بازیگرانی با اهمیت برابر در شبکه‌ها عمل می‌کنند و باید از طریق دیدگاه مفهومی و روش شناختی متناسب مورد تجزیه و تحلیل قرار گیرند (Law, ۱۹۹۲). بنابراین، شناسایی و کاربرست انواع مؤلفه‌های ارزیابی و روابط آن‌ها در سطوح مختلف باید در دستور کار سیاست‌گذاران ارزیابی قرار بگیرد.

اسناد بالادستی و سیاست‌های ارزیابی اتخاذ شده از سوی سیاست‌گذاران بر تعیین مسیر پژوهش متخصصان، از انتخاب برونداد تا ارزیابی آن مؤثر است و در ارائه چشم‌اندازها و رسالت‌ها در این زمینه ایفای نقش می‌نماید (Ant onaki s & Hooi j berg, ۲۰۰۸). از دیدگاه نظری، علم‌سنجی را می‌توان حلقه واسط میان **تولید علم و تصمیم‌گیری علمی** دانست؛ چراکه از طریق سنجش کمی و کیفی عملکرد پژوهشی، شواهد لازم برای ارزیابی اثربخشی نظام علمی را فراهم می‌کند (Bor nmann & Dani el, ۲۰۰۸). از این منظر، ارزیابی علم نه تنها ابزار سنجش بروندها، بلکه فرایندی برای **بازاندیشی**

¹ Output

² Format

³ Platform

⁴ Activity

⁵ Tony Becher

⁶ Actor-network theory

در جهت گیری های علمی و سیاستی است. پس قابل تأکید است که هدف اصلی علم سنجی ارائه خدمت به سیاست علم است. علم سنجی در ارزیابی هایی که انجام می دهد، به شناسایی عملکردها و برنامه های در دست اجرای علم و فناوری می پردازد و شرایط لازم را برای حمایت از آنها فراهم می سازد، روند صعودی و نزولی رشد دستاوردها را روشن می کند، و در تعیین مسئولیت موفقیت ها و شکست ها با توجه به هدف های تعیین شده برنامه های علم و فناوری مفید است (نوروزی چاکلی، ۱۳۹۰). بنابراین، این دو مقوله با هم ممزوج هستند و با استفاده از پژوهش های علم سنجی که به نوعی کتاب سنجی به علاوه سیاست علم است می توان در تدوین چشم انداز سیاست گذاری در ارزیابی سهمیم بود.

پیشینه پژوهش

هر حوزه علمی مستعد پژوهش در ابعاد مختلفی است و هرچه طیف ذینفعان و دامنه اثرگذاری آن گسترده تر باشد زیرشاخه های آن نیازمند بررسی تخصصی تر است. با پذیرش اهمیت علوم انسانی و نقش آن در توسعه جامعه، نگاه ها از فضای صرفاً علمی فراتر رفته و تأثیر این علوم در عرصه های مختلف مورد توجه قرار گرفته است. جایگاه علوم انسانی، ویژگی های انتشاراتی علوم انسانی، ارزیابی عملکرد پژوهشی، و ارزیابی بروندهای علمی علوم انسانی از زوایای مورد مطالعه بوده است. یکی از جنبه های قابل توجه علوم انسانی ارزیابی بروندهای علمی آن است که در سطح خرد تا کلان اثرگذار است. مرور پیشینه نشان می دهد که در منابع فارسی، عموماً با نگاه کلی تر و در منابع خارج از کشور به صورت جزئی تر به بروندهای علوم انسانی و ارزیابی آنها توجه گردیده است. در ادامه به برخی از موارد مرتبط اخیر اشاره خواهد شد.

کبیری خوزانی و همکاران (۱۴۰۴) با استفاده از فنون کتاب سنجی به بررسی کمی مفاهیم ارزیابی بروندهای پژوهشی علوم انسانی و سیر تحول آنها پرداخته اند. براساس یافته ها وجود کلیدواژه های هسته و پرکاربردی چون تأثیر پژوهش، سنجش اثر و اثر اجتماعی در سال های اخیر نشان دهنده تغییر رویکرد از ارزیابی صرفاً کمی به ترکیب روش های کمی و کیفی است. عباس پور و همکاران (۱۴۰۴) هدف پژوهش خود را طراحی و اعتباریابی الگوی کیفیت پژوهش دانشگاهی در دانشگاه های علوم انسانی و اجتماعی اعلام کرده اند. با ابتناء به داده های پژوهش، مدیریت و قوانین پژوهش به عنوان شرایط علی، تعالی کیفیت پژوهش به عنوان مقوله اصلی، ویژگی های فنی و جهانی شدن به عنوان شرایط زمینه ای، نیروی انسانی به عنوان شرط مداخله گر، نظارت و ارزیابی به عنوان راهبردهای عمل-تعامل، و در نهایت کارایی و اثربخشی و پذیرش نتایج پژوهش ها به عنوان پیامد شناسایی شدند. جنوی و مردانی (۱۴۰۳) نیز مطالعه ای برای بررسی سیاست های ارزیابی اعضای هیئت علمی حوزه علوم انسانی در دانشگاه های منتخب جهان با رویکرد سیاست پژوهی، استخراج معیارها و شاخص های ارزیابی و مقایسه آنها با ابعاد و شاخص های آیین نامه ارتقا در ایران داشته اند. معیارهای نظام ارزیابی اعضای هیئت علمی علوم انسانی دانشگاه های جهان در سه بعد آموزش، پژوهش و خدمت طبقه بندی شدند. مقایسه با آیین نامه داخلی نشان داد که تنوع شاخص ها در بعد پژوهش برای دانشگاه های جهان نسبت به آیین نامه داخلی بسیار بیشتر است. ابراهیمی درچه و همکاران (الف-۱۴۰۲) به شناسایی آسیب های ارزیابی بروندهای پژوهشی علوم انسانی با تأکید بر رشته های زبان و ادبیات و ارائه راهکارهایی برای بهبود آن پرداخته اند. ابراهیمی درچه و همکاران (ب-۱۴۰۲) در پژوهشی دیگر با توجه به کارکرد علوم انسانی و تفاوت آن با علوم دیگر، به دنبال بررسی سیاست ها و معیارهای ارزیابی بروندهای علمی

رشته‌های زبان و ادبیات از دیدگاه متخصصان بوده‌اند. ابراهیمی و میرشاه‌ولایتی (۱۴۰۲) الگوی سنجش سطح آمادگی کاربرپذیری پژوهش‌های علوم انسانی معادل با اثرات علمی، تأثیرات سیاستی، آثار اقتصادی و اجتماعی و آثار عملیاتی را ارائه نموده‌اند. نجاتی و همکاران (۱۴۰۲) با شناسایی و تبیین شاخص‌های نخبگی، چارچوبی برای کشف نخبگان علمی در حوزه علوم انسانی ارائه نموده‌اند. بخش قابل توجهی از این چارچوب به کمیت و کیفیت بروندهای علمی مربوط می‌شود. کمساری و همکاران (۱۴۰۱) معیارهای برای سنجش و رتبه‌بندی کتاب‌های تخصصی علوم انسانی ارائه داده‌اند. صفرپور و همکاران (۱۴۰۰) با کاوش تجربیات و دیدگاه‌های اعضای هیئت علمی حوزه علوم انسانی و هنر در مورد آیین‌نامه ارتقا بهترین راهکار پیشنهادی جهت رفع مجموعه مشکلات را تفکیک آیین‌نامه حوزه علوم انسانی از حوزه غیرعلوم انسانی دانسته‌اند. در پژوهش مرتبط دیگری که عصاره و همکاران (۱۳۹۸) برای طراحی مدل مفهومی و عملیاتی ارزیابی و رتبه‌بندی علمی گروه‌های آموزشی حوزه علوم انسانی دانشگاه‌ها و مؤسسات آموزش عالی ایران انجام داده‌اند نیز بعد پژوهشی و تولید علم کیفی اولویت اول را در میان ابعاد و معیارها داشته‌اند.

مولر و همکاران (Miller et al., ۲۰۲۴) کیفیت از درون^۱ یا از زبان محققان علوم انسانی برای ارزیابی کیفیت پژوهش‌های علوم انسانی را بررسی کرده‌اند. پژوهشگران علوم انسانی اغلب برای قضاوت در مورد کیفیت پژوهش‌ها به معیارهای ضمنی تکیه می‌کنند که می‌توان آن‌ها را به نوآوری، برتری پژوهشگر و تأثیر طبقه‌بندی کرد. ابراهیمی درجه و همکاران (Ebrahimi Dorcheh et al., ۲۰۲۴) به شناسایی و اولویت‌بندی معیارهای ارزیابی بروندهای علمی رشته‌های زبان و ادبیات در هشت دسته پرداخته‌اند: بستر خلق، ارائه و انتشار؛ ساختار نگارش؛ محتوا؛ تأثیر در محیط آنلاین؛ تأثیر علمی؛ تأثیر اجتماعی؛ تأثیر اقتصادی؛ و تأثیر فرهنگی. ونگ و ونگ (L. Wang & Wang, ۲۰۲۴) ارزیابی کارایی علوم انسانی با توجه به بافت را مطالعه کرده‌اند و نتایج حاکی از آن است که هر منطقه باید اقدامات متقابل متفاوتی را برای بهبود کارایی پژوهش‌های علمی اتخاذ کند. ونگ (R. Wang, ۲۰۲۳) در پژوهش خود نشان می‌دهد که مشکلات اصلی سازوکار فعلی ارزیابی عملکرد پژوهش‌های علمی علوم انسانی و اجتماعی شامل شاخص‌های ارزیابی بیش از حد کمی، معیارهای ارزیابی غیرجزئی و موضوعات ارزیابی غیرمتنوع است. وکویک و همکاران (Vučković et al., ۲۰۲۳) ارزیابی کیفیت پژوهش‌های علوم انسانی به‌صورت چندبعدی و متناسب با ماهیت این علوم را گوشزد کرده‌اند. موهانن و همکاران (Mihonen et al., ۲۰۲۳) ارزیابی تأثیر اجتماعی علوم انسانی را مورد بحث قرار داده‌اند. تأثیر اجتماعی به روش‌های مختلفی در رویه‌های مختلف ارزیابی در سطح ملی، دانشگاهی، پروژه‌ای و پژوهشی در نظر گرفته می‌شود که همه بر کار دانشگاهیان تأثیر می‌گذارد. دی جانگ و اسپاین (De Jong & Spaapen, ۲۰۲۳) ارزیابی تکوینی در علوم انسانی را شرح داده‌اند. برخلاف مدل‌های سنتی، ارزیابی تکوینی بر رویکردی جامع تأکید دارد که هم کیفیت پژوهش و هم تأثیر اجتماعی را در بر می‌گیرد و به واحدهای پژوهشی اجازه می‌دهد شاخص‌های خود را انتخاب کنند. یانگ و همکاران (Yang et al., ۲۰۲۱) تأثیرگذاری آلت‌متریکس در ارزیابی پژوهش‌های علوم انسانی را بررسی کرده‌اند. مولینی (Molini, ۲۰۲۱) با ارزیابی نوآوری در دانشکده هنر و علوم انسانی دشواری نوآوری در علوم انسانی در سطح جهانی را یادآور شدند. صنیعی و همکاران (Sani'ee et al., ۲۰۲۲) روندهای مؤثر و نیروهای محرک در آینده ارزیابی عملکرد پژوهش را

¹ Quality from within

تعیین کرده‌اند. در این مطالعه کیفی، تحلیل روند از طریق جستجو و مرور ۴۲ منبع در پایگاه‌های اطلاعاتی بین‌المللی و مصاحبه انجام شده‌است. بر اساس یافته‌ها، روندها شامل ابعاد اجتماعی (۲۷ کد)، فناوری (۳۸ کد)، اقتصادی (۳۰ کد)، زیست‌محیطی (۵ کد)، و سیاسی (۴۴ کد) است.

پیشینه نشان می‌دهد که در پژوهش‌های اولیه به ویژگی‌های علوم انسانی و نیاز به ارزیابی متناسب با آن‌ها اشاره شده اما هرچه جلوتر می‌آییم به چگونگی آن توجه بیشتری شده‌است. آنچه در این زمینه نباید فراموش گردد این است که تا سیاست مناسبی برای ارزیابی تعریف نشود برنامه‌ریزی متناسبی نیز انجام نخواهد شد. اما تعریف چنین سیاستی مستلزم شناخت مضامین، اصول و قواعد ارزیابی است؛ بنابراین، مرور آرای پژوهشگران در زمینه ارزیابی برون‌دادهای علمی علوم انسانی که موضوع این پژوهش است مهم و ضروری است و می‌تواند گامی مثبت و زیربنایی در این زمینه باشد. از دلایلی که پژوهشگران را بر آن داشته تا به طور ویژه به علوم انسانی بپردازند، تفاوت ماهیت و ویژگی رشته‌های آن با حوزه‌های دیگر حتی علوم اجتماعی و البته تفکیک این دو در فهرست رشته‌های وزارت علوم، تحقیقات و فناوری است. تفکیک نمایه استنادی هنر و علوم انسانی^۱ از نمایه استنادی علوم اجتماعی^۲ در پایگاه وب‌آوساینس و حوزه موضوعی^۳ هنر و علوم انسانی از علوم اجتماعی در پایگاه اسکوپوس نیز می‌تواند مؤید گستردگی و اهمیت بررسی جداگانه هر یک از این حوزه‌ها باشد. هرچند به علت احتساب این دو حوزه ذیل علوم نرم، کاربرد توأم آن‌ها در تعداد چشمگیری از مطالعات دیده می‌شود اما پژوهش‌هایی هم تمایز الگوهای انتشاراتی علوم انسانی و اجتماعی را برجسته نموده‌اند. برای نمونه سایورتن سن (Sivertsen, ۲۰۱۶) "الگوهای بین‌المللی سازی و معیارهای سنجش پژوهش در علوم انسانی و اجتماعی" را بررسی نموده‌است. او بیان می‌کند که الگوهای انتشار بین رشته‌های علوم انسانی و اجتماعی متفاوت است در حالی که در کشورهای مختلف برای هر رشته مشابه هستند. ورلیسن و ویرن (Verl eyzen & Wæren, ۲۰۱۶) پژوهشی با عنوان "خوشه‌بندی الگوهای انتشار نویسندگان بزرگ در علوم انسانی و اجتماعی" انجام داده‌اند. بر اساس یافته‌ها، مقاله مجله به‌طور واضح رایج‌ترین نوع انتشار در همه رشته‌ها است. علوم اجتماعی بیشتر بر زبان انگلیسی و هم‌تألیفی متکی است و علوم انسانی بر زبان‌های غیرانگلیسی و تک‌نویسندگی تکیه دارد. کیویک (Kyvik, ۲۰۰۳) اذعان می‌کند که الگوی انتشار در علوم اجتماعی به‌میزان فزاینده‌ای به الگوی انتشار در علوم طبیعی نزدیک شده‌است. این تفاوت‌ها نشان از نیاز به ارزیابی متفاوت برون‌دادهای علمی این دو حوزه دارد.

