



Historical Evolution of the Concept and Measurement of Scientific Progress: Implications for Reforming Science Assessment in Iran

Habibollah Tabatabaeian

Faculty Member, Allameh Tabataba'i University, Tehran, Iran. (abatabaeian@atu.ac.ir) 

Abstract

This study examines how the concept and measurement of scientific progress have evolved over the past half-century and what this evolution implies for reforming science assessment in Iran. It adopts a historical-analytical, multi-source approach combining a structured literature review, complementary bibliometric evidence, policy-document analysis, and normative comparison. A Scopus search conducted on 11 August 2026, covering relevant records from 1964 onward, yielded 1,883 documents after filtering and single-reviewer title and abstract screening. The findings indicate that science assessment evolves through interactions among epistemic criteria, the social organization of science, measurement tools, and policy expectations. More recently, quality, impact, and responsibility have been added to earlier criteria. In Iran, development-plan objectives have expanded from capacity building and scientific production toward innovation and problem orientation, while official measurement remains predominantly quantitative and ranking-based. This mismatch is conceptualized as a concept–measurement gap. Differences between the attainment of financial indicators and selected outputs further underscore the need to distinguish resources, capacity, quality, and impact. Normative comparison also identifies domestic foundations that may support assessment reform. The study proposes a six-dimensional framework combining quantitative monitoring, qualitative judgment of quality and impact, and contextual and ethical adjustment. Pilot implementation and empirical validation of this framework are recommended.

Keywords: Scientific Progress, Science Measurement, Research Assessment, Science And Technology Policy, Responsible Research Assessment.

How to Cite this Paper:

Tabatabaeian, H. (2026). **Historical Evolution of the Concept and Measurement of Scientific Progress: Implications for Reforming Science Assessment in Iran.** *Journal of Science & Technology Policy*, 19(3), 87-105. {In Persian}. doi: 10.22034/jstp.2026.12372.2068





سال نوزدهم، شماره ۳، پاییز ۱۴۰۵

مقاله پژوهشی

فصلنامه علمی - پژوهشی
سیاست علم و فناوری

تحول تاریخی مفهوم و سنجش پیشرفت علمی: دلالت‌هایی برای بازنگری نظام ارزیابی علم در ایران

سید حبیب‌الله طباطبائی

عضو هیئت علمی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران. (tabatabaeian@atu.ac.ir) 

چکیده

این پژوهش به این پرسش پاسخ می‌دهد که مفهوم و سنجش پیشرفت علمی طی نیم قرن گذشته چگونه تحول یافته است و این تحول چه دلالت‌هایی برای بازنگری نظام ارزیابی علم در ایران دارد. روش پژوهش تاریخی - تحلیلی و چندمنبعی است و مرور ساختاریافته ادبیات، شواهد مکمل کتاب‌سنجی، تحلیل اسناد سیاستی و مقایسه هنجاری را تلفیق می‌کند. جست‌وجوی اسکوپوس در ۱۱ اوت ۲۰۲۶، با پوشش مدارک مرتبط از سال ۱۹۶۴، پس از اعمال محدودیت‌ها و غربالگری عنوان و چکیده، به مجموعه‌ای شامل ۱۸۸۳ مدرک انجامید. یافته‌ها نشان می‌دهند تحول ارزیابی علم، حاصل تعامل معیارهای معرفتی، سازمان اجتماعی علم، ابزارهای سنجش و انتظارات سیاستی است که در سال‌های اخیر، لایه‌های جدید کیفیت، اثرگذاری و مسئولیت به معیارهای پیشین افزوده شده‌اند. در ایران نیز اهداف برنامه‌های توسعه از ظرفیت‌سازی و تولید علمی به نوآوری و مسئله‌محوری گسترش یافته، اما سنجش رسمی همچنان عمدتاً کمی و رتبه‌محور است. این ناهم‌زمانی به صورت «شکاف مفهوم - سنجش» صورت‌بندی می‌شود. تفاوت تحقق شاخص‌های مالی و برخی برون‌دادها نیز ضرورت تفکیک منابع، ظرفیت، کیفیت و اثر را نشان می‌دهد. همچنین مقایسه هنجاری، ظرفیت‌هایی را برای پشتیبانی بومی از اصلاح ارزیابی آشکار می‌کند. دلالت سیاستی پژوهش، طراحی چارچوبی شش‌بعدی با ترکیب پایش کمی، داوری کیفیت و اثر، و تعدیل زمینه‌ای و اخلاقی است که اجرای آزمایشی و اعتبارسنجی تجربی آن نیز پیشنهاد می‌شود.

کلیدواژه‌ها: پیشرفت علمی، سنجش علم، ارزیابی پژوهش، سیاست علم و فناوری، ارزیابی مسئولانه پژوهش.

برای استنادات بعدی به این مقاله، قالب زیر به نویسندگان محترم مقالات پیشنهاد می‌شود:

طباطبائی، سید حبیب‌الله. (۱۴۰۴). تحول تاریخی مفهوم و سنجش پیشرفت علمی: دلالت‌هایی برای بازنگری نظام ارزیابی علم در ایران. *سیاست علم و فناوری*,

۱۹(۳)، ۸۷-۱۰۵.

doi: 10.22034/jstp.2026.12372.2068



۱- مقدمه

مفهوم «پیشرفت علمی» از مفاهیم بنیادین در سیاست‌گذاری علم، فناوری و نوآوری است؛ بااین‌حال، درباره چیرستی آن و معیارهای مناسب برای سنجش آن اجماع کاملی وجود ندارد. پیشرفت علمی می‌تواند در سطح معرفتی به افزایش دانش معتبر و فهم بهتر پدیده‌ها اشاره داشته باشد؛ در سطح نظام علمی با گسترش ظرفیت تولید و انتشار دانش سنجدیده شود؛ و در سطح اجتماعی و اقتصادی بر توانایی علم برای حل مسائل، ایجاد نوآوری و بهبود شرایط زندگی دلالت کند. تفاوت میان این تلقی‌ها موجب می‌شود انتخاب شاخص‌های سنجش پیشرفت علمی صرفاً یک مسئله فنی نباشد، بلکه به تعریف علم، اهداف نظام علمی و انتظارات جامعه از پژوهش وابسته باشد.

در اواخر قرن بیستم، توسعه پایگاه‌های اطلاعات علمی و روش‌های کتاب‌سنجی، امکان سنجش کمی تولیدات علمی را فراهم کرد و شاخص‌هایی مانند تعداد انتشارات، استنادها و شاخص‌های تأثیر به تدریج در ارزیابی پژوهشگران، دانشگاه‌ها و نظام‌های علمی جایگاه مهمی یافتند [۱ و ۲]. بااین‌حال، استفاده جانشین‌وار از این سنجدها با نقدهای فزاینده‌ای روبه‌رو شد. ابتکارهایی مانند بیانیه سانفرانسیسکو درباره ارزیابی پژوهش (DORA)، مانیفست لیدن و ائتلاف پیشبرد ارزیابی پژوهش (CoARA) بر استفاده مسئولانه از سنجدها، داوری کیفی، توجه به تفاوت رشته‌ها و مأموریت‌ها و شناسایی تنوع دستاوردهای پژوهشی تأکید کرده‌اند [۳، ۴ و ۵].

در ایران نیز «پیشرفت علمی» طی دهه‌های اخیر جایگاهی برجسته در اسناد و سیاست‌های علم و فناوری یافته است. قوانین برنامه‌های پنجدساله توسعه نشان می‌دهند که دامنه اهداف علمی کشور از ظرفیت‌سازی آموزشی و پژوهشی و تأمین منابع مالی، به تولید علمی، فناوری، نوآوری، تجاری‌سازی و مسئله‌محوری گسترش یافته است. بااین‌حال، استمرار شاخص‌های کمی و رتبه‌ای در کنار ورود معیارهای جدید این پرسش را ایجاد می‌کند که تحول نظام سنجش علمی در ایران تا چه حد با تحول مفهومی و روش‌شناختی ارزیابی علم در سطح بین‌المللی هم‌راستا بوده است.

در این پژوهش، تحول «مفهوم پیشرفت علمی» از تحول «ابزارهای سنجش» تفکیک می‌شود. مسئله اصلی، چگونگی

پیوند معیارهای معرفتی، نهادهای تولید دانش و انتظارات سیاستی با شاخص‌های ارزیابی است. البته شایان توجه است که تغییر تاریخی نیز به معنای تغییر وزن این منطقی‌هاست و ظهور سنجدهای جدید، اعتبار پرسش‌های معرفت‌شناختی پیشین را از میان نمی‌برد.

اصلاح ارزیابی همچنین به زمینه نهادی و هنجاری وابسته است. در این مقاله بررسی منابع اسلامی و جهت‌گیری‌های رسمی درباره علم در ایران، به شناسایی پشتوانه‌ها و محدودیت‌های هنجاری اصلاحات اختصاص دارد؛ از این منابع، شاخص فنی استخراج نمی‌شود و همچنین هم‌پوشانی ارزشی نیز دلیل کارآمدی اجرایی یک الگو تلقی نمی‌گردد.