روش‌شناسی پژوهش

پژوهش حاضر از نظر نتیجه و پیامد، بنیادی و از نظر فرایند اجرا، ترکیبی (کمی و کیفی) از نوع هم‌زمان است. جهت پاسخ به پرسش‌های پژوهش نیاز به شناسایی و غربال منابع مرتبط با ارزیابی برون‌دادهای علمی علوم انسانی بود که از طریق مرور دامنه^۴

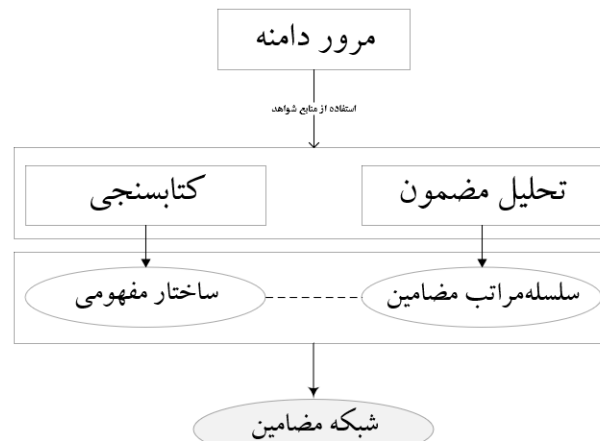
¹ Arts & humanities citation index (A&HCI)

² Social sciences citation index (SSCI)

³ Subject area

⁴ Scoping review

انجام شد. سپس منابع با استفاده از روش تحلیل مضمون^۱ گذراری گردید. به طور همزمان، ساختار مفهومی داده‌ها با استفاده از روش کتابسنگی^۲ ترسیم شد. در نهایت، با ترکیب خروجی این دو مرحله، شبکه مضامین موضوع مورد نظر ترسیم گردید. مراحل پژوهش در شکل ۱ آمده است. دلیل انتخاب هر یک از روش‌های یادشده در ادامه در چهار بخش خواهد آمد.



شکل ۱. مراحل پژوهش

مرور دامنه

مرور دامنه معمولاً برای تشخیص^۳ و روشن کردن تعاریف و مرزهای مفهومی یک موضوع یا زمینه استفاده می‌شود (Peters et al., ۲۰۱۵). چارچوبی که برای انجام یک مرور دامنه مطرح می‌شود عبارت از طرح پرسش پژوهش برای هدایت راهبرد جستجو؛ شناسایی مطالعات مرتبط؛ انتخاب مطالعات؛ ترسیم نمودار داده‌ها؛ و گردآوری، تحلیل، و ارائه گزارش روایی ادبیات (سنتز نتایج) است (Arksey & O'Malley, ۲۰۰۵). از آنجا که مرور دامنه برای پاسخ به پرسش‌های گسترده‌تر نسبت به پرسش‌های مرور سیستماتیک به کار می‌رود، اهداف و رویکرد روش‌شناختی متفاوتی داشته و برای آن از نسخه محدودتری از پروتکل PRISMA-ScR استفاده می‌شود که نام دارد (Tricco et al., ۲۰۱۸). اجزای مربوط به این پروتکل در ارتباط با این پژوهش در ادامه می‌آید.

برای بازبینی^۴ منابع، مقالاتی درباره اندازه‌گیری یا تمرکز بر ابعاد خاص ارزیابی برون‌دادهای علمی با تأکید بر علوم انسانی، که در چارچوب مفهومی پژوهش توسعه یافته‌اند (مانند علوم انسانی، ارزیابی^۴ پژوهش، برون‌داد پژوهشی و علمی) مورد نیاز بود. مقالات مجلات و همایش‌های بررسی شده در صورتی در بازبینی گنجانده می‌شدند که: از ابتدا تا تاریخ جستجو در پایگاه منتشر

¹ Thematic analysis (TA)

² Reconnaissance

³ Review

⁴ Assessment/ evaluation/ measurement

شده بودند، با روش‌های کمی یا کیفی انجام شده بودند، به زبان انگلیسی نوشته شده بودند، و موضوعات مرتبط را شامل می‌شدند. مقالات در صورتی که با چارچوب مفهومی مطالعه مطابقت نداشتند یا بر مطالعه کمی (کتابسنجی یا علم‌سنجی) یک موضوع تمرکز داشتند، برای مثال تحلیل استنادی مقالات مرتبط با یک موضوع، گنجانده نشده‌است.

برای شناسایی اسناد مرتبط بالقوه، پایگاه‌های اسکوپوس و وب‌اوساینس بدون محدودیت زمانی در خرداد ۱۴۰۴ جستجو شد. راهبرد جستجو در دو پایگاه توسط پژوهشگران توسعه داده شد و طبق جدول ۱ متناوب با قواعد هر پایگاه بهینه گردید. جستجو در سه فیلد عنوان، چکیده و کلیدواژه‌ها انجام شد. یافته‌های شناسایی شده از اسکوپوس شامل ۳۷۴ رکورد و وب‌اوساینس شامل ۳۲۵ رکورد با فرمت CSV و XLS به اکسل صادر و ادغام گردید. در این مرحله موارد تکراری حذف شد (۲۲۳ رکورد). در مرحله غربالگری ابتدا منابع بدون چکیده حذف شد (۲ رکورد)، سپس منابع دارای عنوان نامرتب (۲۲۹ رکورد) و منابع دارای چکیده نامرتب (۱۵۱ رکورد) حذف گردید، و در نهایت منابعی که تمام متن آن‌ها در دسترس نبود (۱۴ رکورد) و منابع دارای تمام متن اما نامرتب (۲۷ رکورد) حذف شد. آنچه به مرحله مشمول^۱ رسید ۵۳ منبع بود. اطلاعات مربوط به این منابع در جدول ۴ در پیوست آمده‌است.

جدول ۱. راهبرد جستجو

پایگاه	راهبرد جستجو
اسکوپوس	(TITLE-ABS-KEY (humanities) AND TITLE-ABS-KEY ("research evaluat*" OR "research assess*" OR "research measur*" OR "research output evaluat*" OR "research output assess*" OR "research output measur*" OR "scientific output evaluat*" OR "scientific output assess*" OR "scientific output measur*" OR "research output" OR "scientific output"))
وب‌اوساینس	(TS=(humanities)) AND TS=("research evaluat*" OR "research assess*" OR "research measur*" OR "research output evaluat*" OR "research output assess*" OR "research output measur*" OR "scientific output evaluat*" OR "scientific output assess*" OR "scientific output measur*" OR "research output" OR "scientific output")

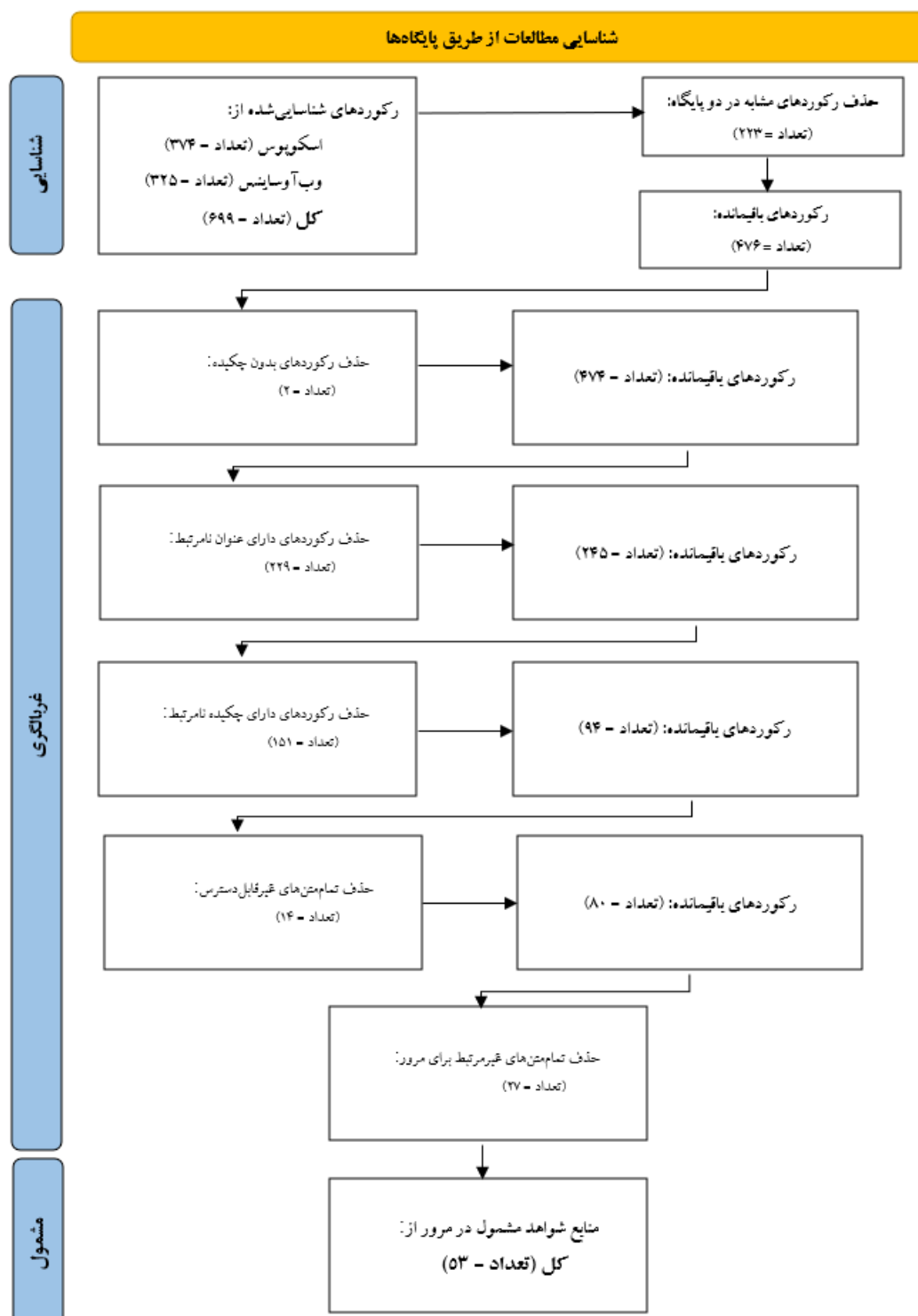
یک فرم نمودار داده‌آیجاد شد تا مشخص شود کدام متغیرها باید استخراج شود. بر اساس آن، برای هر مقاله ویژگی‌های کتاب‌شناختی (عنوان، نویسنده، سال و...) و محتوایی (کلیدواژه‌ها، چکیده، هدف، روش و یافته‌های کلیدی) گردآوری شد. در مرور فایل‌های تمام‌متن، محدودیت‌ها، ظرفیت‌ها، نقدها و پیشنهادات مربوط به ارزیابی برون‌دادهای پژوهشی علوم انسانی استخراج شد. برای سنتز نتایج، مقالاتی که در آن‌ها به شاخص یا معیار، رویکرد، روش، مدل یا چارچوب، و ابزار ارزیابی اشاره شده بود به مرحله مشمول وارد شدند. مطالعات بر اساس هدف و یافته‌ها کدگذاری گردید. مراحل شناسایی منابع همراه با جزئیات هر کدام در شکل ۲ قابل مشاهده است.

¹ Identification

² Screening

³ Included

⁴ Data charting



شکل ۲. مراحل شناسایی منابع شواهد

تحلیل مضمون

مراحل تحلیل مضمون عبارت از آشنایی با داده‌ها؛ کدگذاری؛ جستجوی مضامین (الگوی منسجم و معنادار در داده‌ها)؛ بررسی مضامین و روابط آن‌ها؛ تعریف و نام‌گذاری مضامین؛ نوشتن گزارش و روایت مضامین است. مضامین باید "برونداد" نهایی کدگذاری داده‌ها و توسعه تم‌های تکراری باشد (Clarke & Braun, ۲۰۱۳). اگر پرسش پژوهش مربوط به "ماهیت" چیزی که پژوهشگر به دنبال درک آن است باشد، آنگاه تحلیل مضمون بازتابی از براون و کلارک روشی است که برای پاسخ به آن مناسب است (Braun & Clarke, ۲۰۲۲). تحلیل داده در این روش به صورت گام به گام است به این معنی که ابتدا همه داده‌ها گردآوری و سپس به طور یکجا تحلیل می‌شود. با توجه به ماهیت تحلیل مضمون بازتابی، در این پژوهش برای رسیدن به مضامین از روش کدگذاری نظری، استقرایی و توصیفی برای توسعه ساختار سلسله‌مراتبی مضامین (مقوله، حوزه^۳ و طبقه^۴) استفاده شد.

کدگذاری منابع طی سه مرحله باز، محوری و انتخابی انجام گردید. در مرحله اول، تحلیل و شکستن داده‌های گردآوری شده که کدگذاری باز^۵ سطح اول است، پیام اصلی و کلیدی مستتر در هر جمله به عنوان واحد معنایی استخراج گردید و داده‌هایی که بار معنایی یکسانی داشتند زیر یک کد قرار داده شد. در مرحله دوم که کدگذاری محوری^۶ سطح دوم است کدها با یکدیگر مقایسه و بر اساس شباهت، طبقه‌بندی شدند. در مرحله آخر که کدگذاری انتخابی^۷ سطح سوم نامیده می‌شود این طبقات نیز مورد بازبینی و مقایسه قرار گرفته و بر اساس ویژگی‌های مشترک در سطحی انتزاعی‌تر طبقه‌بندی کلی‌تر شدند. این طبقه‌ها که اشتراکات را به تصویر می‌کشند، از نظر داخلی بیشترین تجانس و از نظر خارجی بیشترین عدم تجانس را دارا هستند. در کدگذاری، محتوای متن در یک توصیف تلگرافی خلاصه می‌شود؛ نام یا برچسب کدها و طبقه کدها باید کلمه یا عبارتی باشد که محتوای داده‌های کیفی را در بر گیرد. این نام باید مختصر اما به اندازه کافی خاص باشد تا معنادار بوده و ایده اصلی مشترک همه بخش‌های کدگذاری شده را شامل شود. در این پژوهش نیز نام هر مضمون سازنده و پایه (حوزه و طبقه) طوری انتخاب شده‌است که محتوای کدهای زیرمجموعه خود را پوشش دهد.

کتابسنجی

^۱ Reflexive thematic analysis (RTA)

^۲ Category

^۳ Domain

^۴ Taxonomy

^۵ Open

^۶ Axial

^۷ Selective

از آنجا که خواندن یک متن گسترده برای کشف الگوها و روندها زمان‌بر است از راهبرد تحلیل واژگان در کتابسنجی برای پی بردن به ساختار مفهومی و نمایش گرافیکی یک حوزه معین استفاده می‌شود. ساختار مفهومی^۱ در نرم‌افزار بیبلیومتریکس^۲ (از کتابخانه‌های زبان برنامه‌نویسی R) از ابزارهای ترسیم شبکه هم‌رخدادی^۳ واژگان (خوشه‌بندی)، نقشه مضمون^۴ (ماتریس) و تکامل مضمون^۵ در طول زمان (سیر زمانی) و به‌عنوان یک روش کمی مکمل روش‌های کیفی است (Aria et al., ۲۰۲۲). شکل ۳ مربوط به این ساختار است. از آنجا که در تحلیل مضمون تنها به درک مضامین و الگویابی آن‌ها بسنده می‌شود در این پژوهش برای بررسی عمیق‌تر انواع موضوعات پنهان در متون و نیز شناخت تحولات هرکدام در طول زمان در راستای تکمیل نتایج حاصل از تحلیل مضمون از ساختار مفهومی نیز بهره برده شد.