بر این اساس، سؤال اصلی پژوهش چنین صورت‌بندی می‌شود: «الگوی تاریخی تحول مفهوم و سنجش پیشرفت علمی در ادبیات و سیاست‌های علمی جهان چیست؛ این تحول چگونه در سیاست‌های علم و فناوری ایران بازتاب یافته است، و چه دلالت‌هایی برای بازنگری نظام سنجش پیشرفت علمی در ایران دارد؟»

پژوهش در پاسخ به این سؤال چهار هدف ذیل را دنبال می‌کند: (۱) شناسایی تحولات اصلی در تلقی از پیشرفت علمی و معیارهای ارزیابی آن؛ (۲) تبیین رابطه تحول مفهوم پیشرفت با گسترش ارزیابی چندبعدی و مسئولانه؛ (۳) تحلیل تحول شاخص‌های علم و فناوری ایران و مقایسه آن با روندهای بین‌المللی؛ و بالاخره (۴) بررسی ظرفیت مبانی هنجاری بومی برای پشتیبانی از اصلاح نظام سنجش پیشرفت علمی.

۲- روش پژوهش

پژوهش تاریخی - تحلیلی حاضر، چهار مسیر مکمل را تلفیق می‌کند: (۱) مرور ساختاریافته ادبیات با شواهد کتاب‌سنجی؛ (۲) تحلیل تاریخی-مفهومی؛ (۳) تحلیل اسناد سیاست علم ایران؛ و (۴) مقایسه مبانی هنجاری بومی. مسیر نخست ساختار ادبیات را توصیف می‌کند، دو مسیر بعدی تحول مفهوم و منطق سنجش را تبیین می‌کنند و مسیر چهارم به دلالت‌های اصلاح در زمینه ایران می‌پردازد.

۲-۱ مرور ساختاریافته ادبیات و تحلیل کتاب‌سنجی

به‌منظور شناسایی ساختار و روند گسترش ادبیات مرتبط با

تحلیل با Excel، BibExcel و VOSviewer انجام شد. روند زمانی، فراوانی کلیدواژه‌ها و نقشه هم‌رخدادی، شواهد توصیفی مکمل برای بحث مفهومی‌اند؛ از آن‌ها رابطه علی یا ترتیب تاریخی ظهور مکاتب استنتاج نمی‌شود. تفسیر شبکه در بخش ۳-۴ به برجسب‌ها و پیوندهای قابل مشاهده محدود است و جایگزین گزارش فنی خوشه‌بندی نیست.

۲-۲ تحلیل تاریخی-مفهومی

در مرحله دوم، یافته‌های ادبیات در چارچوب تحلیل تاریخی-مفهومی تفسیر شدند. واحد اصلی تحلیل «تلقی از پیشرفت علمی» و «منطق مناظر سنجش آن» بود. در این مرحله چهار سطح از یکدیگر تفکیک شد: سطح معرفت‌شناختی؛ سطح اجتماعی- نهادی؛ سطح سنجش و ارزیابی؛ و سطح سیاستی- اثرمحور. هدف، انتساب هر دوره به یک مکتب فلسفی واحد یا ادعای پایان یافتن رویکردهای پیشین نبود؛ بلکه بررسی این موضوع بود که در دوره‌های مختلف چه منطق‌هایی در تعریف و ارزیابی پیشرفت علمی برجستگی بیشتری یافته‌اند.

۲-۳ تحلیل اسناد سیاست علم و فناوری ایران

در مرحله سوم، قوانین برنامه‌های توسعه پس از انقلاب اسلامی با تمرکز بر مؤلفه‌های علم، پژوهش، فناوری و نوآوری تحلیل شدند. برای هر برنامه، سه مؤلفه «هدف و جهت‌گیری سیاستی»، «شاخص سنجش‌پذیر» و «سازوکار پایش» از یکدیگر تفکیک شد. شاخص‌ها بر اساس تمرکز بر ظرفیت و نهاده، برون‌داد علمی، فناوری و نوآوری، اثر اقتصادی و اجتماعی و مسئله‌محوری دسته‌بندی شدند. هر جا داده‌های رسمی، معتبر و هم‌تعریف در دسترس بود، هدف و عملکرد با یکدیگر مقایسه شد؛ و هر جا رقم پایان دوره یا داده قابل مقایسه وجود نداشت، محدودیت به‌صراحت در جداول عملکرد گزارش شد.

مفهوم و سنجش پیشرفت علمی، مرور ساختاریافته‌ای از مطالعات نمایه‌شده در پایگاه Scopus انجام شد. جست‌وجو در ۱۱ اوت ۲۰۲۶ انجام گرفت و دامنه تحلیلی از نخستین مدرک مرتبط شناسایی شده در سال ۱۹۶۴ تا تاریخ انجام جست‌وجو را دربر گرفت. جست‌وجو در عنوان، چکیده و کلیدواژه انجام شد. عبارت پایه جست‌وجو به شکل زیر بود:

TITLE-ABS-KEY ("scientific progress" OR "progress of science" OR "scientific advancement" OR "scientific development") AND TITLE-ABS-KEY (measur* OR evaluat* OR assess* OR indicator* OR metric* OR bibliometric* OR scientometric*)

محدودیت‌های جست‌وجو شامل زبان انگلیسی، انواع مدرک مقاله، مقاله مروری و فصل کتاب، و حوزه‌های علوم اجتماعی (SOCI)، هنر و علوم انسانی (ARTS)، مدیریت و کسب‌وکار (BUSI)، ریاضیات (MATH)، فیزیک (PHYS)، اقتصاد (ECON) و علوم تصمیم (DECI) بود. طبعاً این انتخاب برای تمرکز بر مفهوم و ارزیابی علم انجام شد و پوشش همه کاربردهای اصطلاح پیشرفت علمی را تضمین نمی‌کند.

جست‌وجوی اولیه ۶۰۰۶ رکورد بازایی کرد. با اعمال محدودیت‌های زبان، نوع مدرک و حوزه موضوعی، ۴۰۰۱ رکورد کنار گذاشته شد و ۲۰۰۵ رکورد به غربالگری عنوان و چکیده رسید. پس از حذف ۱۲۲ مدرک نامرتبط، ۱۸۸۳ مدرک برای تحلیل باقی ماند. این اعداد مستند به خروجی جست‌وجو هستند. غربالگری را یک پژوهشگر انجام داده است؛ بنابراین ارزیابی مستقل دوم و سنجش توافق بین ارزیابان انجام نشده است. مطالعه با عنوان «مرور ساختاریافته» گزارش می‌شود و ادعای اجرای کامل مرور نظام‌مند یا تبعیت کامل از PRISMA را ندارد. معیارها در جدول الف آمده‌اند؛ تعداد حذف به تفکیک هر فیلتر گزارش نشده و مجموع ۴۰۰۱ نباید به سهم‌های جداگانه تعبیر شود (جدول ۱).

جدول ۱) معیارهای ورود و خروج مدارک

مؤلفه	معیار ورود	معیار خروج
پایگاه و زمان	Scopus؛ مدارک مرتبط از ۱۹۶۴ تا ۱۱ اوت ۲۰۲۶	خارج از دامنه جست‌وجوی تعریف شده
زبان و مدرک	انگلیسی؛ مقاله، مرور، فصل کتاب	سایر زبان‌ها و انواع مدرک
حوزه موضوعی	SOCI، ARTS، BUSI، MATH، PHYS، ECON، DECI	حوزه‌های خارج از فیلترهای انتخاب شده
ارتباط محتوایی	ارتباط مستقیم عنوان یا چکیده با مفهوم یا ارزیابی پیشرفت علمی	کاربرد حاشیه‌ای اصطلاح یا محتوای نامرتبط

۲-۴ تحلیل مبانی هنجاری بومی

مقایسه هنجاری بر منابع مشخص متمرکز است و مرور جامع متون اسلامی یا همه بیانات سیاسی محسوب نمی‌شود. واحد مقایسه، گزاره مرتبط با اعتبار دانش، عدالت، مسئولیت یا سودمندی اجتماعی است. دلالت منبع از کاربست پیشنهادی آن در ارزیابی پژوهش تفکیک می‌شود. قواعد مقایسه گزاره‌های سیاسی در بخش ۶-۳ توضیح داده شده‌اند.

۲-۵ تلفیق یافته‌ها

در مرحله نهایی، نتایج چهار مسیر در یک چارچوب تطبیقی تلفیق شدند: تلفی از پیشرفت علمی؛ معیارهای متناظر ارزیابی؛ تحول بین‌المللی؛ بازتاب در سیاست‌های ایران؛ مبانی هنجاری بومی؛ دلالت برای اصلاح نظام سنجش. هدف از این مرحله تعیین برتری یک الگوی واحد نیست، بلکه شناسایی نقاط هم‌راستایی، شکاف‌ها و مؤلفه‌های لازم برای یک نظام ارزیابی چندبعدی، زمینه‌مند و مسئولانه است.

۳- یافته‌های مرور ساختاریافته و تحلیل کتاب‌سنجی

شکل ۱ مسیر انتخاب مدارک را بر پایه خروجی واقعی جست‌وجو نشان می‌دهد. حذف پیش از غربالگری ناشی از محدودیت‌های پایگاه و حذف مرحله بعد بر اساس ارتباط عنوان و چکیده با سؤال پژوهش بوده است؛ این دو نوع حذف از یکدیگر تفکیک شده‌اند.