زودآیند ویرایش
نشده

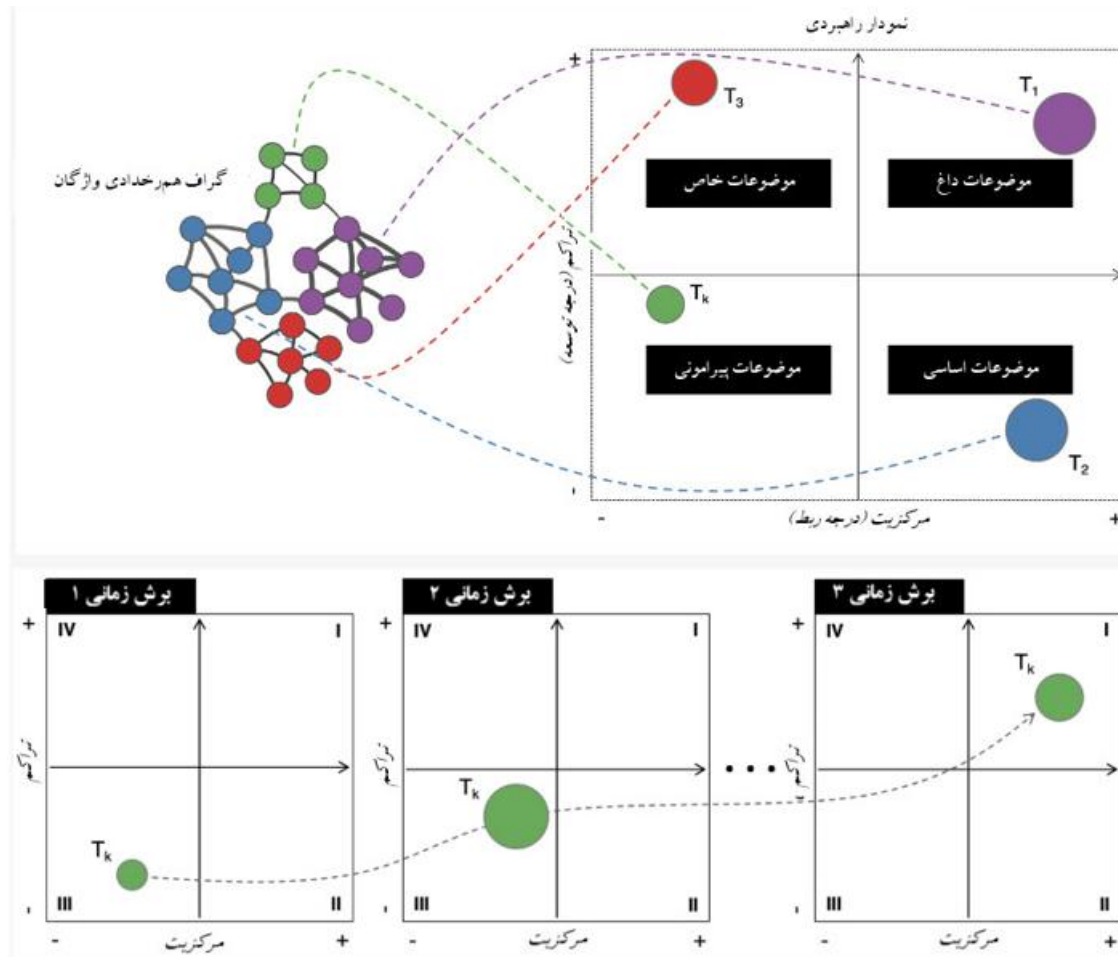
¹ Conceptual structure

² Bibliometrix

³ Co-occurrence network

⁴ Thematic map

⁵ Thematic evolution



شکل ۳. ساختار مفهومی در بیبلیومتریکیس

شبکه مضامین

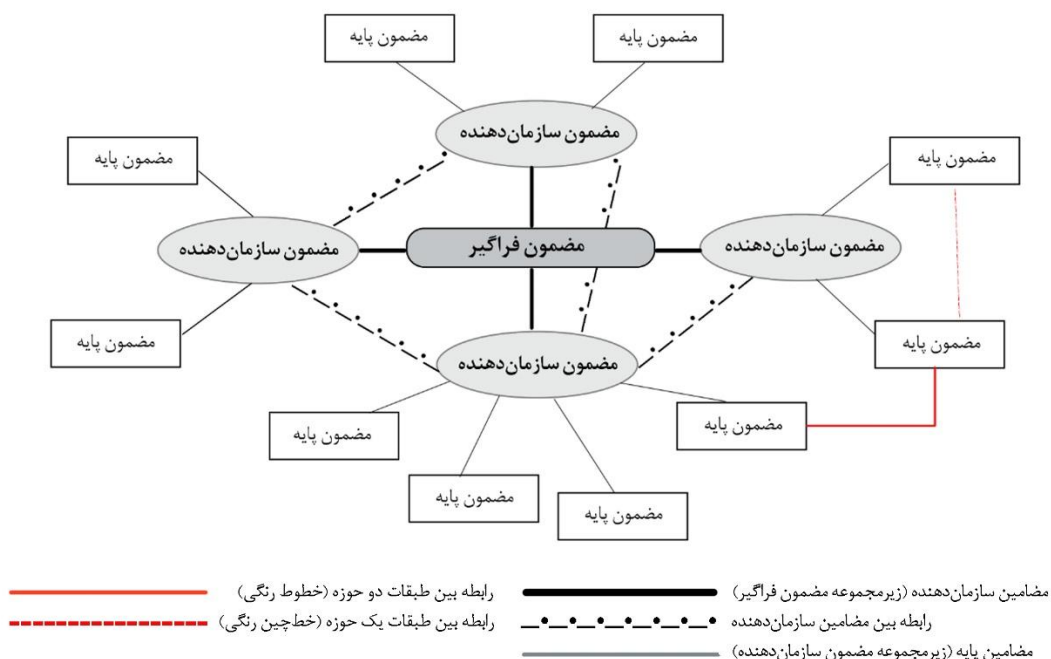
شبکه مضامین به عنوان ابزار م‌صور سازی، تارنمایی است که مضامین اصلی یک متن را خلاصه و الگوی زیربنایی آن را ارائه می‌کند. سه دسته موضوع و مضمون در این شبکه عبارت است از: ۱. مضمون پایه^۱ یا مقدمه و پایین‌ترین موضوعی که از داده‌های متنی مشتق شده است. آن‌ها با هم یک موضوع سازمان‌دهنده^۲ را نشان می‌دهند؛ ۲. مضمون سازمان‌دهنده یک موضوع درجه متوسط است که تم‌های پایه را در دسته‌هایی از موضوعات مشابه سازماندهی می‌کند. گروهی از موضوعات سازمان‌دهنده یک مضمون فراگیر^۳ را تشکیل می‌دهند؛ ۳. مضمون فراگیر: مضامین کلان هستند که استعاره‌های اصلی را در کل داده‌ها دربرمی‌گیرند.

¹ Basic theme

² Organizing theme

³ Global theme

هر مضمون فراگیر هسته یک شبکه موضوعی است. این سه دسته را می‌توان منطبق با مقوله، حوزه و طبقه در تحلیل مضمون دانست. شمای گرافیکی یک شبکه مضامین در شکل ۳ آمده‌است. گام‌های ایجاد شبکه مضمون به‌طور خلاصه عبارت از کدگذاری و شناسایی مضامین (نتیجه تحلیل مضمون)؛ ساخت و توصیف شبکه موضوعی؛ و تفسیر الگوها است (Attride-Stirling, ۲۰۰۱). در پژوهش حاضر، ساختار سلسله‌مراتب کدها حاصل از تحلیل مضمون با ساختار مفهومی حاصل از کتابسنجی ترکیب شده تا شبکه مضامین از آن پدید آید.

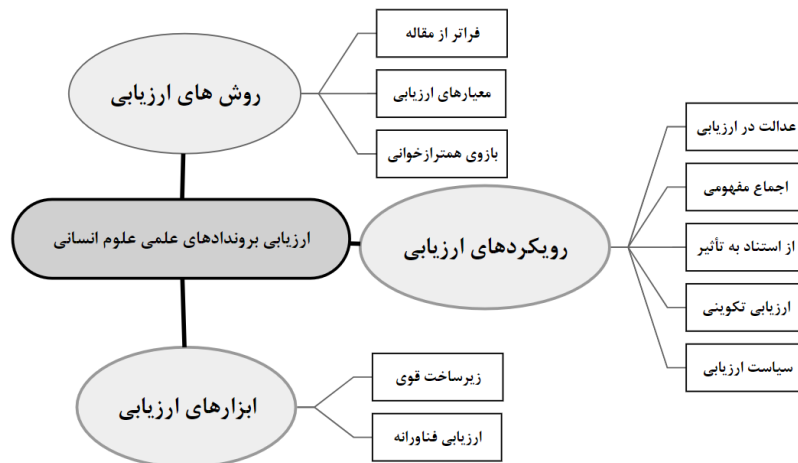


شکل ۳. شمای گرافیکی یک شبکه مضامین

یافته‌های پژوهش

پاسخ به پرسش اول پژوهش. در منابع مرتبط با ارزیابی بروندهای علمی علوم انسانی به چه مضامینی اشاره شده است؟

طی تحلیل مضمون ۵۳ منبع مرتبط با ارزیابی بروندهای علمی علوم انسانی سه طبقه رویکردهای ارزیابی، روش‌های ارزیابی و ابزارهای ارزیابی شناسایی شد. چنانچه در شکل ۴ آمده‌است رویکردهای ارزیابی شامل پنج حوزه عدالت در ارزیابی، اجماع مفهومی، از استناد به تأثیر، ارزیابی تکوینی و سیاست ارزیابی است. روش‌های ارزیابی دارای سه حوزه فراتر از مقاله، معیارهای ارزیابی و بازوی همترازخوانی است. دو حوزه زیرساخت قوی و ارزیابی فناورانه را نیز می‌توان برای طبقه ابزارهای ارزیابی برشمرد. هر یک از حوزه‌ها شامل مقوله‌هایی است که در ادامه به آن اشاره می‌شود.



شکل ۴. مضامین ارزیابی برون داده‌های علمی علوم انسانی

رویکرد ارزیابی که دیدگاه و چارچوب کلی مسیر ارزیابی را مشخص می‌کند به عنوان یکی از طبقات حاصل از تحلیل مضمون در این پژوهش شامل پنج حوزه و ۳۵ مقوله به شرح زیر است:

زیربنای ارزیابی، سیاست ارزیابی است. پژوهشگران این حوزه تأکید می‌کنند که رویکرد ارزیابی در تدوین سیاست باید این باشد که هدف ارزیابی، ارتقای علم باشد نه افراد، و معیار ارزیابی نباید به هدف تبدیل شود. ارزیابی باید ارزش ذاتی، کیفیت، اصالت، سهم و تأثیر پژوهش را در زمینه‌های محلی، ملی و بین‌المللی در نظر بگیرد و از فرهنگ پژوهشی باز، عادلانه و انسان‌محور حمایت کند. ارزیابی کمی نباید باعث فرهنگ سرعت در بین پژوهشگران بشود و ارزیابی کیفی با رویکرد پایین به بالا و متناسب با هر رشته باشد.

تحقق ارزیابی تکوینی منوط به افزایش سواد معیارها و استفاده از نظرات نهادهای ملی، استفاده از نظرات متخصصان کتابداری در تدوین معیارهای ارزیابی تک‌نگاشت‌ها به عنوان برون‌دادهای مهم در علوم انسانی، استفاده از نظر پژوهشگران علوم انسانی، مشارکت فیلسوفان علم در مباحث سیاست علمی و فراتر از آن، داشتن یک گفت‌وگو باز و عمومی در تدوین و توسعه معیارهای ارزیابی است. هم‌نظر شدن ذینفعان بر سر ارزش‌های علوم انسانی جهت سنجش بهتر پژوهش‌ها از جمله رویکردهای ارزیابی است. اجماع بر سر مفهوم تأثیر پژوهش، تعریف دقیق ارزش فرهنگی برای ارزیابی بهتر پژوهش، تعریف کیفیت پژوهش از نگاه پژوهشگران علوم انسانی و مفهوم‌سازی آن جهت تدوین معیارهای ارزیابی در این راستا راهگشاست.

روند پژوهش‌ها نشان‌دهنده آن است که در تدوین و توسعه معیارهای ارزیابی باید از استناد و به عبارت دیگر، از معیارهای کمی به سمت ارزیابی تأثیر حرکت کرد؛ چرا که به توجه به بهره‌وری پژوهشی، استنادات و اعتبار آکادمیک به عنوان تأثیر پژوهش اشاره زیادی شده است. ارزیابی تأثیر اجتماعی به عنوان کیفیت پژوهش میان‌رشته‌ای، با لحاظ کردن تعاملات مولد و تغییراتی که ایجاد می‌کنند و برآورد سهم پژوهش در تأثیر اجتماعی قابل برنامه‌ریزی است. این رویه نیازمند روش‌های ارزیابی تأثیر مبتنی بر بافت و استفاده از داده‌های غیرسیستماتیک به عنوان شواهد تأثیر در ارزیابی است.

البته به دلیل متفاوت بودن برون‌دادهای علوم انسانی از نظر ماهیت، موضوع، روش و الگوی انتشار لازم است ارزیابی آن‌ها

متناسب و عادلانه باشد. چنانچه در منابع اشاره شده است سیاستگذاری ارزیابی با توجه به بافت هر کشور، توسعه چارچوب ارزیابی پژوهش مبتنی بر ارزش‌ها و نیازهای واقعی هر رشته و توجه به الگوی انتشار در هر رشته در هنگام ارزیابی عامل عدالت در ارزیابی خواهد بود. تحقق این امر با متکی نبودن ارزیابی در کشورهای غیرانگلیسی‌زبان به پایگاه‌های استنادی بین‌المللی انگلیسی‌محور، لحاظ کردن مجلات محلی در پلتفرم‌های ارزیابی و کاهش فشار پژوهش بین‌المللی، مهم شمردن پژوهش به زبان محلی و پوشش زبانی گسترده‌تر در پایگاه‌های استنادی کنونی امکان‌پذیر است.

در مقابل، روش نشان‌دهنده تکنیک‌ها و مراحل عملی چگونگی ارزیابی است. روش ارزیابی به‌عنوان یکی از طبقات حاصل از تحلیل مضمون در این پژوهش شامل سه حوزه و ۳۸ مقوله است:

لازم است تنوع برون‌دادهای علمی به‌عنوان واحد تحلیل یا کانال ارتباطی، در ارزیابی علوم انسانی به رسمیت شناخته شود؛ به‌ویژه فراتر از مقاله باید اهمیت زیادی برای کتاب با کاربرد بالا در علوم انسانی قائل شد. طبقه‌بندی انتشارات از زوایای مختلف (مانند تعداد نویسنده) و وزن‌دهی برون‌دادهای علمی متناسب با ماهیت هر رشته همراه با بکارگیری معیار مرتبط با هر نوع واحد تحلیل و طبقه انتشار، تدوین معیارهای ارزیابی ملی تنظیم با شیوه‌های بین‌المللی و عدم وابستگی معیارهای آن‌ها به کیفیت و پوشش پایگاه‌ها ازجمله راه‌های پیشنهادی در ارزیابی متناسب است.

با وجود رویکرد از استناد به تأثیر در ارزیابی و تغییر معیارهای ارزیابی تأثیر از سستی به جدید، روش‌های ارزیابی مرتبط با آن باید به کار گرفته شود. تدوین معیارهای ارزیابی مبتنی بر مفاهیم ذهنی کیفیت به‌صورت دقیق و منحصر به بافت هر رشته و ارزیابی کیفیت متمرکز بر متن باید بیش از پیش مورد توجه قرار بگیرد. توجه به تفاوت رشته‌ها در استفاده از آلت‌متریکس یا سنج‌های جایگزین و شاخص‌های وب و مبتنی بر عملکرد برای ارزیابی برون‌دادهای رسمی و غیررسمی یا غیراستاندارد و ارزش‌های کمتر ملموس و غیرقابل اندازه‌گیری پژوهش را نباید از نظر دور داشت. این روش نیازمند استفاده همزمان و مکمل معیارهای چندبعدی اولویت‌بندی‌شده و هم‌ترازخوانی است. چنانکه از منابع برمی‌آید توجه روزافزونی به معیارهای مسئولانه در ارزیابی شده و در آینده شاهد تدوین پررنگ‌تر معیارهای سازگار با علم باز و دموکراتیک و شفاف‌تر شدن هم‌ترازخوانی در شرایط دسترسی باز خواهیم بود. این‌ها مستلزم استفاده از معیارهای ارزیابی پس از بررسی اعتبار و پایایی و تأثیر آن‌ها بر پژوهش جهت کاهش فاصله نگرش به معیارهای ارزیابی در اصل و عمل است.