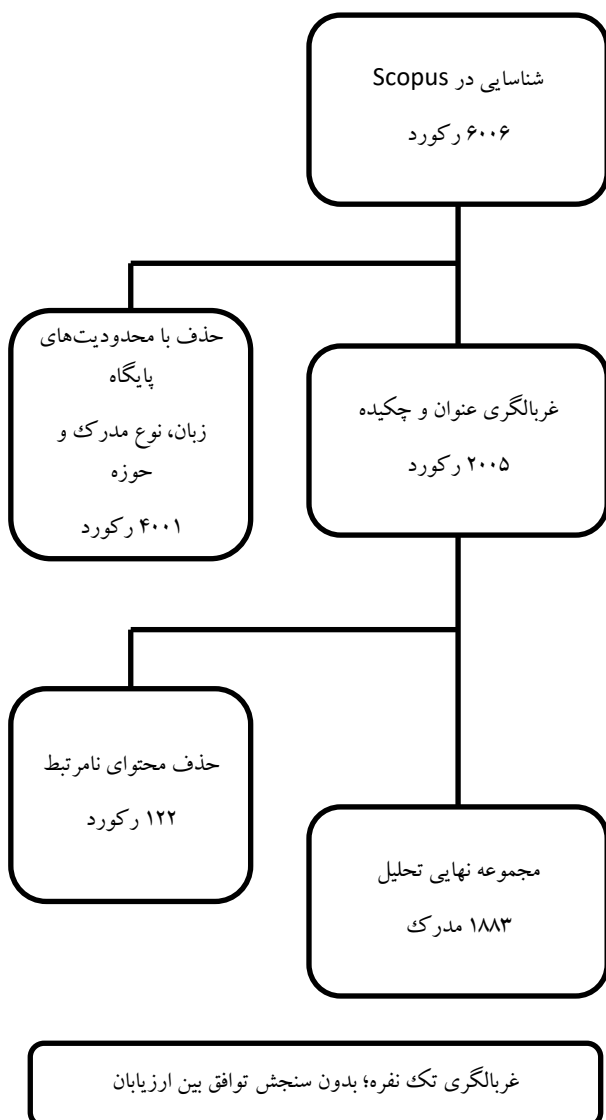
۳-۱ روند زمانی انتشار

داده‌های زمانی موجود نشان می‌دهند که ادبیات مرتبط با مفهوم و سنجش پیشرفت علمی از سال ۱۹۶۴ به بعد روندی روبه‌گسترش داشته است. در فاصله ۱۹۶۴ تا اواخر دهه ۱۹۸۰، تعداد انتشارات سالانه محدود بود؛ از دهه ۱۹۹۰ رشد تدریجی آغاز شد و از ابتدای قرن بیست‌ویکم شتاب بیشتری یافت. بر اساس داده‌های نمودار تعبیه‌شده در فایل تحلیل، تعداد مدارک از ۸۵ مورد در سال ۲۰۱۸ به ۱۱۷ مورد در ۲۰۲۱، ۱۵۲ مورد در ۲۰۲۳، ۲۱۷ مورد در ۲۰۲۴ و ۲۱۸ مورد در ۲۰۲۵ رسیده است. رقم ۱۵۱ مدرک برای ۲۰۲۶ به سال ناقص تا ۱۱ اوت مربوط است. این روند به‌تنهایی بیانگر تغییر در معنای پیشرفت علمی نیست، بلکه افزایش توجه پژوهشی به مسئله تعریف،

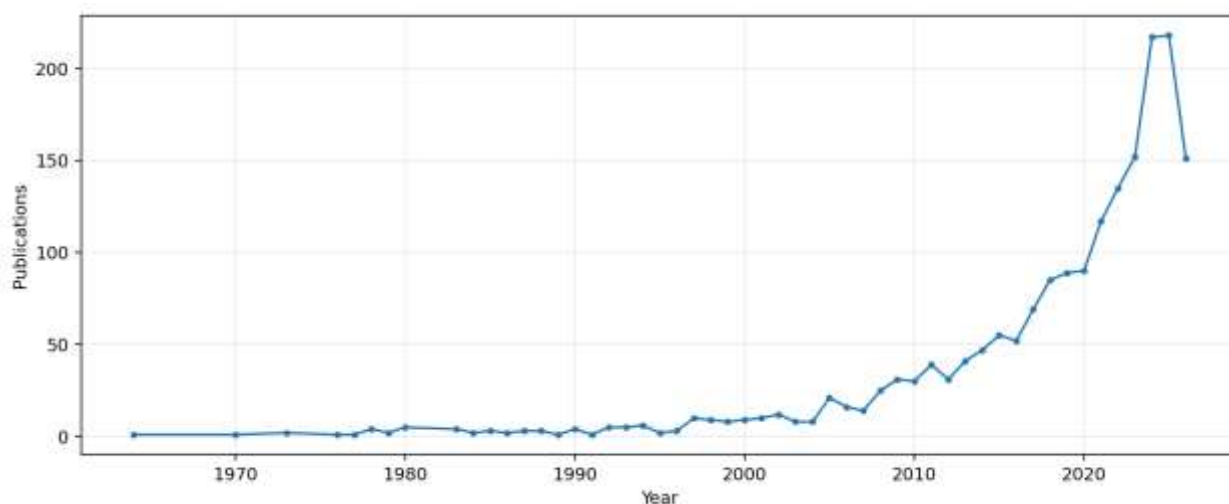
سنجش و ارزیابی پیشرفت علمی را نشان می‌دهد. (شکل ۲)

۳-۲ جهت‌گیری‌های پژوهشی پژوهشگران پرکار

در مجموعه، ۵۳۳۸ پژوهشگر مشارکت داشته‌اند؛ Mansoor Niaz با ۱۱، Lutz Bornmann با ۷ و Mario Coccia با ۶ مدرک از نویسندگان پرکارند. در توصیف محتوایی آثار آنان، به‌ترتیب توجه به تاریخ و فلسفه علم، سنجش و اثرگذاری پژوهش، و پویایی علم و نوآوری برجسته است. این نمونه‌ها تنوع جهت‌گیری ادبیات را نشان می‌دهند و نماینده جامع همه سنت‌های فکری نیستند.



شکل ۱) مسیر انتخاب مدارک در مرور ساختاریافته



شکل ۲) روند زمانی انتشار مطالعات مرتبط با پیشرفت علمی در مجموعه Scopus

۳-۳ گستره جغرافیایی ادبیات

مجموعه شامل آثاری از ۱۳۷ کشور است؛ آمریکا با ۳۵۱، چین با ۲۳۸، بریتانیا با ۱۴۵ و آلمان با ۱۴۰ مدرک بیشترین حضور را دارند. این اعداد گستره مجموعه را توصیف می‌کنند و مبنای رتبه‌بندی کیفیت یا پیشرفت علمی کشورها نیستند.

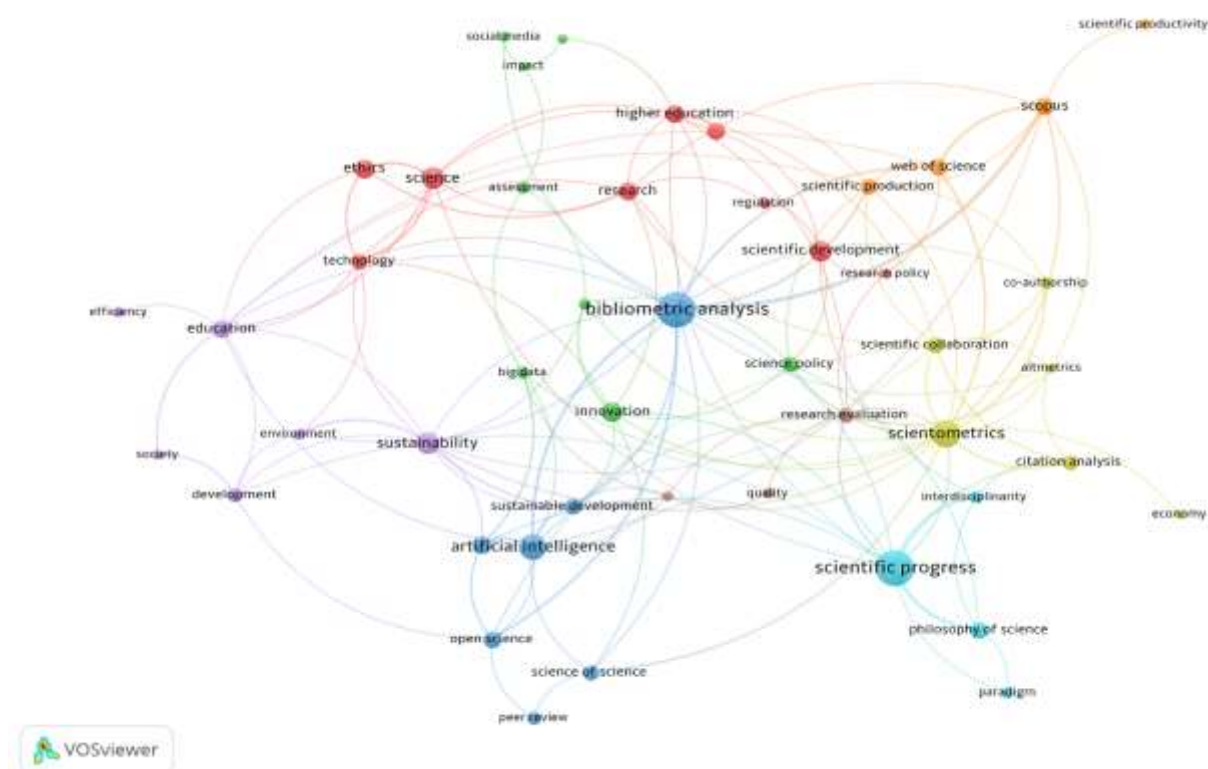
۴-۳ ساختار موضوعی ادبیات

در مجموع ۶۰۳۳ کلیدواژه در مدارک مجموعه شناسایی شد. کلیدواژه‌های پرتکرار نشان می‌دهند ادبیات پیشرفت علمی در پیوند هم‌زمان با کتاب‌سنجی و علم‌سنجی، هوش مصنوعی و کلان‌داده، پایداری و نوآوری، اخلاق و علم باز، فلسفه علم، ارزیابی پژوهش، همکاری علمی، کیفیت و اثرگذاری شکل گرفته است. جدول ۲، سی کلیدواژه پرتکرار را نشان می‌دهد.

در خوانش توصیفی شکل ۳، هم‌جواری «تولید علمی»، Scopus و Web of Science بر زیرساخت مشاهده و شمارش برون‌داد دلالت دارد؛ این ناحیه به سطح سنجش در بخش ۴-۴ مرتبط است. پیوندهای «علم‌سنجی»، «تحلیل استنادی» و «همکاری علمی» نیز اتصال شاخص‌ها به روابط جمعی تولید علم را نشان می‌دهد و با بحث اجتماعی - نهادی بخش ۴-۳ مرتبط می‌شود. هم‌جواری «فلسفه علم»، «پارادایم» و «پیشرفت علمی» حضور پرسش معرفتی بخش ۴-۲ را آشکار می‌کند. در سوی دیگر، «پایداری»، «جامعه» و «توسعه»، و پیوندهای «نوآوری»، «سیاست علم» و «اثر»، زمینه بحث اثرمحوری بخش ۴-۵ را فراهم می‌کنند. حضور «اخلاق»، «داوری هم‌تا»، «علم باز» و «کیفیت» نیز نشان می‌دهد بحث ارزیابی به شمارش محدود نیست.

جدول ۲) سی کلیدواژه پرتکرار در مجموعه مورد بررسی

ردیف	کلیدواژه	رخداد	ردیف	کلیدواژه	رخداد	ردیف	کلیدواژه	رخداد
۱	پیشرفت علمی	۶۵	۱۱	یادگیری ماشین	۳۴	۲۱	تحلیل استنادی	۲۵
۲	تحلیل کتاب‌سنجی	۶۴	۱۲	آموزش عالی	۳۲	۲۲	فناوری	۲۵
۳	علم‌سنجی	۶۰	۱۳	پژوهش	۳۱	۲۳	علم علم	۲۴
۴	هوش مصنوعی	۵۸	۱۴	علم باز	۳۰	۲۴	روش‌شناسی	۲۳
۵	پایداری و توسعه پایدار	۵۵	۱۵	Web of Science	۲۸	۲۵	ارزیابی پژوهش	۲۳
۶	توسعه علمی	۵۱	۱۶	فلسفه علم	۲۷	۲۶	توسعه	۲۳
۷	نوآوری	۴۷	۱۷	سنجش	۲۷	۲۷	داوری	۲۲
۸	اخلاق	۴۰	۱۸	همکاری علمی	۲۷	۲۸	بین‌رشته‌ای	۲۲
۹	Scopus	۴۰	۱۹	تولیدات علمی	۲۷	۲۹	کلان‌داده	۲۲
۱۰	آموزش	۳۸	۲۰	سیاست‌گذاری علم	۲۶	۳۰	ارزیابی	۲۲



شکل ۳) شبکه هم‌رخدادی کلیدواژه‌ها در ادبیات مورد بررسی

شکل ۳ تفسیر مفهومی نواحی شبکه است؛ تعداد خوشه‌ها، عضویت قطعی گره‌ها یا قدرت آماری پیوندها از تصویر استنتاج نمی‌شود. فراوانی کلیدواژه‌ها نیز به‌تنهایی غلبه یک مکتب را اثبات نمی‌کند. سهم این بخش در استدلال مقاله، نشان دادن هم‌حضور مسائل معرفتی، سنجشی، نهادی و اثرمحور است؛ دوره‌بندی بخش بعد بر ادبیات تاریخی و اسناد تکیه دارد. در مقایسه با ایران، همین تنوع مبنای طرح این پرسش است که آیا تنوع اهداف سیاستی با تنوع معیارهای ارزیابی همراه شده است.

۴- چارچوب تحلیلی تحول مفهوم و سنجش پیشرفت علمی

۴-۱ تفکیک سطوح تحلیل

چهار سطح تحلیل عبارت‌اند از (۱) معیار اعتبار و رشد دانش، (۲) شرایط اجتماعی و نهادی تولید آن، (۳) ابزارهای مشاهده و ارزیابی، و (۴) اهداف و پیامدهای سیاستی. این سطوح مکمل‌اند، اما هم‌ارز نیستند؛ شاخص عملکرد نمی‌تواند به‌تنهایی تعریف پیشرفت معرفتی را تعیین کند. بخش‌های ۲-۴ تا ۵-۴

۴-۲ سطح معرفت‌شناختی پیشرفت به‌مثابه تحول در دانش و فهم

در سطح معرفت‌شناختی، مسئله اصلی معیار پیشرفت خود دانش علمی است. روایت‌های اثبات‌گرایانه و اثبات‌گرایی منطقی بر مشاهده، صورت‌بندی تجربی و نظم معرفتی تأکید داشتند [۱۵-۱۳]. با این حال، از این سنت نمی‌توان به‌طور مستقیم نتیجه گرفت که شاخص‌هایی مانند تعداد مقاله و استناد محصول فلسفه اثبات‌گرایی بوده‌اند؛ تاریخ سنجش علم نشان می‌دهد توسعه زیرساخت‌های اطلاعاتی، نیازهای مدیریتی و سیاستی و امکان محاسبه در مقیاس وسیع نیز در شکل‌گیری سنجش‌های کمی نقش مهمی داشته‌اند [۱ و ۲].

پوپر، کوهن و لاکاتوش نیز پاسخ‌های متفاوتی به مسئله پیشرفت علم ارائه کرده‌اند. پوپر پیشرفت را با نقد و ابطال نظریه‌های ناموفق پیوند می‌دهد [۱۶]؛ کوهن بر علم عادی، ناهنجاری و تغییر پارادایم‌ها تأکید می‌کند و تصویری غیرخطی از تحول علم ارائه می‌دهد [۱۷]؛ و لاکاتوش با مفهوم برنامه‌های پژوهشی، ارزیابی تغییر علمی را در سطح

مجموعه‌های نسبتاً پایدار نظری و روش‌شناختی دنبال می‌کند [۱۸]. این دیدگاه‌ها یک نظریه واحد نیستند؛ اهمیت آن‌ها برای پژوهش حاضر در نشان دادن چندمعنایی بودن «پیشرفت علمی» حتی در سطح معرفت‌شناختی است.

۳-۴ سطح اجتماعی-نهادی پیشرفت علم به مثابه فرایندی جمعی

در سطح اجتماعی-نهادی، علم نه صرفاً مجموعه‌ای از گزاره‌ها، بلکه فعالیتی سازمان‌یافته در شبکه‌ای از هنجارها، نهادها، منابع و روابط اجتماعی است. جامعه‌شناسی علم بر هنجارهای اجتماعی و سازمان علم تأکید می‌کند [۱۹ و ۲۰] و مطالعات اجتماعی علم و فناوری، توجه را به تولید اجتماعی دانش و نقش زمینه‌های نهادی و سیاسی معطوف کرده‌اند [۲۱ و ۲۲] از این منظر، پیشرفت علم صرفاً نتیجه عملکرد فردی پژوهشگران نیست و شرایط نهادی، شبکه‌های همکاری، اعتماد و ساختارهای پشتیبان نیز در امکان یا محدودیت آن نقش دارند.

۴-۴ سطح سنجش و ارزیابی از اندازه‌گیری کمی تا ارزیابی مسئولانه

با توسعه پایگاه‌های استنادی و روش‌های کتاب‌سنجی و علم‌سنجی، امکان مشاهده کمی ابعادی از فعالیت علمی در مقیاس وسیع فراهم شد. تعداد انتشارات، استنادها، شاخص‌های تأثیر و شبکه‌های همکاری به ابزارهای مهم پایش نظام علم تبدیل شدند. این تحول را باید عمدتاً تحول در «ظرفیت سنجش» دانست، نه اثبات اینکه پیشرفت علمی ذاتاً معادل افزایش این سنجه‌هاست [۱ و ۲].

سه ابتکار اصلاح ارزیابی، نقش‌های متمایزی دارند. DORA بر داوری خودِ دستاورد و پرهیز از کاربرد ضریب تأثیر مجله به‌جای کیفیت فرد یا مقاله تأکید دارد [۳]. مانیفست لیدن اصول استفاده از سنجه‌ها را مشخص می‌کند: پشتیبانی از داوری تخصصی، توجه به مأموریت و رشته، و شفافیت داده و محاسبه [۵]. CoARA این جهت‌گیری را به تعهد سازمانی برای اصلاح معیارها، تخصیص منابع و بازبینی پیشرفت اصلاحات پیوند می‌زند [۴]. اصول هنگ‌کنگ نیز تمامیت پژوهش را وارد ارزیابی می‌کنند [۲۳].