حال این پرسش مطرح می‌شود که با وجود رویکردها و روش‌های ارزیابی مطرح‌شده چه ابزاری برای ارزیابی برون‌دادهای علمی علوم انسانی لازم است. این طبقه شامل دو حوزه و نه مقوله است. ایجاد پایگاه داده ملی به‌ویژه پایگاه داده باز با رویکرد پایین به بالا برای رفع کاستی‌های موجود در پایگاه‌های فعلی و تدارک سیستم ارزیابی شفاف و دسترسی آزاد منابع شاخص‌ها برای ارزیابی متناسب با علوم انسانی جزئی از یک زیرساخت قوی در ارزیابی است. با پیشرفت روزافزون فناوری، ارزیابی فناورانه با مواردی چون طراحی سیستم هوشمند گردآوری داده‌های ارزیابی و استفاده از فنون جدید (ازجمله یادگیری ماشین) در تخمین کیفیت پژوهش نیز مطرح است.

پا سخ به پر سش دوم پژوهش. ساختار مفهومی مضمین ارزیابی برون دادهای علمی علوم انسانی به چه صورت است؟

برای پا سخ به این پرسش چکیده مقالات تحلیل شد، چرا که عنوان ممکن است جامعیت کامل را نداشته باشد و از طرف دیگر، کلیدواژه‌ها داده‌ازدست‌رفته قابل توجهی داشتند (۲۰,۷۵ در صد کلیدواژه‌های نویسنده و ۶۹,۸۱ در صد کلیدواژه‌های پایگاه). پس از بررسی حالت‌های مختلف ترکیب، حالت دوتایی^۲ که معنادارتر و دارای تعداد کلمات قابل قبولی بود برای تحلیل انتخاب شد. برای دقت بیشتر نتایج، دو فایل عبارت‌های حذفی^۳ و عبارت‌های مترادف تهیه و اعمال شد. بنا بر پیشنهاد نرم‌افزار و بررسی حالت‌های مختلف، یک برش زمانی^۴ انجام و بررسی گردید؛ بازه‌های زمانی ۲۰۰۱ تا ۲۰۱۸ و ۲۰۱۹ تا ۲۰۲۵. این تقسیم‌بندی می‌تواند به ترتیب نشان‌دهنده کندی و شتاب تغییرات مباحث مربوط به ارزیابی برون دادهای علمی علوم انسانی در این دو بازه باشد. در ادامه نقشه‌ها، جداول و شبکه‌های مرتبط با هر بازه خواهد آمد.

نقشه^۵ مضمون مربوط به بازه زمانی سال‌های ۲۰۰۱ تا ۲۰۱۸ در شکل ۵ آمده است. برای درک بهتر نقشه لازم است ابتدا ساختار آن مشخص شود. نقشه مضمون دارای دو محور است: محور افقی مربوط به چگالی و نشان‌دهنده قدرت داخلی هر یک از خوشه‌ها و معیار توسعه‌یافتگی موضوعات است؛ و محور عمودی مربوط به مرکزیت و نشان‌دهنده تعامل یک خوشه با خوشه‌های دیگر و معیار اهمیت موضوعات است. این دو محور نقشه را به چهار ربع تقسیم می‌کنند: ربع بالا سمت راست موضوعات محرک هستند که به خوبی توسعه یافته، ارتباط خوبی دارند و برای ساختار یک زمینه پژوهشی مهم هستند؛ ربع بالا سمت چپ موضوعات بسیار تخصصی^۶ هستند که ارتباط کمتری دارند و اصطلاحاً منزوی هستند؛ ربع پایین سمت راست موضوعات اساسی^۷ هستند که بالغ، محوری و عرضی هستند؛ ربع پایین سمت چپ موضوعات در حال ظهور^۸ یا رو به زوال و اصطلاحاً موضوعات پیرامونی هستند. هر یک از خوشه‌ها کلمات متعددی را در بر می‌گیرد که نام خوشه برگرفته از کلمه دارای بیشترین فراوانی و یا کلمه‌ای است که مفاهیم دیگر را پوشش دهد.

¹ N-grams

² Bigrams

³ Stop list

⁴ Time slice

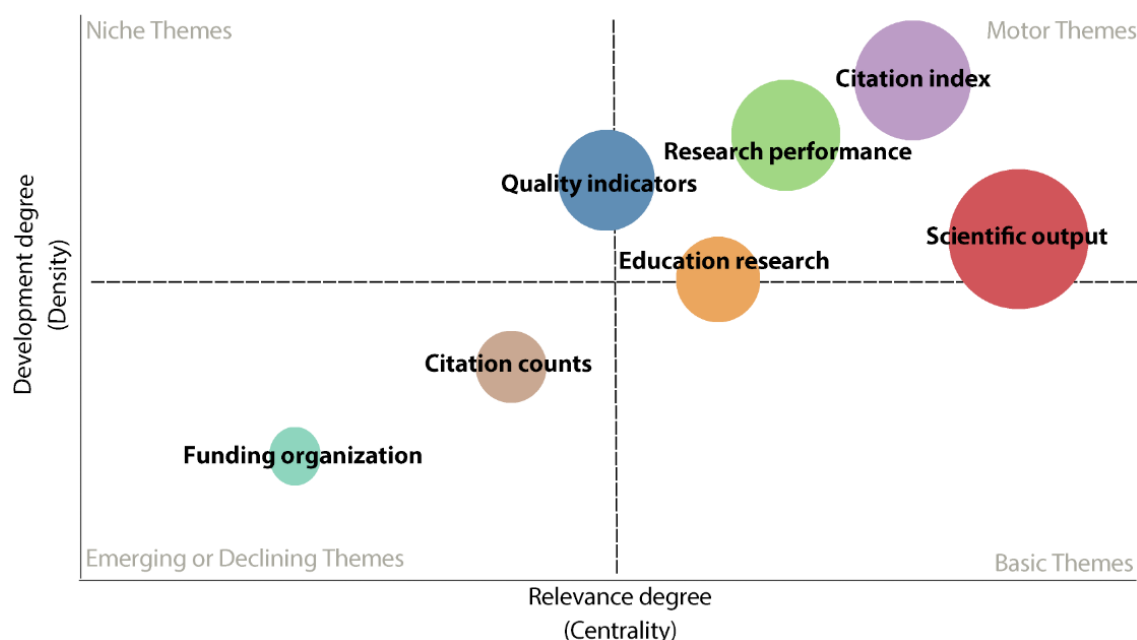
⁵ Map

⁶ Motor

⁷ Niche

⁸ Basic

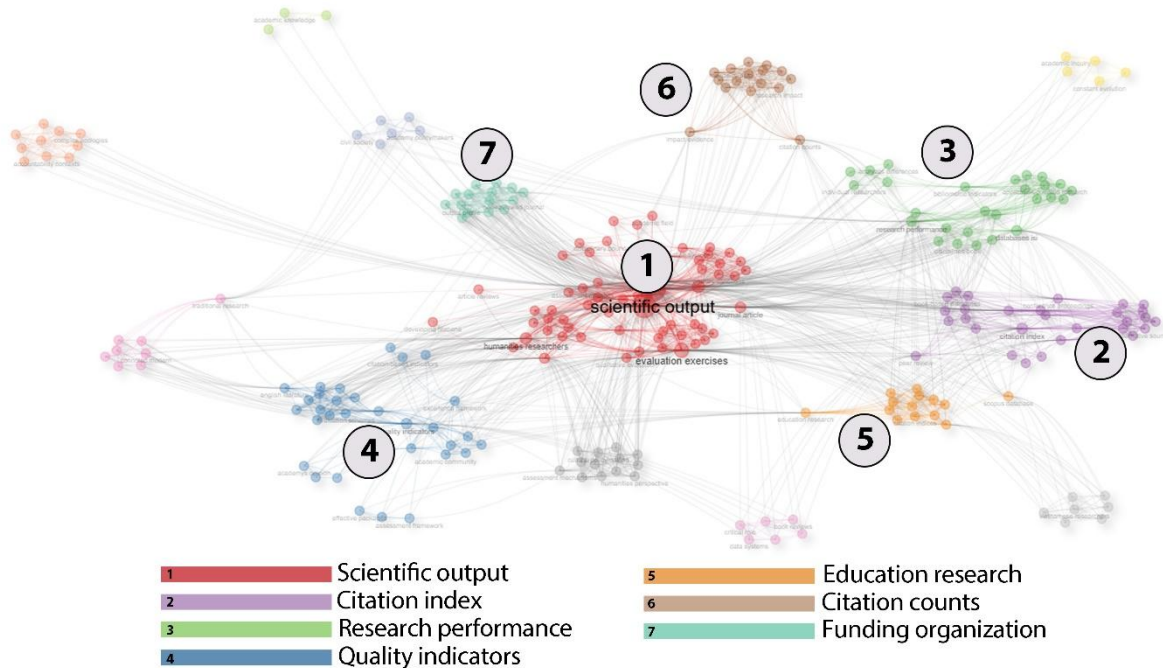
⁹ Emerging



شکل ۵. نقشه مفهومی بازه زمانی ۲۰۰۱ تا ۲۰۱۸

در نقشه مفهومی در شکل ۵ اکثر خوشه‌ها در ربع بالا سمت راست هستند و دو مورد در ربع پایین سمت چپ قرار دارند؛ بنابراین موضوعات مرتبط با ارزیابی برون داده‌های علمی علوم انسانی بین سال‌های ۲۰۰۱ تا ۲۰۱۸ جزء مضامین محرک بوده، به خوبی توسعه یافته و ارتباط و تعامل بالایی با هم دارند. اما چنانکه مشخص است و شکل شماره ۷ هم مؤید آن خواهد بود این موضوعات گرایش به تبدیل شدن به موضوعات اساسی و تخصصی دارند. این را می‌توان نشان از جافتادگی موضوعات در پژوهش‌های این حوزه و بالتبع توجه گسترده‌تر به آن‌ها در عرصه ارزیابی علوم انسانی دانست.

شبکه مربوط به بازه زمانی ۲۰۰۱ تا ۲۰۱۸ در شکل ۶ آمده است. بیشترین مرکزیت مربوط به خوشه Scientific output و بیشترین چگالی مربوط به خوشه Citation index است. چنانکه در جدول ۲ آمده است خوشه Scientific output مفاهیم اساسی و مهم در ارزیابی علوم انسانی را شامل می‌شود و اکثر آن‌ها طبق شکل ۷ در سال‌های ۲۰۱۹ تا ۲۰۲۵ به ربع پایین سمت راست یعنی مضامین اساسی منتقل می‌شود. بالا بودن چگالی در خوشه Citation index می‌تواند به دلیل نزدیکی و همبستگی زیاد مضامین آن باشد؛ از جمله آن‌ها پایگاه‌های کتابسنجی، تأثیر مبتنی بر استناد و ضریب تأثیر است.



شکل ۶. شبکه بازه زمانی ۲۰۰۱ تا ۲۰۱۸

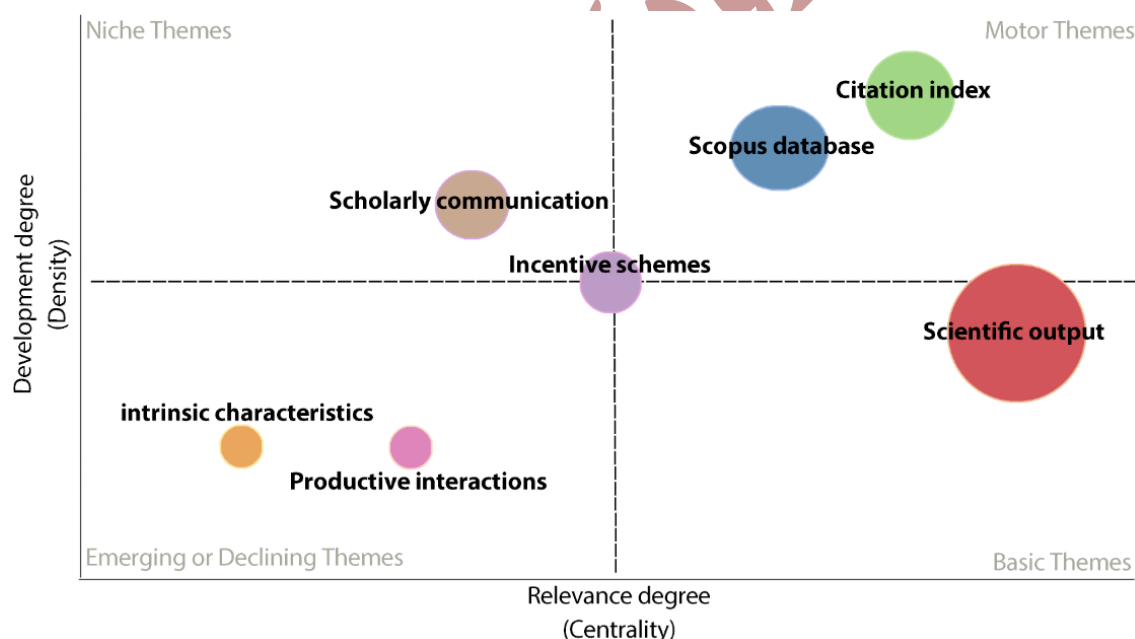
اطلاعات بیشتر مربوط به هفت خوشه اصلی شبکه شکل ۶ طبق نقشه مفهومی شکل ۵ در جدول ۲ می‌آید. به منظور خوانایی بهتر نقشه، رنگ هر خوشه در جدول به رنگ خوشه‌های نقشه و شبکه درآمده است.

جدول ۲. کلمات نقشه و خوشه بازه زمانی ۲۰۰۱ تا ۲۰۱۸

خوشه	مرکزیت	چگالی	کلمات (فراوانی)
Scientific output	9	6	Scientific output (20), Research evaluation (18), Humanities disciplines (11), Humanities researchers (6), Scholarly book (4), Journal article (4), Evaluation criteria (4), Bottom-up evaluation (3), Impact indicator (3), SSH researchers (3), University press (3), publication patterns (2), Evaluation system (2), Qualitative evaluation (2), Alternative metrics (2), Evaluation exercises (1)
Citation index	8	9	Citation index (4), Citation coverage (3), Google scholar (2), Peer review (2), Book-based disciplines (2), Google books (2), Book citations (2), Citation-based impact (2), Conference proceedings (2), Impact factors (2), Journal-based citation (2), Relevant outputs (2)
Research performance	7	8	Research performance (5), Databases ISI (3), Citation data (2), Bibliometric databases (2), Bibliometric indicators (2), Scientific information (2)
Quality indicators	5	7	Quality indicators (5), Research quality (3), Research excellence (3), English literature (2), Excellence framework (2), Intrinsic characteristics (2)

Education research (2), Scopus database (2), Neoliberal approach (2)	5	6	Education research
Citation counts (2), Impact evidence (2)	4	4	Citation counts
Funding organization (2)	2	2	Funding organization

نقشه مضمون مربوط به بازه زمانی سال‌های ۲۰۱۹ تا ۲۰۲۵ در شکل ۷ قابل مشاهده است. بیشترین جابجایی از مضامین محرک، محرک-اساسی و محرک-تخصصی به ربع‌های دیگر بوده است. در وهله دوم تعداد زیادی از کلمات در بازه زمانی اول نبوده ولی در بازه زمانی دوم اضافه شده است. از طرف دیگر، در بازه زمانی دوم ربع محرک و محرک-اساسی بیشترین دریافت را از ربع‌های دیگر داشته است. تعداد زیادی از کلمات به‌ویژه از ربع محرک و محرک-اساسی نیز حذف شده‌اند. نکته قابل توجه دیگر، افزایش مضامین نوظهور در ربع پایین سمت چپ است. ارزیابی تأثیر از طریق تعاملات مولد در علوم انسانی و از طرف دیگر، ویژگی‌های ذاتی علوم انسانی برای ارزیابی متناسب با ماهیت این علوم اخیراً توجه بیشتری به خود جلب کرده است.



شکل ۷. نقشه مفهومی بازه زمانی ۲۰۱۹ تا ۲۰۲۵

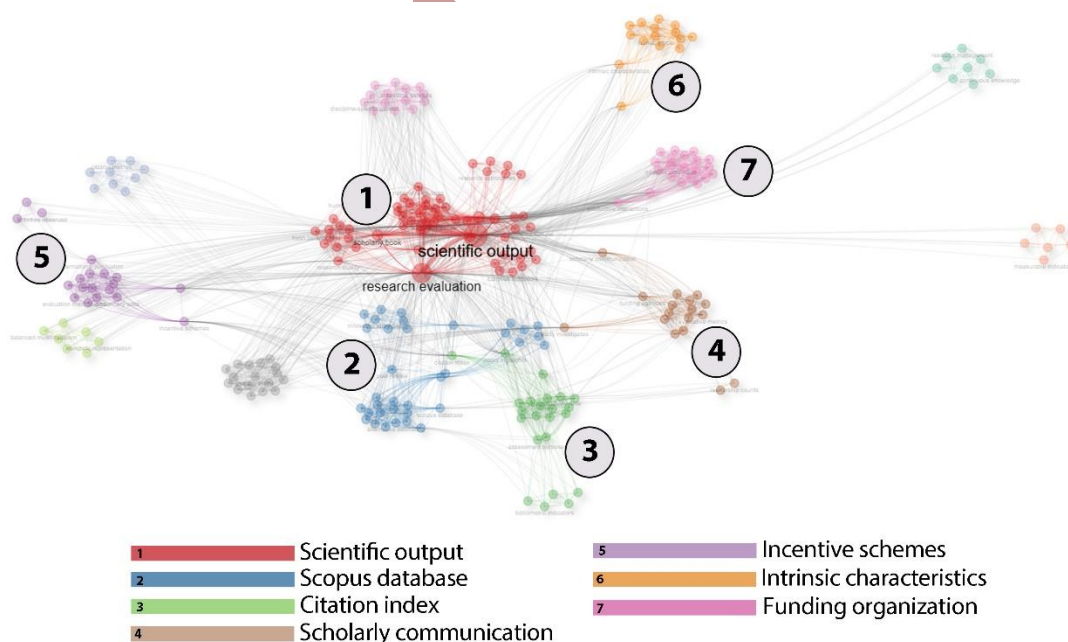
خوشه‌ها در گذر به این بازه زمانی تغییراتی داشته‌اند: ۱. برخی کلمات در همان خوشه یا همان ربع باقی مانده و این نشان‌دهنده حفظ جایگاه آن‌ها در ادبیات این حوزه است. کلمات مربوط به ارزیابی استنادی نیاز به در نظر داشتن انواع برون‌دادهای علمی فراتر از مقاله در ارزیابی علوم انسانی در این دسته است. ۲. برخی کلمات به خوشه دیگر در همان ربع یا خوشه‌ای در ربع دیگر منتقل شده است. ۳. برخی کلمات حذف شده‌اند. ۴. برخی کلمات جدید اضافه شده است. مصادیق این چهار دسته در جدول ۳ آمده است.

جدول ۳. تغییرات کلمات نقشه بازه زمانی ۲۰۱۸-۲۰۰۱ و ۲۰۲۵-۲۰۱۹

نوع مضمون	خوشه بازه ۲۰۱۹-۲۰۲۵	نوع مضمون	خوشه بازه ۲۰۰۱-۲۰۱۸	کلمات
Motor	Scopus database	-	-	Academic excellence, Quantitative indicators
Motor	Citation index	-	-	Assessment purposes, Mendeley data, Policy highlights, Research impact
Niche-Motor-Emerging-Basic	Incentive schemes	-	-	Evaluation methods, Incentive schemes, International publication
Niche	Scholarly communication	-	--	Funding agencies, Scholarly communication, Study investigates, Covariates project
Emerging	Productive interactions	-	--	Impact assessment, Productive interactions
Basic-Motor	Scientific output	-	-	National languages, Societal impact, Current discussion, Evaluation agencies
Emerging	Intrinsic characteristics	-	-	Research findings
-	-	Motor-Basic	Scientific output	Alternative metrics, Bottom-up evaluation, Evaluation exercises, Evaluation system, Journal article, Qualitative evaluation, Publication patterns, SSH researchers, Impact indicator
-	-	Motor	Research performance	Bibliometric indicators, Citation data, Databases ISI, Research performance, Scientific information
-	-	Motor	Citation index	Book citations, Book-based disciplines, Citation coverage, Conference proceedings, Google books, Google scholar, Impact factors, Journal-based citation, Relevant outputs
-	-	Motor-Basic	Education research	Education research, Neoliberal approach
-	-	Motor-Niche	Quality indicators	English literature, Quality indicators
-	-	Emerging	Funding organization	Funding organization
-	-	Emerging	Citation counts	Impact evidence
Niche	Scholarly communication	Emerging	Citation counts	Citation counts
Motor	Citation index	Motor	Research performance	Bibliometric databases

Motor	Citation index	Motor	Citation index	Citation index, Citation-based impact, Peer review
Basic-Motor	Scientific output	Motor-Basic	Scientific output	Evaluation criteria, Humanities disciplines, Humanities researchers, Research evaluation, Scholarly book, Scientific output, University press
Motor	Scopus database	Motor-Basic	Education research	Scopus database
Basic-Motor	Scientific output	Motor-Niche	Quality indicators	Research quality
Emerging	Intrinsic characteristics	Motor-Niche	Quality indicators	Intrinsic characteristics
Motor	Scopus database	Motor-Niche	Quality indicators	Excellence framework, Research excellence

شبکه مربوط به بازه زمانی ۲۰۱۹ تا ۲۰۲۵ در شکل ۸ آمده است. بیشترین مرکزیت را خوشه scientific output و بیشترین چگالی را خوشه citation index دارد. نتایج این خوشه با خوشه مربوط به بازه ۲۰۰۱ تا ۲۰۱۸ یکسان است؛ شاید کلمات داخل خوشه‌ها تغییراتی کرده باشد اما برچسب آن‌ها براساس پربسامدترین کلمه آن تعیین می‌شود.



شکل ۸. شبکه بازه زمانی ۲۰۱۹ تا ۲۰۲۵

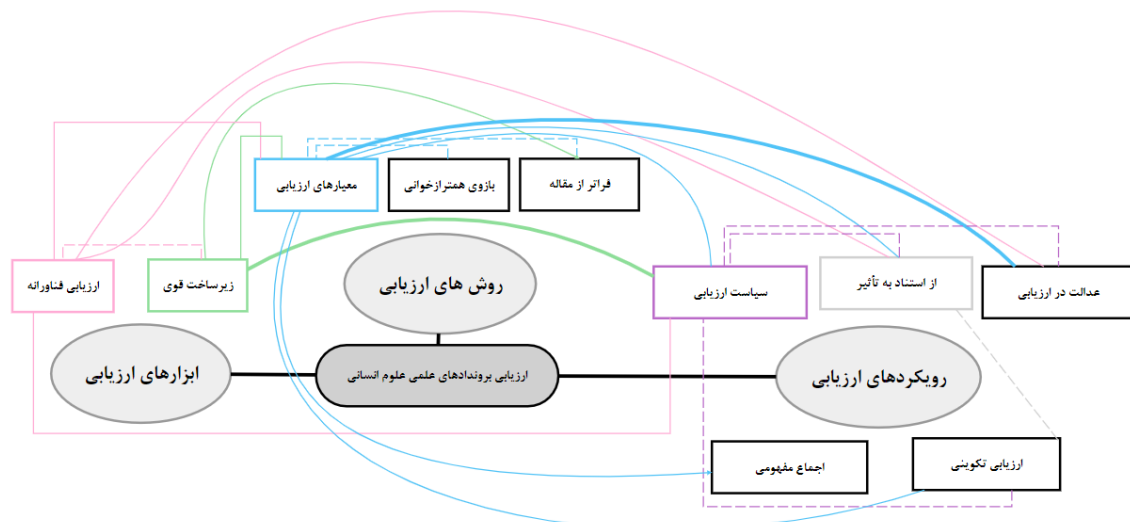
اطلاعات بیشتر مربوط به هفت خوشه اصلی شبکه شکل ۸ طبق نقشه مفهومی شکل ۷ در جدول ۴ می‌آید. به منظور خوانایی بهتر نقشه، رنگ هر خوشه در جدول به رنگ خوشه‌های نقشه و شبکه درآمده است.

جدول ۴. کلمات نقشه و خوشه بازه زمانی ۲۰۱۹ تا ۲۰۲۵

خوشه	مرکزیت	چگالی	کلمات (فراوانی)
Scientific output	7	3	Scientific output (17), Research evaluation (12), Humanities disciplines (9), Societal impact (5), Scholarly book (4), University press (4), Research quality (3), Evaluation criteria (3), Evaluation agencies (2), National languages (2), Current discussion (2), Humanities researchers (2)
Scopus database	5	6	Scopus database (2), Peer review (2), Academic excellence (2), Excellence framework (2), Quantitative indicators (2), Research excellence (2)
Citation index	6	7	Citation index (3), Mendeley data (2), Research impact (2), Assessment purposes (2), Bibliometric databases (2), Citation-based impact (2), Policy highlights (2)
Scholarly communication	3	5	Scholarly communication (2), Citation counts (2), Covariates project (2), Funding agencies (2), Study investigates (2)
Incentive schemes	4	4	Incentive schemes (2), Evaluation methods (2), International publication (2)
Intrinsic characteristics	1	1.5	Intrinsic characteristics (2), Research findings (2)
Productive interactions	2	1.5	Productive interactions (2), Impact assessment (2)

پاسخ به پرسش سوم پژوهش. شبکه مضامین ارزیابی برون‌دادهای علمی علوم انسانی چگونه است؟

چنانکه در شکل ۱ هم آمد در مرحله سوم، حاصل ترکیب نتایج تحلیل مضمون و کتابسجی شبکه مضامین است. سلسله‌مراتب شبکه مضامین در شکل ۹ مربوط به سلسله‌مراتب مضامین در شکل ۴ است. مضمون فراگیر، مضامین سازمان‌دهنده و مضامین پایه در شبکه منطبق با طبقه، حوزه و مقوله در تحلیل مضمون است. برای ترسیم روابط حوزه‌ها در سطح طبقه از کلمات بخش کتابسجی استفاده شد. در این مرحله تک‌تک کلمات خوشه‌های مرحله دوم در مقوله‌های مرحله اول جستجو و مشخص شد که متعلق به چه طبقاتی است. سپس بین طبقات دارای کلمه مشترک ارتباط برقرار گردید. اگر دو طبقه متعلق به یک حوزه بوده باشد رابطه آن‌ها با خط چین نشان داده شده و اگر دو طبقه متعلق به دو حوزه بوده باشد رابطه آن‌ها با خط نشان داده شده است. ضخامت خطوط **رنگی** نشان‌دهند تکرار رابطه بین دو طبقه است. رنگ خطوط فقط برای تفکیک بیشتر شبکه است و معنای خاصی را متبادر نمی‌کند. این روابط در شکل ۹ به تصویر درآمده است.



شکل ۹. شبکه مضامین ارزیابی برون داده‌های علمی علوم انسانی

در هر سه طبقه رویکردها، روش‌ها و ابزارهای ارزیابی بین حوزه‌ها ارتباط داخلی وجود دارد. در طبقه رویکردها حوزه سیاست ارزیابی بیشترین ارتباط را با حوزه‌های دیگر دارد و این نشان از زیربنایی بودن آن است. در طبقه روش‌ها حوزه معیارهای ارزیابی با دو حوزه دیگر در ارتباط است و این می‌تواند به دلیل اهمیت چگونگی ارزیابی باشد.

دو ارتباط قوی‌تر وجود دارد؛ ۱. ارتباط حوزه عدالت در ارزیابی از طبقه رویکردها و حوزه معیارهای ارزیابی از طبقه روش‌ها که گویای نیاز به ارائه و استفاده عادلانه معیارهای ارزیابی است، و ۲. حوزه سیاست ارزیابی از طبقه رویکردها و حوزه زیرساخت قوی از طبقه ابزارها که هر دو ماهیت زیربنایی دارند.

طبق جدول ۵ بیشترین ارتباط را حوزه معیارهای ارزیابی (فراوانی ۹) و کمترین ارتباط را حوزه بازوی همتر از خوانی (فراوانی ۱) دارد. فراوانی ارتباط در حوزه‌های دیگر به ترتیب عبارت است از: سیاست ارزیابی (۶)، ارزیابی فناوریانه (۵)، زیرساخت قوی (۴)، از استاندارد به تأثیر (۴)، عدالت در ارزیابی (۳)، ارزیابی تکوینی (۳)، فراتر از مقاله (۲)، و اجماع مفهومی (۱). بررسی‌ها نشان داد حوزه‌های دارای ارتباط بیشتر مربوط به مضامین با چگالی و مرکزیت بالاتر (مضامین محرک) در نقشه مضامین هستند.

جدول ۵. ماتریس روابط شبکه مضامین ارزیابی برون داده‌های علمی علوم انسانی

راهنما		
* بدون رابطه		
* رابطه بین طبقات دو حوزه		
* رابطه قوی بین طبقات دو حوزه		
ابزارهای ارزیابی	روش‌های ارزیابی	رویکردهای ارزیابی

- رابطه بین طبقات یک حوزه									
ارزیابی فناوریانه	زیرساخت قوی	بازوی همتراز خوانی	معیارهای ارزیابی	فراتر از مقاله	سیاست ارزیابی	ارزیابی تکوینی	از استاندارد به تأثیر	اجماع مفهومی	عدالت در ارزیابی
*	*	*	*	*	-	*	*	*	*
رویکردهای ارزیابی	*	*	*	*	*	*	*	*	عدالت در ارزیابی
	*	*	*	*	*	*	*	*	اجماع مفهومی
	*	*	*	*	-	-	*	*	از استاندارد به تأثیر
	*	*	*	*	-	*	*	*	ارزیابی تکوینی
	*	*	*	*	*	*	*	*	سیاست ارزیابی
روش‌های ارزیابی	*	*	-	*	*	*	*	*	فراتر از مقاله
	*	*	-	*	*	*	*	*	معیارهای ارزیابی
	*	*	*	*	*	*	*	*	بازوی همتراز خوانی
ابزارهای ارزیابی	-	*	*	*	*	*	*	*	زیرساخت قوی
	*	*	*	*	*	*	*	*	ارزیابی فناوریانه

بحث و نتیجه‌گیری

طی تحلیل مضمون منابع مرتبط با ارزیابی برون‌دادهای علمی علوم انسانی، سه طبقه اصلی شامل رویکردهای ارزیابی، روش‌های ارزیابی و ابزارهای ارزیابی شناسایی شد. رویکردهای ارزیابی دربرگیرنده پنج حوزه عدالت در ارزیابی، اجماع مفهومی، از استاندارد به تأثیر، ارزیابی تکوینی و سیاست ارزیابی است. روش‌های ارزیابی شامل سه حوزه فراتر از مقاله، معیارهای ارزیابی و بازوی همتراز خوانی بوده و ابزارهای ارزیابی نیز با دو حوزه زیرساخت قوی و ارزیابی فناوریانه شناخته شدند. این ساختار سه‌سطحی نشان می‌دهد که ارزیابی برون‌دادهای علمی علوم انسانی نه تنها به گردآوری داده‌ها و شاخص‌ها محدود نمی‌شود بلکه مستلزم توجه به سیاست‌ها، اصول اخلاقی و چارچوب‌های مفهومی است که کیفیت و تأثیر پژوهش را تضمین کند.