برای ایران، پیشنهاد این مقاله استفاده مرحله‌ای از این سه نقش است: لیدن راهنمای مناسب‌تری برای بازطراحی فوری

سنجه‌های موجود است؛ DORA قاعده‌ای روشن برای پرهیز از جان‌سپین کردن اعتبار مجله به‌جای کیفیت دستاورد فراهم می‌کند؛ و منطق تعهد و پیگیری CoARA می‌تواند الگوی اجرای سازمانی باشد. این ترجیح، داوری سیاستی مبتنی بر مسئله مقاله است و دربرگیرنده ادعای برتری تجربی یک ابتکار در ایران نیست.

۴-۵ سطح سیاستی-اثرمحور از تولید دانش به ارزش‌آفرینی در سطح سیاستی، پرسش از آنچه علم تولید می‌کند به این پرسش گسترش می‌یابد که علم چه مسئله‌ای را حل می‌کند و چه ارزشی برای جامعه ایجاد می‌کند. سنت‌های پراگماتیستی بر کارکرد دانش و پیوند معرفت و عمل تأکید داشته‌اند [۲۴ و ۲۵]؛ اما نباید میان این سنت فلسفی و چارچوب‌های جدید سیاست نوآوری رابطه تبارشناختی مستقیم فرض کرد. در سیاست علم معاصر، نظام‌های ملی نوآوری، تعامل دانشگاه - صنعت - دولت و سیاست‌های مأموریت‌گرا بر پیوند علم با فناوری، نوآوری و مسائل عمومی تأکید دارند [۲۸-۲۶]. در ادبیات داخلی نیز، مرور تاریخی سوزنچی کاشانی به تحول نقش علم، فناوری و نوآوری در توسعه می‌پردازد [۲۹] و میرعمادی کاربرد مفهوم نظام ملی نوآوری و محدودیت‌های آن را برای سیاست‌گذاری، از جمله در ایران، بررسی می‌کند [۳۰].

۴-۶ روابط متقابل سطوح و تغییر برجستگی تاریخی آن‌ها رابطه چهار سطح را می‌توان در یک چرخه توضیح داد: معیارهای معرفتی روشن می‌کنند چه چیزی دانش معتبر تلقی می‌شود؛ نهادها منابع، قواعد و فرصت تولید این دانش را فراهم می‌کنند؛ ابزارهای سنجش بخش‌هایی از فعالیت را قابل مشاهده می‌سازند؛ و سیاست‌گذاری بر اساس اهداف عمومی به تخصیص منابع و اصلاح قواعد بازمی‌گردد. رابطه فقط یک‌طرفه نیست: پیوند شاخص با پاداش می‌تواند رفتار نهادها و پژوهشگران را تغییر دهد و ترجیح انتشار بیشتر را بر اعتبار یا اهمیت مسئله مقدم سازد. در این حالت میان آنچه مطلوب دانسته می‌شود و آنچه پاداش می‌گیرد تنش ایجاد می‌شود؛ نقد ارزیابی مسئولانه متوجه همین تنش است [۵ و ۲۰]. این چرخه یک تبیین مفهومی از سازوکارهای ممکن است، نه مدل علی آزمون‌شده در داده‌های این پژوهش.

در بازسازی تاریخی حاضر، «غلبه» به برجستگی در گفتمان یا

گذاشته شد. همچنین میان طراحی سیاستی و تحقق عملکردی تفکیک می‌شود؛ بنابراین، وجود یک هدف یا شاخص در قانون به معنای تحقق آن تلقی نمی‌شود.

۵-۱ از ظرفیت‌سازی به اندازه‌گیری منابع پژوهش

در برنامه اول، جهت‌گیری غالب بر بازسازی ظرفیت علمی و آموزش عالی، تربیت پژوهشگر و گسترش ارتباطات علمی متمرکز بود و هنوز مجموعه‌ای منسجم از شاخص‌های علم‌سنجی شکل نگرفته بود. در برنامه‌های دوم و سوم، سهم تحقیق و توسعه از تولید ناخالص داخلی و ترکیب منابع مالی به‌عنوان شاخص‌های کلان وارد برنامه شد. بنابراین، منطق سنجش از ظرفیت‌سازی به سمت اندازه‌گیری نهاده‌ها و منابع نظام پژوهش گسترش یافت.

۵-۲ ورود فناوری، نوآوری و رقابت‌پذیری

در برنامه چهارم، دامنه سیاست علم از تأمین منابع فراتر رفت و انتقال و جذب فناوری، رقابت‌پذیری، پژوهش مأموریت‌گرا، همکاری بین‌المللی و استفاده از شاخص‌های جهانی وارد چارچوب برنامه شد. این تغییر نشان می‌دهد که در سطح طراحی سیاستی، پیشرفت به‌تدریج با نظام نوآوری و کاربرد دانش پیوند خورده است؛ با این حال، ورود یک هدف به قانون نباید با تحقق عملی آن یکسان فرض شود.

۵-۳ تثبیت سنجش کمی و گسترش دامنه شاخص‌ها

برنامه پنجم تنوع شاخص‌ها را افزایش داد و در کنار منابع پژوهشی، تولید علمی بین‌المللی، اختراع، شرکت‌های دانش‌بنیان، صادرات فناوری و کیفیت آموزش عالی را مطرح

ابزارهای سیاستی اشاره دارد، نه حذف سطوح دیگر. در میانه قرن بیستم، مناقشه فلسفی درباره پیشرفت عمدتاً در سطح معرفت‌شناختی صورت‌بندی می‌شد، درحالی‌که مدیریت علم هم‌زمان به سنجش نهاد و برون‌داد گرایش داشت [۲، ۱۶ و ۱۷]. از دهه‌های ۱۹۶۰ و ۱۹۷۰، گسترش آمار و پایگاه‌های علمی، سطح سنجش را در مدیریت برجسته‌تر کرد و جامعه‌شناسی علم نیز محدودیت نگاه فردمحور را توضیح داد [۱ و ۱۹]. از دهه ۱۹۹۰، نظام‌های نوآوری و تعامل دانشگاه، صنعت و دولت، پیوند نهادی و سیاستی را پررنگ‌تر کردند [۲۶ و ۲۸]. از دهه ۲۰۱۰، اصلاح ارزیابی و سیاست مأموریت‌گرا، کیفیت، اثر و مسئولیت را در برابر غلبه شاخص‌های جانشین قرار داده‌اند [۵ و ۲۷]. البته بایستی دقت نمود که این دوره‌ها هم‌پوشان‌اند و برتری عملی منطق کمی در همه کشورها پایان نیافته است.

جدول ۳ خلاصه این تمایزهاست. دقت شود که دامنه تحلیل به سنت‌های مرتبط با سؤال پژوهش محدود است و تاریخ جامع فلسفه یا جامعه‌شناسی علم را پوشش نمی‌دهد.

۵- تحول منطق سنجش پیشرفت علمی در سیاست‌های

توسعه ایران

برای بررسی بازتاب تحول مفهوم و سنجش پیشرفت علمی در سیاست‌گذاری ایران، قوانین برنامه‌های توسعه پس از انقلاب اسلامی تحلیل شدند. در این بررسی میان «هدف و جهت‌گیری سیاستی»، «شاخص سنجش‌پذیر» و «سازوکار پایش» تمایز

جدول ۳) چارچوب تحلیلی چهارسطحی مفهوم و سنجش پیشرفت علمی

سطح تحلیل	پرسش اصلی	نمونه دیدگاه‌ها/چارچوب‌ها	تلقی از پیشرفت	دلالت برای ارزیابی
معرفت‌شناختی	علم چگونه از نظر معرفتی پیشرفت می‌کند؟	اثبات‌گرایی؛ پوپر؛ کوهن؛ لاکاتوش	افزایش دانش معتبر، نقد و آزمون، حل مسائل و تحول چارچوب‌ها	کیفیت، اعتبار، استحکام شواهد و قدرت تبیین
اجتماعی-نهادی	چه شرایطی تولید و رشد علم را شکل می‌دهد؟	جامعه‌شناسی علم؛ مطالعات اجتماعی علم و فناوری	پیشرفت وابسته به هنجارها، نهادها، شبکه‌ها و زمینه اجتماعی	توجه به زمینه، همکاری، ساختار و عدالت نهادی
سنجش و ارزیابی	فعالیت و عملکرد علمی چگونه سنجیده شود؟	کتاب‌سنجی؛ علم‌سنجی؛ CoARA؛ Leiden؛ DORA	سنجش چندبعدی عملکرد و کیفیت	سنجش‌های کمی + داور کیفی و استفاده مسئولانه
سیاستی-اثرمحور	علم چه ارزشی برای جامعه ایجاد می‌کند؟	سیاست علم؛ نظام نوآوری؛ مأموریت‌گرایی	حل مسئله، نوآوری و اثر اجتماعی/اقتصادی	اثر اجتماعی، سیاستی، فناوریانه و مسئله‌گشایی

تجاری‌سازی و مسئله‌محوری دانست. از آنجا که برنامه هفتم در حال اجراست، این یافته صرفاً ناظر بر جهت‌گیری سیاستی است و نمی‌توان آن را شاهد تحقق‌گذار به نظام اثرمحور دانست. (جدول ۴)

۵-۵ اهداف در برابر عملکرد

اهداف و عملکرد بر پایه گزارش‌های مشخص در جداول ۵ و ۶ مقایسه شده‌اند. درصد تحقق فقط برای تعریف، واحد و سال مرجع قابل مقایسه محاسبه می‌شود؛ رتبه‌ها به درصد تبدیل نمی‌شوند. داده فاقد سال پایان دوره یا تعریف همسان، با قید محدودیت گزارش شده و داده‌های برنامه هفتم صرفاً میان‌دوره‌ای اند. GERD کل با اعتبارات پژوهشی دولت یکسان نیست؛ نسبت‌های مبتنی بر تفاوت تعریف، ارزیابی رسمی تحقق محسوب نمی‌شوند.

کرد. برنامه ششم این روند را با اهداف عددی دقیق‌تری برای رتبه تولید مقاله، سهم پژوهش از تولید ناخالص داخلی، ثبت اختراع، همکاری علمی بین‌المللی، تعداد پژوهشگران و سرانه مقاله ادامه داد. از منظر قابلیت پایش، این تحول نشان‌دهنده رسمی‌تر شدن نظام سنجش است؛ اما از منظر مفهومی، هم‌زمان بیانگر تثبیت منطق شاخص‌محور و رتبه‌محور در ارزیابی پیشرفت علمی است.

۵-۴ حرکت سیاستی به سوی اثرگذاری و مسئله‌محوری

در برنامه هفتم، در کنار استمرار شاخص‌هایی مانند رتبه علمی، تولید مقاله و اختراع، معیارهایی مرتبط با درآمد قراردادهای پژوهشی، تجاری‌سازی اختراعات، پایان‌نامه‌های نیازمحور، فروش شرکت‌های دانش‌بنیان و سهم محصولات فناورانه مورد توجه قرار گرفته‌اند. این تغییر را می‌توان نشانه گسترش طراحی سیاستی از سنجش تولیدات علمی به سمت پیامدهای اقتصادی،

جدول ۴) تحول منطق سنجش پیشرفت علمی در برنامه‌های توسعه ایران

برنامه	منطق غالب سنجش	نمونه شاخص‌ها/اهداف	سازوکار پایش	وضعیت تحلیل عملکرد
اول (۱۳۶۸-۱۳۷۲)	ظرفیت‌سازی	دانشجو، پژوهشگر، کیفیت تحقیق، ارتباط علمی	سازوکار یکپارچه مستقل تعیین نشده	فاقد شاخص ملی هم‌تعریف برای ارزیابی درصدی؛ تحلیل این برنامه بر ظرفیت‌سازی و احکام کیفی متمرکز است.
دوم (۱۳۷۴-۱۳۷۸)	منابع و نهاده‌ها	هدف R&D/GDP؛ سهم بودجه پژوهش	سازمان برنامه و بودجه و شورای پژوهش‌ها	داده پایان دوره هم‌تعریف محدود است.
سوم (۱۳۷۹-۱۳۸۳)	منابع + فرایند/تقاضا	R&D/GDP؛ ترکیب اعتبارات؛ پژوهش سفارشی	صندوق‌ها و سازوکارهای تأمین مالی	مقایسه با GERD کل و با تصریح تفاوت تعریف انجام شده است.
چهارم (۱۳۸۴-۱۳۸۸)	علم + فناوری + نوآوری	R&D؛ فناوری؛ رقابت‌پذیری؛ مأموریت‌گرایی	شبکه واحدهای پژوهشی و فناوری	به علت تفاوت تعریف GERD و اعتبارات عمومی، درصد تحقق رسمی محاسبه نشده است.
پنجم (۱۳۹۰-۱۳۹۴)	نظام جامع چندشاخصی	مقاله؛ اختراع؛ دانش‌بنیان؛ صادرات فناوری؛ کیفیت آموزش	نظام جامع پایش و گزارش سالانه	دو برآورد رسمی/بین‌المللی برای عملکرد گزارش شده‌اند.
ششم (۱۳۹۶-۱۴۰۰)	داشبورد کمی هدف‌دار	رتبه، h-index، مقاله، R&D، اختراع، همکاری بین‌المللی	اهداف سالانه و گزارش به مجلس	برای بسیاری از شاخص‌ها داده پایان ۱۴۰۰ در دسترس نیست.
هفتم (۱۴۰۳-۱۴۰۷)	اثر اقتصادی و مسئله‌محوری	تجاری‌سازی، قرارداد، نیازمحوری، فروش دانش‌بنیان، رتبه و مقاله	سنجش‌های عملکردی و گزارش سالانه	در حال اجرا؛ پایش میان‌دوره‌ای سال‌های ۱۴۰۳ و ۱۴۰۴ گزارش شده و ارزیابی تحقق نهایی زود هنگام است.

جدول ۵) ارزیابی کلان برنامه‌های اول تا هفتم

برنامه	هدف یا شاخص قانونی اصلی	عملکرد مستند	ارزیابی قابل دفاع
اول	توسعه ظرفیت‌های علمی و تحقیقاتی؛ پیوند دانشگاه، تولید و پژوهش؛ تدوین نظام تحقیقات علمی. اهداف غالباً نهادی و کیفی‌اند.	برای هدف‌های کیفی این برنامه، شاخص ملکی هم‌تعریف و خط پایه‌ای که ارزیابی کمی پایان دوره را ممکن کند یافت نشد.	ارزیابی درصدی ممکن نیست. برنامه باید بر مبنای نهادسازی، توسعه ظرفیت و اجرای احکام کیفی بررسی شود. اطلاعات کمی در دسترس نیست.
دوم	افزایش سهم هزینه‌های تحقیق از GDP به ۱٫۵٪ در ۱۳۷۸؛ تأمین ۵۰٪ هزینه‌های تحقیق از درآمدهای عمومی.	گزارش رسمی ارزیابی علم و فناوری، سهم اعتبارات تحقیقاتی دولتی از GDP را در ۱۳۷۵ برابر ۰٫۳۱٪، در ۱۳۷۷ برابر ۰٫۳۰٪ و در ۱۳۷۹ برابر ۰٫۲۹٪ گزارش می‌کند. رقم کل R&D برای پایان ۱۳۷۸ با همان تعریف هدف یافت نشد.	شواهد موجود از فاصله جدی با هدف حکایت دارد؛ اما درصد نهایی تحقق محاسبه نمی‌شود. اطلاعات کمی هم‌تعریف برای پایان برنامه در دسترس نیست.
سوم	افزایش هزینه تحقیقات تا پایان برنامه به ۱٪ GDP در بخش دولتی و ۰٫۵٪ در بخش غیردولتی؛ مجموع هدف‌ها ۱٫۵٪ GDP.	GERD کل ایران در ۲۰۰۴ برابر ۵۰۴۵۳٪، GDP گزارش شده است. تفکیک هم‌تعریف دولتی/غیردولتی برای پایان برنامه یافت نشد.	نسبت توصیفی GERD کل به مجموع هدف‌ها حدود ۳۳٫۶٪ است؛ به دلیل تفاوت تعریف و سال، این رقم درصد رسمی تحقق نیست. تحقق تفکیکی قابل محاسبه نیست.
چهارم	افزایش یکنواخت سرمایه‌گذاری دولت در پژوهش و فناوری تا حداقل ۲٪ GDP از اعتبارات عمومی؛ اختصاص ۱٪ درآمد عملیاتی شرکت‌های دولتی، بانک‌ها و مؤسسات انتفاعی به پژوهش.	GERD کل ایران در ۲۰۰۹ برابر ۲۷۲۵۸٪، GDP است. این شاخص کل هزینه تحقیق و توسعه است و با «اعتبارات عمومی دولت» یکسان نیست.	GERD کل پایین‌تر از سطح هدف است؛ به علت تفاوت هزینه تحقق یافته و اعتبارات عمومی، این مقایسه قرینه محدودیت مالی است و درصد تحقق رسمی محاسبه نمی‌شود.
پنجم	افزایش سالانه ۰٫۵ واحد درصدی سهم تحقیق و پژوهش تا دستیابی به ۳٪ GDP در پایان برنامه.	سری به‌روزرشده UIS/بانک جهانی برای ۲۰۱۵ مقدار ۰٫۳۹۳۸۳٪ GDP را ثبت می‌کند. گزارش UNCTAD (۲۰۱۶) با استاندارد به مرکز پژوهش‌های مجلس، رقم ۰٫۴۷٪ را گزارش کرده است.	نسبت توصیفی دو رقم به سطح هدف ۳٪، به ترتیب ۱۳٫۱٪ و ۱۵٫۷٪ است. تفاوت تعریف و سال مرجع مانع تلقی آن‌ها به عنوان درصد رسمی تحقق است.
ششم	جدول ۱۳ ماده ۶۶ شامل ۱۲ شاخص: رتبه تولید علم، سهم اعتبارات پژوهشی، ثبت اختراع بین‌المللی، فناوری متوسط و بالا، همکاری بین‌المللی، پژوهشگر و نشریه.	عملکرد تفصیلی در جدول ۵ آمده است. برای بسیاری از شاخص‌ها، آخرین رقم قابل اتکا مربوط به ۱۳۹۸ یا ۱۳۹۹ است و رقم پایان ۱۴۰۰ منتشر نشده یا هم‌تعریف نیست.	ارائه یک درصد تجمیعی برای «تحقق علم و فناوری» معتبر نیست. رقم ۵۷٪ گزارش شماره ۱۸۳۰۹ مربوط به احکام بخش تجمیعی آموزش عمومی، آموزش عالی و علم و فناوری و مبتنی بر سال‌های ۱۳۹۶-۱۳۹۸ است، نه تحقق نهایی ۱۲ شاخص جدول ۱۳.
هفتم	از جمله: رتبه ۱۴ جهان در کمیت تولید علم؛ سهم ۲٪ اعتبارات پژوهش و فناوری بخش دولتی از GDP؛ توسعه همکاری علمی بین‌المللی، درآمد پژوهشی و پایان‌نامه‌های تقاضا محور.	در ۱۴۰۳ رتبه کمی تولید علم ۱۷ بود (هدف سال اول ۱۵). نسبت اعتبارات پژوهش و فناوری بخش دولتی/GDP در محاسبه مرکز پژوهش‌ها ۰٫۱۹٪ برای ۱۴۰۳ و حدود ۰٫۲۶٪ برای ۱۴۰۴ برآورد شده است. از ۱۲ هدف کمی سال اول: ۱ تحقق کامل، ۱ تحقق زیاد، ۶ تحقق کم و ۴ عدم تحقق.	در شاخص مالی، ۰٫۱۹٪ معادل ۹٫۵٪ سطح هدف نهایی ۲٪ است؛ این فقط فاصله سال اول با هدف نهایی است و نباید «درصد تحقق برنامه هفتم» نامیده شود. رتبه علمی نیز به درصد تبدیل نمی‌شود.