یافته‌های حاضر همسو با بسیاری از مطالعات پیشین است با این تفاوت که موارد قبلی هرکدام از زاویه‌ای خاص به موضوع ارزیابی برون‌دادهای علمی علوم انسانی پرداخته‌اند اما پژوهش پیش رو نگاهی کلی و چندجانبه داشته‌است. پژوهشی که از نظر یافته‌ها قرابت بیشتری با مطالعه حاضر دارد مقاله عباس‌پور و همکاران (۱۴۰۴) است. به طور خاص، پررنگ بودن سیاست ارزیابی در رویکردهای ارزیابی نشان از نیاز به تغییر نگاه در ارزیابی در علوم انسانی دارد چرا که هر گاه سیاست و رویکردی جوابگوی وضعیت موجود نباشد باید تحولی ایجاد شده و سیاست و رویکرد جدیدی اتخاذ شود. تعدد پژوهش‌هایی که به این یافته دست پیدا کرده‌اند (علی‌نژاد چمازکتی و میرحق‌جو لنگرودی، ۱۴۰۲؛ مرزبان و همکاران، ۱۴۰۲؛ De Jong & Spaapen, ۲۰۲۳) نشان از اهمیت و ضرورت این مضمون دارد.

براساس نتایج کتابسنجی، موضوعات مرتبط با ارزیابی بروندهای علمی علوم انسانی بین سال‌های ۲۰۰۱ تا ۲۰۱۸ جزء مضامین محرک به شمار می‌روند که گرایش به تبدیل شدن به موضوعات اساسی و تخصصی دارند. این امر بیانگر جافتادگی تدریجی موضوعات در پژوهش‌های این حوزه و توجه گسترده‌تر به آن‌ها در عرصه سیاست‌گذاری و عمل است. در کنار آن، نمایه‌های استنادی از دغدغه‌های پژوهشگران این حوزه است و بحث بر سر جامعیت و کفایت پوشش آن برای علوم انسانی همواره از موضوعات داغ بوده و هست.

شبکه مضامین ترسیم‌شده نشان می‌دهد که حوزه معیارهای ارزیابی با بیشترین ارتباط (فراوانی ۹) در مرکز توجه قرار دارد. مضمون معیارهای ارزیابی که مربوط به روش‌های ارزیابی است به طور ویژه در علوم انسانی مورد بحث قرار گرفته و بازتاب اصلی توجه به ماهیت متفاوت علوم انسانی و لزوم ارزیابی متناسب با آن در سال‌های اخیر بوده است. برای نمونه، مولر و همکاران (۲۰۲۴) (Miller et al., ۲۰۲۴) ارائه معیار توسط فعالان علوم انسانی را گوشزد نموده‌اند. ابراهیمی درچه و همکاران (Ebrahimi Dorcheh et al., ۲۰۲۴) با شناسایی و اولویت‌بندی معیارهای ارزیابی بروندهای علمی رشته‌ای خاص از علوم انسانی نشان داده‌اند که نه تنها ارزیابی علوم انسانی با علوم دیگر متفاوت است بلکه ارزیابی در هر رشته از علوم انسانی نیز با دیگری متفاوت خواهد بود. پژوهش‌های ابراهیمی درچه و همکاران (الف-۱۴۰۲؛ ب-۱۴۰۲) و ابراهیمی و میرشاه‌ولایتی (۱۴۰۲) نشان داده که تدوین معیارهای ارزیابی بر اساس ماهیت رشته‌های علوم انسانی و توجه به ابعاد مختلف تأثیر پژوهش برای افزایش عدالت و اعتبار ارزیابی ضروری است. همچنین مطالعات نجاتی و همکاران (۱۴۰۲) و کمساری و همکاران (۱۴۰۱) بر اهمیت زیرساخت داده‌ای و طراحی معیارهای ملی متناسب با ویژگی‌های هر رشته تأکید دارند. این نتایج نشان می‌دهد که درک مفهومی و عملی ارزیابی بروندهای علمی علوم انسانی باید چندبعدی، مبتنی بر شواهد و انعطاف‌پذیر باشد تا با تفاوت‌های میان‌رشته‌ای سازگار شود.

ارزیابی علوم انسانی هم وجوه مشترکی با سایر علوم دارد و هم زوایای متمایزی را بازتاب می‌دهد؛ به‌ویژه تأکید بر عدالت، ارزش فرهنگی و نقش سیاست‌گذاری، ویژگی منحصر به فرد ارزیابی در این حوزه است. نتایج حاضر تأکید می‌کند که برای سیاست‌گذاری در این حوزه باید مانند یک پازل عمل شود؛ بدین معنا که رویکردها، روش‌ها و ابزارها به‌صورت همزمان و متناسب با یکدیگر تغییر یابند. در این میان، سیاست ارزیابی در کانون رویکردها و معیارهای ارزیابی در محور روش‌ها قرار دارند. پیوند عدالت در ارزیابی و معیارهای ارزیابی بیانگر آن است که بهترین راه برای طراحی نظامی متناسب با ماهیت علوم انسانی، بهره‌گیری از معیارهای اختصاصی و متناسب با این علوم است. تحقق این امر مستلزم وجود زیرساخت‌های قوی برای فراهم‌آوردن داده‌های متنوع، صحیح، کامل و شفاف است؛ زیرا بدون چنین بستر داده‌ای، امکان دستیابی به ارزیابی معتبر فراهم نخواهد شد.

به‌طور کلی، تفاوت ماهوی میان علوم انسانی و علوم سخت باید در ارزیابی‌ها به رسمیت شناخته شود. توجه به این تفاوت نه تنها از تحریف یا کاریکاتوری شدن نظام ارزیابی جلوگیری می‌کند، بلکه امکان طراحی ابزارها و روش‌های خلاقانه و کارآمدتر را نیز فراهم می‌آورد. مقایسه با پیشینه نشان داد که پژوهش حاضر ضمن همسویی با یافته‌های گذشته، یک گام فراتر نهاده و روابط متقابل میان مؤلفه‌ها را در قالب یک شبکه مضامین آشکار ساخته است. اگر مقاله کبیری خوزانی و همکاران (۱۴۰۴) به‌عنوان مرتبط‌ترین پژوهش در پیشینه را یک مطالعه صرفاً کمی و بر اساس بسامد کلیدواژه‌ها بدانیم، پژوهش پیش رو یک مطالعه کمی و

کیفی است که علاوه بر شناسایی مضامین به بررسی روابط آن‌ها نیز پرداخته‌است. نوآوری این پژوهش در ترکیب و ارائه تحلیل مضمون و کتابسنجی است که پیش‌تر در مطالعات مشاهده نشده بود. با الهام از یافته‌ها می‌توان نقاط قوت و ضعف، فرصت‌ها و تهدیدهای وضعیت موجود را واکاوی و برای وضعیت مطلوب برنامه‌ریزی نمود. نتیجه این تغییر می‌تواند خودکفایی ارزیابی علوم انسانی در سطح سیاست‌گذاری و کاهش اضطراب ارزیابی در سطح پژوهشگران باشد.

با این حال، محدودیت‌هایی نیز در پژوهش حاضر وجود دارد. نخست، تحلیل مضمون و کتابسنجی تنها بر منابع منتشرشده متمرکز شده در پایگاه‌های بین‌المللی و احتمال آن هست که منابع در سطح ملی و دیدگاه‌های غیرمنتشره یا تجربیات میدانی برخی پژوهشگران علوم انسانی لحاظ نشده باشد. دوم، استفاده از کلیدواژه‌ها و چکیده‌ها برای تحلیل شبکه مضامین محدودیت‌هایی در جامعیت داده‌ها ایجاد می‌کند؛ چرا که ممکن است برخی مضامین مهم در متن کامل مقالات ذکر شده باشند اما در این فیلدها نیامده باشند. سوم، تحلیل حاضر به صورت مرور نظام‌مند و کمی-کیفی بوده و پژوهش میدانی یا مصاحبه با ذینفعان در آن اثر داده نشده‌است. این محدودیت‌ها باید هنگام تفسیر یافته‌ها و تعمیم آن‌ها به کل جامعه پژوهشی علوم انسانی در نظر گرفته شود. علاوه بر این، نتایج پژوهش بیشتر بازتاب‌دهنده روندهای تاریخی و الگوهای مستند است و ممکن است برخی تغییرات لحظه‌ای یا نوظهور در حوزه ارزیابی، از دید پژوهش خارج مانده باشد.

پیشنهادهای اجرایی پژوهش

۱. پایش دوره‌ای شبکه مضامین و بازنگری سیاست‌ها تا همواره با تحولات علمی هماهنگ باشد و از عقب‌ماندگی ارزیابی جلوگیری شود؛
۲. تهیه چک‌لیست براساس شبکه مضامین جهت بررسی نقاط قوت و ضعف مراکز سیاست‌گذاری و اجرایی ارزیابی برون‌دادهای علمی انسانی؛
۳. سیاست‌گذاری ارزیابی برون‌دادهای علمی علوم انسانی با توجه به شبکه مضامین به‌دست‌آمده؛
۴. اطلاع‌رسانی یافته‌های پژوهش به ذینفعان در قالب وبینارهای آموزشی.

پیشنهاد برای پژوهش‌های آتی

۱. مرور پژوهش‌های داخلی جهت کمک به بومی‌سازی شبکه مضامین؛
۲. بررسی تخصصی ارزیابی برون‌دادهای علمی هر یک از رشته‌های علوم انسانی؛
۳. پرداختن به مضامین ارزیابی برون‌دادهای بین‌رشته‌ای که ممکن است به‌دلیل تفکیک علوم انسانی از علوم اجتماعی و علوم دیگر در این پژوهش نادیده گرفته شده باشد؛
۴. مصاحبه با متخصصان و مدیران و پژوهشگران این حوزه جهت افزایش، کاهش یا غریبال مضامین؛
۵. اولویت‌بندی طبقات، حوزه‌ها و مقوله‌های حاصل از تحلیل مضمون؛
۶. وزن‌دهی ارتباطات در شبکه مضامین؛

۷. اعتبارسنجی نظری و عملی شبکه مضامین به دست آمده.

تقدیر و تشکر

این پژوهش تحت حمایت بنیاد علم ایران (به شماره ۴۰۳۳۴۸۱) و دانشگاه اصفهان (به شماره ۲۱۲/۱۴۰۲) در دوره پسادکتر انجام شده است.

تعارض منافع

نویسندگان اعلام می دارند که در خصوص انتشار این مقاله تضاد منافع وجود ندارد. علاوه بر این، موضوعات اخلاقی از جمله سرقت ادبی، رضایت آگاهانه، سوء رفتار، جعل داده ها، انتشار و ارسال مجدد و مکرر و همچنین، سیاست مجله در قبال استفاده از هوش مصنوعی از سوی نویسندگان رعایت شده است.

فهرست منابع

- ابراهیمی درجه، ا.، منصوری، ع.، پشوتنی زاده، م.، میرباقری فرد، ع.، و شعبانی، ا. (الف-۱۴۰۲). آسیب ها و راهکارهای ارزیابی برون داده های پژوهشی علوم انسانی: مطالعه موردی رشته های زبان و ادبیات. *پژوهش نامه علم سنجی*، ۹(۲)، ۷۵-۹۶.
<https://doi.org/10.22070/RSCI.2022.16133.1583>
- ابراهیمی درجه، ا.، منصوری، ع.، پشوتنی زاده، م.، میرباقری فرد، ع.، و شعبانی، ا. (ب-۱۴۰۲). سیاست ها و معیارهای ارزیابی برون داده های علمی علوم انسانی از دیدگاه متخصصان: مطالعه موردی رشته های زبان و ادبیات. *پژوهش و مدیریت اطلاعات*، ۳۹(۱)، ۶۳-۱۰۰.
<https://doi.org/10.22034/JIPM.2023.705519>
- ابراهیمی، ا.، و میرشاه ولایتی، ف. (۱۴۰۲). الگوی سنجش سطح آمادگی کاربردپذیری پژوهش های علوم انسانی و اجتماعی. *سیاست علم و فناوری*، ۱۶(۱)، ۱-۱۵.
<https://doi.org/10.22034/JSTP.2023.11192.1593>
- جنوی، ا.، و مردانی، ا. (۱۴۰۳). مطالعه تطبیقی معیارهای ارزیابی اعضای هیئت علمی دانشگاه ها در حوزه علوم انسانی در ایران و جهان. *مطالعات کتابداری و سازماندهی اطلاعات*، ۳۵(۲)، ۹۷-۱۴۲.
<https://doi.org/10.30484/NASTI.NFO.2024.3546.2263>
- رفاهی کمساری، ع.، اصنافی، ا.، و حاجی زین العابدینی، م. (۱۴۰۱). معیارهای ارزیابی و رتبه بندی محتوایی کتاب های تخصصی در حوزه علوم انسانی از نظر متخصصان. *کتابداری و اطلاع رسانی*، ۳۵(۳)، ۱۴۳-۱۶۵.
<https://doi.org/10.30481/LIS.2020.199112.1617>
- صفریپور، ا.، عباس پور، ج.، و سلیمی، ق. (۱۴۰۱). کاوش تجربیات و دیدگاه های اعضای هیأت علمی حوزه علوم انسانی و هنر دانشگاه شیراز در مورد آیین نامه ارتقا. *مطالعات دانش پژوهی*، ۱(۱)، ۲۹-۵۴.
<https://doi.org/10.22034/JKRS.2021.47805.1006>

عباس پور، ع.، مهرابی، گ.، دلاور، ع.، و خورسندی طاسکوه، ع. (۱۴۰۴). طراحی و اعتباریابی الگوی کیفیت پژوهش دانشگاهی در دانشگاه‌های علوم انسانی و اجتماعی. *اندازه گیری تربیتی*، ۱۶ (۶۰)، ۷-۳۶.

https://jematu.ac.ir/article_۱۹۰۰۷.html?lang=fa

علی نژاد چمازکتی، ف.، و میرحق جو لنگرودی، س. (۱۴۰۲). تحلیل علم‌سنجی و الگوهای هم‌نویسندگی و استناد مقالات منتشرشده در نشریات وب آو ساینس کشورهای ایران و ترکیه. *پژوهش نامه علم سنجی*، ۹ (۲)، ۴۱۳-۴۴۴.

<https://doi.org/۱۰.۲۲۰۷۰/RSCI.۲۰۲۲.۱۴۷۶۵.۱۵۱۳>

عصاره، ف.، عقیقیان، ف.، و نورمحمدی، ح. (۱۳۹۸). طراحی مدل مفهومی و عملیاتی ارزیابی و رتبه‌بندی علمی گروه‌های آموزشی حوزه علوم انسانی دانشگاه‌ها و مؤسسات آموزش عالی ایران. *پژوهش نامه علم سنجی*، ۵ (۱)، ۲۳-۴۶.

<https://doi.org/۱۰.۲۲۰۷۰/RSCI.۲۰۱۹.۴۱۹۲.۱۲۷۳>

فاضلی، نعمت‌الله (۱۳۹۶). *علوم انسانی و اجتماعی در ایران: چالش‌ها، تحولات و راهبردها*. پژوهشگاه علوم انسانی و مطالعات فرهنگی.

کبیری خوزانی، ن.، منصوری، ع.، پشتوتی‌زاده، م.، و ابراهیمی درجه، ا. (۱۴۰۴). تحلیل کتاب‌سنجی مفاهیم ارزیابی برون‌دادهای پژوهشی علوم انسانی. *پژوهش نامه علم سنجی*، ۱۱ (۱)، ۲۵-۴۸.

<https://doi.org/۱۰.۲۲۰۷۰/RSCI.۲۰۲۴.۱۸۷۰۲.۱۷۱۰>

مرزبان، ا.، رضایان، ا.، و رضایان، ع. (۱۴۰۲). آینده آیین‌نامه ارتقای مرتبه علمی اعضای هیئت علمی در ایران: یک رویکرد سناریویی. *پژوهش و برنامه ریزی در آموزش عالی*، ۲۹ (۱)، ۲۰۷-۲۳۳.

https://journal.irphe.ac.ir/article_۷۰۷۵۴۶.html

نجاتی، ف.، نوروزی چاکلی، ع.، و نورمحمدی، ح. (۱۴۰۲). شناسایی و تبیین شاخص‌های نخبگی و ارائه چارچوبی برای کشف نخبگان علمی در حوزه علوم انسانی. *پژوهش نامه علم سنجی*، ۹ (۲)، ۶۱۷-۶۵۴.

<https://doi.org/۱۰.۲۲۰۷۰/RSCI.۲۰۲۲.۱۶۰۲۹.۱۵۷۷>

Abbaspour, A., Mehrabi, G., Delavar, A., & Khorsandi Taskooh, A. (2025). Designing and validation of a model for academic research quality in social sciences and humanities universities. *Educational Measurement*, 16(60), 7-36. https://jem.atu.ac.ir/article_19007.html?lang=en [In Persian].

Alinezhad Chamazkoti, F., & Mirhaghjoo Langerudi, S. (2023). Scientometric analysis and co-authoring patterns and citations of articles in iranian and turkish journals indexed in WoS. *Scientometrics Research Journal*, 9(2), 413-444. <https://doi.org/10.22070/RSCI.2022.14765.1513> [In Persian].

Antonakis, J., & Hooijberg, R. (2008). Cascading a new vision: Three steps for real commitment. *Perspectives for Managers*, 157, 1. <https://B2n.ir/hz4054>

Aria, M., Cuccurullo, C., D'Aniello, L., Misuraca, M., & Spano, M. (2022). Thematic Analysis as a New Culturomic Tool: The Social Media Coverage on COVID-19 Pandemic in Italy. *Sustainability*, 14(6), 3643. <https://doi.org/10.3390/su14063643>

- Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Attride-Stirling, J. (2001). Thematic networks: an analytic tool for qualitative research. *Qualitative Research*, 1(3), 385–405. <https://doi.org/10.1177/146879410100100307>
- Bonaccorsi, A., Daraio, C., Fantoni, S., Folli, V., Leonetti, M., & Ruocco, G. (2017). Do social sciences and humanities behave like life and hard sciences? *Scientometrics*, 112(1), 607–653. <https://doi.org/10.1007/s11192-017-2384-0>
- Bornmann, L., & Daniel, H.-D. (2008). *What do citation counts measure? A review of studies on citing behaviour*. *Journal of Documentation*, 64(1), 45–80. <https://doi.org/10.1108/00220410810844150>
- Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, 9(1), 3–26. <https://doi.org/10.3390/su14063643>
- Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The Psychologist*, 26(2), 120–123. <https://doi.org/https://https://B2n.ir/pe9942>
- de Jong, S., & Spaapen, J. (2023). Formative evaluation in the Netherlands: An opportunity for the arts, humanities and social sciences. In M. Ochsner, S. Researcher, & S. Centre (Eds), *Accountability in Academic Life* (pp. 172–185). Edward Elgar Publishing. <https://doi.org/10.4337/9781800885738.00021>
- De Rijcke, S., Cosentino, C., Crewe, R., D'Ippoliti, C., Motala-Timol, S., Binti A Rahman, N., Rovelli, L., Vaux, D., & Yupeng, Y. (2023). *The future of research evaluation: A synthesis of current debates and developments*. <https://doi.org/10.24948/2023.06>
- Donovan, C. (2007). The qualitative future of research evaluation. *Science and Public Policy*, 34(8), 585–597. <https://doi.org/10.3152/030234207X256538>
- Ebrahimi Dorcheh, E., Mansouri, A., Pashootanizadeh, M., Mirbagherifard, A. A., & Shabani, A. (2024). Harms and solutions for evaluating humanities research outputs: A case study of language and literature. *Scientometrics Research Journal*, 9(2), 75–96. <https://doi.org/10.22070/RSCI.2022.16133.1583> **[In Persian]**.
- Ebrahimi Dorcheh, E., Mansouri, A., Pashootanizadeh, M., Mirbagheri fard, A. 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://doi.org/10.22034/JIPM.2023.705519> **[In Persian]**.

- Ebrahimi Dorcheh, E., Mansouri, A., Pashootanizadeh, M., Mirbagherifard, A.-A., & Shabani, A. (2024). Determining and prioritizing the evaluation criteria of humanities scientific outputs: A case study of language and literature fields. *Journal of Scientometric Research*, 13(1), 148–160. <https://doi.org/10.5530/jscires.13.1.13>
- Ebrahimi Dorcheh, E., Mansouri, A., Pashootanizadeh, M., Mirbagherifard, A.-A., & Shabani, A. (2024). Determining and Prioritizing the Evaluation Criteria of Humanities Scientific Outputs: A Case Study of Language and Literature Fields. *Journal of Scientometric Research*, 13(1), 148–160. <https://doi.org/10.5530/jscires.13.1.13>
- Ebrahimi, E., & Mirshahvelayati, F. (2023). A model to measure the applicability readiness level of humanities and social sciences research. *Journal of Science and Technology Policy*, 16(1), 1-15. <https://doi.org/10.22034/JSTP.2023.11192.1593> [In Persian].
- Faggiolani, C., & Solimine, G. (2018). Mapping the role of the book in evaluation at the individual and department level in Italian SSH. A multisource analysis. In A. Bonaccorsi (Eds.), *The Evaluation of Research in Social Sciences and Humanities: Lessons from the Italian Experience* (pp. 33–53). Springer. https://doi.org/10.1007/978-3-319-68554-0_2
- Greenhalgh, T., Raftery, J., Hanney, S., & Glover, M. (2016). Research impact: a narrative review. *BMC Medicine*, 14(1), 78. <https://doi.org/10.1186/s12916-016-0620-8>
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Raïfols, I. (2015). *Bibliometrics: The Leiden Manifesto for research metrics*. *Nature*, 520(7548), 429-431. <https://doi.org/10.1038/520429a>
- Hug, S. E., Ochsner, M., & Daniel, H.-D. (2013). Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history. *Research Evaluation*, 22(5), 369–383. <https://doi.org/10.1093/reseval/rvt008>
- Janavi, E., & Mardani, A. (2024). Comparative study of evaluation criteria for humanities faculty members of universities in Iran and the world. *Librarianship and Information Organization Studies*, 35(2), 97-142. <https://doi.org/10.30484/NAS-TINFO.2024.3546.2263> [In Persian].
- Kabiri Khozani, N., Mansouri, A., Pashootanizadeh, M., & Ebrahimi Dorcheh, E. (2025). Bibliometric analysis of the evaluation concepts of humanities research outputs. *Scientometrics Research Journal*, 11(1), 25-48. <https://doi.org/10.22070/RSCI.2024.18702.1710> [In Persian].
- Kulczycki, E., Guns, R., Pölönen, J., Engels, T. C. E., Rozkosz, E. A., Zuccala, A. A., Bruun, K., Eskola, O., Starčič, A. I., Petr, M., & Sivertsen, G. (2020). Multilingual publishing in the social sciences and humanities: A seven-country European study. *Journal of*

- the Association for Information Science and Technology*, 71(11), 1371–1385. <https://doi.org/10.1002/asi.24336>
- Kyvik, S. (2003). Changing trends in publishing behaviour among university faculty, 1980–2000. *Scientometrics*, 58(1), 35–48. <https://doi.org/10.1023/A:1025475423482>
- Lauronen, J. P. (2022). Tension in Interpretations of the social impact of the social sciences: walking a tightrope between divergent conceptualizations of research utilization. *SAGE Open*, 12(2), 1–16. <https://doi.org/10.1177/2158244022108996>
- Law, J. (1992). Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systems Practice*, 5(4), 379–393. <https://doi.org/10.1007/BF01059830>
- Marzban, E., Rezayan, A., & Rezayan, A. H. (2023). The futures of the faculty promotion regulations in Iran: A scenario approach. *Research and Planning in Higher Education*, 29(1), 207–233. https://journal.irphe.ac.ir/article_707546.html **[In Persian]**
- Moliní, F. (2021). The Evaluation of Innovation Within a Faculty of Arts and Humanities. *Journal of Higher Education Theory and Practice*, 21(6). <https://doi.org/10.33423/jhetp.v21i6.4384>
- Muhonen, R., Himanen, L., & Pölönen, J. (2023). Evaluation of societal impact of the Social Sciences and Humanities in Finland. In M. Ochsner, S. Researcher, & S. Centre (Eds.), *Accountability in Academic Life* (pp. 95–112). Edward Elgar Publishing. <https://doi.org/10.4337/9781800885738.00016>
- Müller, K., Salö, L., & Sörlin, S. (2024). Quality from within: Entry points to research quality in the humanities. *Research Evaluation*, 33, rvae029. <https://doi.org/10.1093/reseval/rvae029>
- Nejati, F., Noroozi Chakoli, A., & Nourmohammadi, H. (2024). Identifying and explaining elite indicators and providing a framework for discovering scientific elites in the Humanities Field. *Scientometrics Research Journal*, 9(2), 617–654. <https://doi.org/10.22070/RSCI.2022.16029.1577> **[In Persian]**.
- Olmos-Peñuela, J., Benneworth, P., & Castro-Martinez, E. (2014). Are 'STEM from Mars and SSH from Venus'? Challenging disciplinary stereotypes of research's social value. *Science and Public Policy*, 41(3), 384–400. <https://doi.org/10.1093/scipol/sct071>
- Osareh, F., Afifian, F., & Nourmohammadi, H. A. (2019). Designing a conceptual and operational model for scientific evaluation and ranking the of educational departments of humanities in universities and institutes of higher education in Iran. *Scientometrics Research Journal*, 5(1), 23–46. <https://doi.org/10.22070/RSCI.2019.4192.1273> **[In Persian]**.

- Pedersen, D. B., Grønvad, J. F., & Hvidtfeldt, R. (2020). Methods for mapping the impact of social sciences and humanities—A literature review. *Research Evaluation*, 29(1), 4–21. <https://doi.org/10.1093/reseval/rvz033>
- Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, 13(3), 141–146. <https://doi.org/10.1097/XEB.0000000000000050>
- Reale, E., Avramov, D., Canhial, K., Donovan, C., Flecha, R., Holm, P., Larkin, C., Lepori, B., Mosoni-Fried, J., Oliver, E., Primeri, E., Puigvert, L., Scharnhorst, A., Schubert, A., Soler, M., Soòs, S., Sordé, T., Travis, C., & Van Horik, R. (2018). A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. *Research Evaluation*, 27(4), 298–308. <https://doi.org/10.1093/reseval/rvx025>
- Refahi Komsari, A., Asnafi, A. R., & Haji Zeinlobaedini, M. (2022). Criteria for evaluating and ranking the content of specialized books in the field of humanities according to specialists' viewpoint. *Library and Information Sciences*, 25(3), 143–165. <https://doi.org/10.30481/LIS.2020.199112.1617> [In Persian].
- Rovira-Esteva, S., & P. Orero (2012). Evaluating quality and excellence in translation studies research: Publish or perish, the Spanish way. *International Journal of Translation*, 58(3), 264–288. <https://www.researchgate.net/publication/263110817>
- Safarpour, E., Abbaspour, J., & Salimi, G. (2022). Exploring the experiences and views of Shiraz University's Art and Humanities Faculty Members on Promotion Regulation. *Journal of Knowledge-Research Studies*, 1(1), 29–54. <https://doi.org/10.22034/JKRS.2021.47805.1006> [In Persian].
- Sani'ee, N., Nemati-Anaraki, L., Sedghi, S., Noroozi Chakoli, A., & Goharinezhad, S. (2022). The Effective Trends and Driving Forces in The Future of Research Performance Evaluation: A Qualitative Study. *Medical Journal of The Islamic Republic of Iran*, 36(1), 393–406. <https://doi.org/10.47176/mjiri.36.55>
- Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*, 107(2), 357–368. <https://doi.org/10.1007/s11192-016-1845-1>
- Toledo, E. G. (2016). Research assessment in humanities and social sciences in review. *Revista Española de Documentación Científica*, 41(3), 1–14. <https://doi.org/10.3989/redc.2018.3.1552>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... Straus, S. E. (2018).

- PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- Verleysen, F. T., & Weeren, A. (2016). Clustering by publication patterns of senior authors in the social sciences and humanities. *Journal of Informetrics*, 10(1), 254–272. <https://doi.org/10.1016/j.joi.2016.01.004>
- Vučković, D., Pekovic, S., Popović, S., & Janinovic, J. (2023). Assessing the Appraisal of Research Quality in Social Sciences and Humanities: A Case Study of the University of Montenegro. *Business Systems Research*, 14(1), 131–152. <https://doi.org/10.2478/bsrj-2023-0007>
- Wang, L., & Wang, T. (2024). Research efficiency evaluation and regional differences analysis of humanities and social sciences of colleges and universities in China. *Soft Computing*, 1–12. <https://doi.org/10.1007/s00500-024-09662-7>
- Wang, R. (2023). Research on Problems of Scientific Research Performance Evaluation Mechanism of Humanities and Social Sciences. *Frontiers in Educational Research*, 6(18). <https://doi.org/10.25236/FER.2023.061823>
- Wilsdon, J. (2015). *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management*. <https://digital.casalini.it/9781473978775>
- Wróblewska, M. N. (2021). Research impact evaluation and academic discourse. *Humanities and Social Sciences Communications*, 8(1), 1–12. <https://doi.org/10.1057/s41599-021-00727-8>
- Yang, S., Zheng, M., Yu, Y., & Wolfram, D. (2021). Are Altmeter.com scores effective for research impact evaluation in the social sciences and humanities? *Journal of Informetrics*, 15(1), 1–21. <https://doi.org/10.1016/j.joi.2020.101120>

پیوست

جدول ۵. منابع شواهد (۵۳ رکورد)

ID	Author(s)	Year	Title	Aim/ Purpose	Method	General Finding
1	Moed, Henk F. et al.	2002	Towards research performance in the humanities	Providing a general methodology for the development of bibliometric performance indicators	Case study	Identification of the characteristics of the field and the unit of analysis, along with the provision of appropriate indicators, constitutes a core stage of the methodology
2	Baneyx, Audrey	2008	“Publish or perish” as citation metrics used to analyze scientific output in the humanities: International case studies in economics, geography, social sciences, philosophy, and history	Examining the use of alternative data sources, such as Google scholar, in calculating citation metrics	Comparative study	Using Google scholar provides easy access to resources and more comprehensive citation coverage in the humanities
3	Huang, Mu-hsuan & Chang, Yu-wei	2008	Characteristics of research output in social sciences and humanities: From a research evaluation perspective	Investigating the nature of research outputs in the humanities and their implications for evaluation	Case study	Fair evaluation of research should take into account the diversity of outputs across different fields
4	Kousha, Kayvan & Thelwall, MikeM	2009	Google book search: Citation analysis for social science and the humanities	Investigating whether Google book search (GBS) can fill the gap in research evaluation	Comparative study	is a valuable Google book search (GBS) new source of citation data in the humanities
5	Bridges, David	2009	Research quality assessment in education: impossible science, possible art?	Examining the challenges and possibilities of evaluating research quality	Case study	Approaches based on measurable indicators (such as publication venue, number of citations, and downloads) are considered quasi-scientific; instead, an evaluation approach based on quality appreciation, discernment, or even connoisseurship is proposed