جدول ۶) ارزیابی تفصیلی شاخص‌های جدول ۱۳ برنامه ششم

(هرگاه داده ۱۴۰۰ در دسترس نبوده، آخرین سال دارای رقم معتبر درج و مقایسه فقط با هدف همان سال انجام شده است.)

ردیف	شاخص	هدف ۱۴۰۰	آخرین عملکرد قابل اتکا	ارزیابی
۱	رتبه جهانی تولید کمی مقالات	رتبه ۱۲	رتبه ۱۵ در ۲۰۲۱ بر پایه گزارش رسمی سازمان امور دانشجویان/وزارت علوم. [۳۶]	سه رتبه فاصله با هدف؛ عدم تحقق. برای مقیاس رتبه‌ای درصد محاسبه نمی‌شود.
۲	رتبه جهانی شاخص هرش	رتبه ۴۰	رتبه ۴۱ در ۲۰۲۱. [۳۶]	یک رتبه فاصله با هدف؛ نزدیک به هدف ولی محقق نشده. درصد محاسبه نمی‌شود.
۳	سهم اعتبارات پژوهش و فناوری بخش دولتی از GDP	۱٫۵٪	۰٫۴۳٪ در ۱۳۹۹؛ هدف همان سال ۱٫۳۵٪. برای ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	تحقق نسبت به هدف همان سال: ۳۱٫۹٪. ارزیابی نهایی ۱۴۰۰ ممکن نیست.
۴	اختراعات و ابداعات ثبت شده سالانه در مراجع بین‌المللی	۵۰	۱۲۵ مورد در ۱۳۹۸؛ هدف همان سال ۴۳. برای ۱۳۹۹ و ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	عملکرد ۱۳۹۸ معادل ۲۹۰٫۷٪ هدف همان سال و فراتر از هدف نهایی بوده است؛ پایداری تا ۱۴۰۰ قابل احراز نیست.
۵	محصولات با فناوری متوسط به بالا از کل محصولات صنعتی	۵۰٪	۳۸٫۵٪ در ۱۳۹۸؛ هدف همان سال ۴۵٪. برای ۱۳۹۹ و ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	تحقق نسبت به هدف همان سال: ۸۵٫۶٪؛ پایین‌تر از هدف ۱۳۹۸ و فاقد داده پایان برنامه.
۶	سهم محصولات با فناوری متوسط به بالا از GDP	۵٪	۴٫۶۲٪ در ۱۳۹۸؛ هدف همان سال ۳٫۴٪. برای ۱۳۹۹ و ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	عملکرد ۱۳۹۸ از هدف همان سال بالاتر بوده است؛ به دلیل نبود داده پایان ۱۴۰۰، تحقق هدف نهایی ۵٪ قابل تأیید نیست.
۷	رتبه منطقه‌ای صادرات محصولات با فناوری بالا	رتبه ۳	رتبه ۱۱ در ۱۳۹۸؛ هدف همان سال رتبه ۴. برای ۱۳۹۹ و ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	هفت رتبه فاصله با هدف همان سال؛ عدم تحقق روشن. درصد محاسبه نمی‌شود.
۸	سرانه سالانه مقالات Scopus به اعضای هیئت علمی تمام وقت	۰٫۹۵ مقاله/نفر	۰٫۷۵ در ۱۳۹۹؛ هدف همان سال ۰٫۸۵. برای ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	تحقق نسبت به هدف همان سال: ۸۸٫۲٪؛ داده پایان برنامه در دسترس نیست.
۹	مقالات مشترک با محققان خارجی از کل	۳۵٪	۲۸٪ در ۱۳۹۸؛ هدف همان سال ۳۰٪. برای ۱۳۹۹ و ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	تحقق نسبت به هدف همان سال: ۹۳٫۳٪؛ نتیجه نهایی ۱۴۰۰ قابل محاسبه نیست.
۱۰	پژوهشگران به ازای یک میلیون نفر	۲۶۰۰ نفر	۲۸۷۱ نفر در ۱۳۹۸؛ هدف همان سال ۲۲۹۰. برای ۱۳۹۹ و ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	عدد عملکرد از هدف همان سال بالاتر است؛ اما منبع، ناسازگاری خط پایه و روش آمارگیری را تصریح کرده است، بنابراین نتیجه‌گیری قطعی درباره تحقق معتبر نیست.
۱۱	مقالات نمایه شده در ISC؛ دو برابر وضع موجود	۵۸۶٫۳۸۸ مقاله	۳۸۶٫۰۴۶ مقاله در ۱۳۹۷؛ هدف همان سال ۴۱۰٫۴۷۲. داده ۱۳۹۸ ناقص و برای ۱۳۹۹-۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	تحقق نسبت به هدف ۱۳۹۷: ۹۴٫۰٪. به دلیل فقدان داده کامل، تحقق هدف دوبرابرشدن در پایان برنامه قابل ارزیابی نیست.
۱۲	نشریات ایرانی نمایه شده در پایگاه بین‌المللی علم و دارای ضریب تأثیر	۷۰ نشریه	۷۱ نشریه در ۱۳۹۹؛ هدف همان سال ۶۵. برای ۱۴۰۰ اطلاعات کمی در دسترس نیست. [۳۴]	هدف نهایی تا ۱۳۹۹ از نظر تعداد عبور کرده بود (۱۰۹٫۲٪ هدف همان سال)؛ رقم پایان ۱۴۰۰ در دسترس نیست.

۵-۶- تفسیر تفاوت الگوهای تحقق

جدول ۶ نشان می‌دهد بعضی شاخص‌های شمارشی، از جمله اختراعات ثبت شده و نشریات نمایه شده، در سال‌های گزارش شده از هدف همان سال عبور کرده‌اند؛ سهم اعتبارات

پژوهشی دولت در ۱۳۹۹ فقط ۳۱٫۹ درصد هدف همان سال بوده است. این ناهمگونی به معنای جبران کمبود مالی با رشد برونداد نیست؛ تعداد اختراع، کیفیت و بهره‌برداری آن را اندازه نمی‌گیرد و افزایش نشریات نیز معادل بهبود کیفیت همه

سالانه، امکان مقایسه و سهولت پایش اداری به استمرار شاخص‌های کمی کمک می‌کنند؛ باین‌حال، اثبات علل نهادی، بودجه‌ای و سازمانی این استمرار نیازمند مطالعه‌ای مستقل درباره فرایند سیاست‌گذاری و ذی‌نفعان است.

۶- مبانی هنجاری بومی برای بازنگری در سنجش پیشرفت علمی در ایران

این بخش به جزء سوم سؤال پژوهش، یعنی دلالت‌های بازنگری ارزیابی در ایران، پاسخ می‌دهد. تحلیل پیشین نشان داد اهداف رسمی گسترده‌تر از سنجش‌های رایج‌اند؛ اکنون پرسش آن است که اصولی مانند کیفیت، عدالت و حل مسئله چه پشتوانه و چه محدودیت هنجاری در زمینه ایران دارند. سنت فکری برای توضیح اعتبار و زمینه اجتماعی علم، منابع قرآنی برای اصول داوری و مسئولیت، و بیانات رسمی برای جهت‌گیری نهادی سیاست علم بررسی شدند. حاصل، سنجش امکان توجیه هنجاری اصلاحات است؛ ولی پذیرش ذی‌نفعان و اثربخشی اجرا به شواهد تجربی جداگانه نیاز دارد. شایان توجه است که مقایسه در سطح اصول انجام شد. دلالت تاریخی یا تفسیری منبع از کاربرد امروزی آن جداست؛ هیچ شاخص علم‌سنجی به متون کلاسیک نسبت داده نمی‌شود و تشابه ارزشی، یکسانی نظری دو سنت را اثبات نمی‌کند.