6	Kousha, Kayvan et al.	2011	Assessing the citation impact of books: The role of Google books, Google scholar, and Scopus	Investigating the potential of using Google books and Google scholar as alternative sources of citation evidence	Comparative study	Google books and Google scholar citations to books were 1.4 and 3.2 times higher, respectively, than those in Scopus
7	Larivière, Vincent & Macaluso, Benoît	2011	Improving the coverage of social science and humanities researchers' output: The case of the Érudit journal platform	Quantifying the importance of articles published in local journals	Comparative study	Including articles published in local journals increases researchers' output by nearly one-third
8	Giménez-Toledo, Elea	2012	Evaluation of scientific books publishers in social sciences and humanities: Results of a survey	Reviewing the evaluation of scientific book publishers	Survey study	Identifying the indicators that determine the quality of book publishers
9	Zuccala, Alesia	2012	Quality and influence in literary work: Evaluating the 'educated imagination'	Evaluating literary works as products of the 'educated imagination'	Survey study	Although citations can be used to track patterns of influence, peer review through book reviews plays a more significant role in evaluating literary quality
10	Mutz, Rüdiger	2012	Types of research output profiles: A multilevel latent class analysis of the Austrian science fund's final project report data	Examining expected research outputs in the humanities	Case study	The humanities should be regarded as an independent field, and a variety of scholarly outputs should be considered for evaluation
11	Bennett, Dawn & Franzmann, Majella	2013	Aristotle and the ERA measuring the immeasurable	Investigating the relationship between research evaluation frameworks and knowledge forms	Case study	Focusing on measurable outcomes risks overlooking the intangible and hidden aspects of research
12	Hug, Sven E	2013	Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history	Examining criteria for evaluating research quality	Delphi study	The need to conceptualize research quality and formulate clear evaluation criteria unique to the context of each field, based on experts' perspectives
13	Ochsner, Michael	2013	Four types of research in the humanities: Setting the stage for research quality criteria in the humanities	Identifying research concepts and criteria of intellectual quality	Survey study	Identifying three structuring dimensions of research concepts, four types of research in the humanities, and extracting fifteen quality criteria for research
14	Hazelkorn, Ellen	2014	Making an impact: new directions for arts and	Examining environmental changes in	Survey study	Although the current global economic

			humanities research	research evaluation		crisis has highlighted the importance of measuring academic and research performance and productivity, each country has adopted distinct policy approaches to research
15	Mohammad, Ehsan & Thelwall, Mike	2014	Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows	Comparing Mendeley readership and citation metrics across different fields	Case study	The correlation between Mendeley readership and citations suggests that Mendeley data can serve as a useful complementary metric for citation analysis and as an early indicator of impact
16	Molas-Gallart, Jordi	2014	Research evaluation and the assessment of public value	Examining research evaluation goals and the challenges of assessing the overall value of research	Case study	- A formative approach to evaluation is needed that supports the social engagement of academic researchers. This approach not only evaluates outcomes but also fosters researchers' ongoing engagement with societal needs, thereby enhancing the public value of research - Research evaluation requires a methodology capable of capturing the diverse and often less tangible ways in which value is created
17	Thelwall, Mike & Delgado, Maria M.	2015	Arts and humanities research evaluation: No metrics please, just data	Arguments for using data with contextual information in evaluation instead of rigid systematic criteria	Survey study	Creative use of non-systematic data types (such as questionnaires) to support research value, rather than relying solely on systematic and standardized measures (such as citations)
18	Chen, Kuang-hua	2015	Exploring alternative metrics of scholarly performance in the social sciences and	Investigating alternative metrics (Altmetrics) for improved evaluation of	Case study	Citation database metrics are inadequate for the humanities, which encompass

			humanities in Taiwan	research performance		diverse publication channels. Therefore, it is necessary to use alternative metrics (Altmetrics) for both official and unofficial channels to better reflect research outputs, considering the differences across fields
19	Oancea, Alis et al.	2015	The ecologies and economy of cultural value from research	A conceptual and empirical exploration of how individuals engaged in academic research respond to the challenges of creating, interpreting, and demonstrating the cultural value of research	Case study	The need for consensus on what constitutes cultural value to improve research evaluation in cultural and social fields by developing multifaceted approaches and tools that account for the complexity of cultural value
20	Kousha, Kayvan & Thelwall, Mike	2015	Web indicators for research evaluation. Part 3: Books and non-standard outputs	Describing web metrics for measuring the impact of books, software, datasets, videos, and other non-traditional academic outputs	Case study	Given the inadequacy of citation-based impact indicators for non-standard academic outputs, web- and performance-based indicators represent a potentially valuable resource for their evaluation
21	Verleysen, Frederik T. & Weeren, Arie	2016	Clustering by publication patterns of senior authors in the social sciences and humanities	Mapping the diversity of publication patterns	Case study	Identifying two widespread publication styles across fields that demonstrate the potential of cluster analysis for research evaluation
22	Holland, Charlotte et al.	2016	Performance anxiety in academia tensions within research assessment exercises in an age of austerity	Investigating the effects and challenges of research evaluations, focusing on changes driven by neoliberal programs and the emphasis on quantitative criteria in assessment	Case study	The emphasis on quantitative criteria in the neoliberal agenda diminishes the quality and diversity of research, undermines researchers' independence and autonomy, and creates pressure to produce measurable results. Therefore, it is necessary to develop appropriate criteria that are sensitive to the specific characteristics of humanities research and

						aligned with international best practices
23	Lauer, Gerhard	2016	The ESF scoping project 'towards a bibliometric database for the social sciences and humanities'	Investigating the potential for developing a research output database to evaluate research performance	Case study	Launching a national database with an emphasis on a bottom-up approach to address the inadequate coverage of international databases
24	Akker, Wiljan van den	2016	Yes we should; research assessment in the humanities	Arguments for establishing a research evaluation system	Review study	The humanities must establish an evaluation system that considers nationalism, diversity of outputs, and appropriate criteria
25	Thelwall, Mike	2017	Do Mendeley reader counts indicate the value of arts and humanities research?	Investigating whether Mendeley readership counts can serve as early indicators of research impact	Case study	Proposing the number of Mendeley readers as an indicator of early impact
26	Bulaitis, Zoe	2017	Measuring impact in the humanities: Learning from accountability and economics in a contemporary history of cultural value	Examining the future of research evaluation in higher education from a humanities perspective	Case study	Despite the increasing focus on measurable or commercial outcomes, there is a growing emphasis on measuring "impact" in research evaluation criteria
27	Buitendag, Johan & Simut, Corneliu C.	2017	Rethinking research impact by theology and religious studies with references to the faculty of theology at the university of Pretoria	Rethinking research impact and proposing solutions to enhance its assessment	Case study	Three factors should be considered and improved for a proper and fair assessment of research outputs: research productivity, citations, and academic reputation
28	Ochsner, Michael et al.	2017	The future of research assessment in the humanities: Bottom-up assessment procedures	Supporting a bottom-up approach to research evaluation	Case study	Adopting a bottom-up approach tailored to each field in evaluation
29	Vuong, Quan-Hoang et al.	2018	An open database of productivity in Vietnam's social sciences and humanities for public use	Providing an open database of researchers' scientific outputs	Case study	Creating a national open database to address shortcomings in current databases, such as data duplication, slow updates, and high access costs
30	Konkiel, Stacy	2018	Approaches to creating 'humane' research	Developing criteria for evaluating	Case study	Achieving consensus on the core values

			evaluation metrics for the humanities	"humane" research in the humanities		of each field within the humanities and developing a framework for measuring research based on the authentic values and needs of the field, rather than relying solely on common but inadequate quantitative indicators
31	.Engels, Tim C.E et al.	2018	Are book publications disappearing from scholarly communication in the social sciences and humanities?	Analyzing the evolution of the share of scientific book publications	Case study	Research evaluation systems should consider the book due to its vital epistemological and social role in the socio-economic system, while also accounting for variations in the book's contribution across different fields and countries
32	Watermeyer, Richard & Chubb, Jennifer	2018	Evaluating 'impact' in the UK's research excellence framework (REF): Liminality, looseness and new modalities of scholarly distinction	Investigating the process of evaluating the economic and social impact of research in universities	Case study	The emergence of a new and intermediate concept of impact as scholarly 'distinction,' which differs from traditional models of academic excellence
33	Muhonen, Reetta	2018	From productive interactions to impact pathways: Understanding the key dimensions in developing SSH research societal impact	Developing a typology to understand the mechanisms of research's social impact	Comparative study	Presenting a typology of research social impact pathways—focusing on productive interactions and the resulting changes—and highlighting the importance of considering the conditions that support impact processes
34	Toledo, Elea Giménez	2018	Research assessment in humanities and Social Sciences in review	Examining changes in methodologies and approaches to research evaluation	Review study	Key points in research evaluation methods include the use of comprehensive data on scientific outputs, greater emphasis on qualitative assessment, promotion of open access, adoption of alternative criteria, implementation of responsible metrics,

						and consideration of the social impact of research. These elements reflect a shift toward more inclusive, transparent, and context-sensitive approaches to evaluating scholarly work
35	Ma, Lai, & Ladisch, Michael	2019	evaluation complacency or evaluation inertia? A study of evaluative metrics and research practices in Irish universities	Exploring the understanding of evaluation criteria and their influence on research practices	Case study	- Evaluation inertia creates a discrepancy between participants' beliefs about how evaluation criteria should be applied and how they actually use them in practice - Evaluation criteria should not be standardized or institutionalized without thoroughly assessing their validity and reliability, as well as their impact on academic life, research practices, and knowledge production - There is a need for metrics literacy and, more importantly, for an open and public discourse on metrics. The conversation about metrics should not be confined to subject matter experts and policymakers, but should include all those who use or are affected by metrics and related issues
36	Sivertsen, Gunnar	2019	Understanding and evaluating research and scholarly publishing in the social sciences and humanities (SSH)	Reviewing and evaluating scientific publications and their representation in data sources	Survey study	The need for more comprehensive representation of publication types in multiple languages within current citation databases, reflecting the diverse publication patterns in the humanities
37	Zahedi, Zohreh & Costas	2020	Do online readerships offer useful assessment tools?	Examining the significance of online reading as an indicator for research	Comparative study	Readership indicators can be valuable for evaluating under-resourced outputs in

	Comesana, Rodrigo			evaluation		citation databases, demonstrating early impact, and engaging non-academic audiences
38	Tomaselli, Keyan G.	2020	Humanities, citations and currency: hierarchies of value and enabled recolonisation	A comparative study of the relevance of journal metrics for the humanities and sciences	Comparative study	Considering the impact of metrics on research integrity, a composite impact factor could be tailored to different fields by accounting for disciplinary cycles, procedural differences, teaching, community service, and other relevant factors
39	Doğan, and Zehra Taşkın, Güleda	2020	Humanities: The outlier of research assessments	Demonstrating the extent to which the humanities differ from other fields using five size-independent bibliometric indicators	Case study	In addition to differences between the humanities and other sciences, its subfields also vary. Therefore, rather than new quantitative indicators, different evaluation approaches—such as bottom-up methods—are needed
40	Janinovic, Jovana	2020	Innovative strategies for creating and assessing research quality and societal impact in social sciences and humanities	Investigating innovative methods for evaluating research quality and social impact	Review study	While most strategies for evaluating transdisciplinary research still require refinement and implementation, social impact measurement has emerged as a reliable, relevant, comprehensive, and practical approach to assessing research quality
41	Gajo, Laurent & Berthoud, Anne-Claude	2020	Issues of multilingualism for scientific knowledge: Practices for assessing research projects in terms of linguistic diversity	Examining the role of language in the construction, transmission, and dissemination of scientific knowledge	Review study	Paying attention to multilingualism in evaluation is essential to preserve the diversity of scientific cultures and research perspectives, thereby avoiding the dominance of English as the mainstream language
42	Pedersen, David	2020	Methods for mapping the impact of social	Examining key methodological components	Review study	- Impact evaluation methods should be

	Budtz		sciences and humanities—A literature review	used to evaluate "research impact" and their respective advantages and disadvantages		developed and applied according to specific contexts, avoiding the use of universal criteria - Evaluation should aim to better understand the nature and scope of impacts by focusing on the contribution of research to social impact, rather than attributing social impact to individual research projects
43	Valzano, Virginia	2020	Open science: New models of scientific communication and research evaluation	Exploring the importance of open science in transforming scientific communication and research evaluation	Review study	Evaluation criteria should be adapted to the principles of open science, prioritizing research access and quality over traditional measures
44	Maryl, Maciej	2020	The case for an inclusive scholarly communication infrastructure for social sciences and humanities	Providing a vision for the scientific communication infrastructure	Review study	Adapting the evaluation system to align with the principles and characteristics of open science
45	Sîle, Linda et al.	2020	Towards complexity-sensitive book metrics for scholarly monographs in national databases for research output	Exploring an approach to book evaluation in research assessment that accounts for the complexity of scholarly monographs	Case study	Bibliographic control of scientific monographs in national databases requires highly detailed metadata for these publications
46	Giménez Toledo, Elea	2020	Why books are important in the scholarly communication system in social sciences and humanities	Describing the state of the academic publishing sector and the role of small- and medium-sized publishers at the national level	Descriptive study	- Books hold a central role in the humanities and should be appropriately recognized in research evaluation - One criterion for evaluating a book is the assessment of its publisher
47	Xu, Xin	2021	The impacts of incentives for international publications on research cultures in Chinese humanities and social sciences	Investigating the impact of financial and job-related incentive schemes for international publications on research culture	Survey study	One-size-fits-all policymaking is challenged; instead, institutions are encouraged to support an open, fair, and people-centered—rather than output-

						centered—research culture when developing evaluation mechanisms
48	Boza-Chua, Alejandro	2022	Use of robotized process automation for research management and evaluation	- Identifying information and communication technology tools that facilitate effective research management - Implementing a system based on robotic process automation to develop a scientific information system	Case study	Designing and implementing a robotic process automation system that significantly reduces data collection time for evaluating researchers' scientific output
49	Thelwall, Mike	2023	Predicting article quality scores with machine learning: The U.K. research excellence framework	Exploring the feasibility of using machine learning to estimate article quality by incorporating multiple bibliometric and metadata inputs	Comparative study	Predictions can be made through machine learning using article quality scores on a three-level scale, based on article metadata, citation data, author job information, and title/abstract text
50	Ebrahimi Dorcheh, Elahe et al.	2024	Determining and prioritizing the evaluation criteria of humanities scientific outputs: A case study of language and literature fields	Determining and prioritizing evaluation criteria	Delphi study	Prioritizing evaluation criteria for scientific outputs according to the nature of each humanities field, based on expert opinions
51	Ebrahimi Dorcheh, Elahe et al.	2024	Determining and weighting effective outputs of humanities research to meet scientific and professional goals: A study of language and literature fields	Determining and weighting types of scientific outputs	Exploratory study	Weighting scientific outputs according to the specific nature of each field within the humanities
52	M€uller, Klara	2024	Quality from within: Entry points to research quality in the humanities	Identifying entry points for further studies on how humanities scholars engage with concepts of quality	Review study	Exploring entry points for further studies on how humanities researchers articulate concepts of quality in their own terms (quality from within)
53	Lisciandra, Chiara	2025	Citation Metrics: A philosophy of science perspective	Analyzing citation criteria from the perspective of Kuhn's model of scientific development	Review study	Citation criteria hinder the development of normal science in some fields and fail to account for field-specific differences