۶-۱ سنت فکری اسلامی و چندبعدی بودن علم

آثار کلاسیک اسلامی مستقیماً درباره نظام‌های ارزیابی کنونی نیستند؛ باین‌حال، برهان در فارابی و ابن‌سینا، مشاهده در بیرونی، سازمان دانش در طوسی و زمینه اجتماعی علم در ابن‌خلدون، امکان مقایسه محدود در سطح اصول را فراهم می‌کنند [۴۰-۳۸ و ۹-۷]. (جدول ۷)

۶-۲ اصول معرفتی و اخلاقی مرتبط با علم در قرآن

در جدول ۸، برداشت‌های مربوط به اعتبار معرفت، عدالت و مسئولیت با منابع تفسیری مشخص همراه شده‌اند. کاربرد این اصول در ارزیابی علم، استنباط مقاله است؛ لذا آیات تعریف فنی شاخص یا روش محاسبه ارائه نمی‌کنند. آیات بر اساس نام سوره و شماره آیه در قرآن کریم مشخص شده‌اند.

۶-۳ دیدگاه رهبران انقلاب به عنوان منبع سیاستی و نهادی

برای روشن‌شدن مبنای مقایسه، گزاره‌های مستند جدول ۹ در

پژوهش‌ها نیست. عبور یک شاخص از هدف می‌تواند علاوه بر عملکرد، از سطح هدف‌گذاری و تغییر پوشش آماری تأثیر بگیرد؛ درباره تعداد پژوهشگران، ناسازگاری خط پایه صراحتاً مانع نتیجه‌گیری قطعی است [۳۴].

یک تبیین احتمالی این است که برخی برون‌دادها از ظرفیت انباشته، همکاری و مشوق‌های حرفه‌ای تغذیه می‌کنند، اما تحقق هدف مالی به تخصیص منابع و هماهنگی دستگاه‌ها وابسته است؛ بنابراین زمان و سازوکار تغییر آن‌ها یکسان نیست. داده‌های این مطالعه برای آزمون این تبیین کافی نیستند. گزارش برنامه هفتم، کمبود منابع و تأخیر در اسناد اجرایی را مطرح می‌کند، اما تصریح دارد که برخی اصلاحات رویه‌ای با ظرفیت موجود نیز قابل انجام بوده‌اند [۳۷] پس همه عدم تحقق‌ها را نمی‌توان به بودجه نسبت داد یا نتیجه برنامه هفتم را به دوره‌های پیشین تعمیم داد.

پیامد تحلیلی این تفاوت، تفکیک پایش نهاده، برون‌داد و اثر و ثبت خط پایه و تعریف ثابت هر شاخص است. مقایسه رشد تعداد مقالات با سهم مالی نیز به‌تنهایی بهره‌وری یا رابطه علی را نشان نمی‌دهد. نبود داده هم‌تعریف، محدودیت واقعی ارزیابی است و نباید به‌عنوان صفر عملکرد یا تحقق هدف تفسیر شود.

۵-۷ همگرایی در اهداف و شکاف در منطق سنجش

تحلیل برنامه‌های توسعه نشان می‌دهد مفهوم پیشرفت علمی در سیاست‌گذاری ایران از ظرفیت‌سازی آموزشی و پژوهشی به تأمین منابع، تولید علمی، فناوری و نوآوری و سپس به اثر اقتصادی و مسئله‌محوری گسترش یافته است. بنابراین، در سطح اهداف سیاستی نوعی همگرایی با تحول بین‌المللی از ارزیابی تک‌بعدی تولید به سوی تلقی گسترده‌تر از اثر علم مشاهده می‌شود. باین‌حال، این همگرایی کامل نیست؛ بخش مهمی از سیاست‌های ارزیابی ایران همچنان در قالب اهداف عددی مربوط به رتبه، مقاله، استناد، اختراع و دیگر سنجش‌های قابل شمارش صورت‌بندی شده است. این وضعیت را می‌توان «شکاف مفهوم - سنجش» نامید: مفهوم سیاستی پیشرفت علمی چندبعدی‌تر شده، اما ابزارهای رسمی سنجش با همان سرعت چندبعدی نشده‌اند. از ساختار اسناد، به‌ویژه برنامه ششم، می‌توان استنباط کرد که قابلیت گزارش‌دهی، هدف‌گذاری

گزاره مستقیم به کار می‌رود. این طبقه‌بندی تفسیر کیفی است و ضریب انطباق یا حاصل آزمون آماری نیست. انتخاب منابع هدفمند و محدود است؛ نتیجه به همه بیانات تعمیم داده نمی‌شود و از فقدان شاهد در این مجموعه، فقدان دیدگاه استنتاج نمی‌گردد.

بازنگری حاضر با شش محور ارزیابی مسئولانه تطبیق داده شدند. واحد بررسی، گزاره همراه با مشخصات منبع است؛ خلاصه مضمون، تاریخ و ارجاع آن ثبت و سپس نسبت آن با محور جدول بررسی شده است. «تصریح موضوعی» برای اشاره مستقیم به همان مسئله، «پشتوانه عام» برای اصل اخلاقی یا سیاستی مرتبط، و «فاقد شاهد در منابع بررسی‌شده» برای نبود

جدول (۷) برخی مؤلفه‌های مرتبط با علم معتبر در سنت فکری اسلامی

متفکر	تلقی از علم معتبر	تأکید روش‌شناختی/نهادی	دلالت قابل استفاده در این مقاله
فارابی	علم نظام‌مند و عقلانی	منطق، برهان و نظم معرفت	اهمیت کیفیت معرفتی و سازمان دانش
ابن سینا	معرفت برهانی و عکلی	تعریف، قیاس، منطق و برهان	اعتبار و استحکام استدلال
ابوریحان بیرونی	معرفت مبتنی بر مشاهده و سنجش	مشاهده، اندازه‌گیری، مقایسه و نقد تعصب	شواهد، دقت و قابلیت بررسی
خواجه نصیرالدین طوسی	پیوند دانش نظری و نهاد علمی	منطق، ریاضیات، مشاهده و سازمان پژوهش	اهمیت زیرساخت و نهاد علمی
ابن خلدون	علم در بستر جامعه و تمدن	تحلیل عقلی و اجتماعی شرایط علم	زمینه‌مندی و شرایط نهادی پیشرفت

جدول (۸) اصول هنجاری و نسبت آن‌ها با ارزیابی علم

(دلالت تفسیری و کاربرد پیشنهادی از یکدیگر تفکیک شده‌اند؛ قیده‌های سیاق و دامنه تعمیم در ستون پایانی آمده‌اند.)

اصل و آیه	دلالت تفسیری و منبع	کاربرد پیشنهادی در ارزیابی	قید تفسیر
معرفت معتبر اسراء، ۳۶	پرهیز از اعتقاد، گفتار و عمل فاقد علم؛ رجوع به متخصص در حوزه تخصص	اتکا به شواهد قابل بررسی و داوری تخصصی.	آیه تعریف فنی شواهد علمی ارائه نمی‌کند.
راستی آزمایی حجرات، ۶	بررسی خبر فاقد وثاقت پیش از اقدام زیان‌بار (ج ۱۸، صص ۴۶۴-۴۶۶)	بررسی منشأ و سازگاری داده‌های عملکرد.	موضوع مستقیم، خبر فاسق است؛ تعمیم به داده‌ها کاربرد تحلیل است.
مطالبه دلیل بقره، ۱۱۱	مطالبه برهان برای ادعای نجات انحصاری ج ۱)	مستندسازی ادعای رتبه، تحقق و اثر.	سیاق دینی است؛ کاربرد سیاستی، استنباط مقاله است.
عدالت و میزان الرحمن، ۷-۹	اقامه وزن به قسط و پرهیز از کاستن و طغیان در میزان	شفافیت معیار، تعریف ثابت و مقایسه منصفانه.	قواعد آماری از روش‌شناسی تخصصی به دست می‌آیند.
انتخاب سنجیده زمر، ۱۸	شنیدن و ترجیح حق و رشد؛ با اختلاف تفسیری درباره دامنه «قول»	مقایسه شواهد و تصریح معیار ترجیح.	«احسن» معادل ترجیح شخصی یا تعمیم نامحدود نیست.
کوشش و مسئولیت نجم، ۳۹-۴۰؛ توبه، ۱۰۵	ظهور سعی و اعمال و مسئولیت نسبت به آن‌ها	توجه به فعالیت و آثار مستند فراتر از ادعا.	سیاق اخلاقی و اخروی است؛ کوشش با اثر یکسان نیست.

جدول ۹) دیدگاه رهبران انقلاب به عنوان منبع سیاستی و نهادی

محور	مضمون منابع بررسی شده	نوع پشتوانه و حدود دلالت
اعتبار و کیفیت	ارتقای تخصص همراه با اخلاق ^۱ ؛ تقدم کیفیت و جهت مقاله بر مقاله‌بودن آن ^۲ .	برای کیفیت مقاله تصریح موضوعی؛ برای تخصص و اخلاق پشتوانه عام.
استفاده از سنجه‌ها	نقد مقاله‌محوری ارزیابی استاد و تأکید بر حل مسئله ^۳ .	تصریح موضوعی درباره نقد معیار تک‌بعدی؛ در منابع منتخب امام خمینی شاهد مستقیم سنجه استنادی وجود ندارد.
تنوع دستاوردها	دانشگاه متناسب با نیاز ملت ^۴ ؛ تنوع تولید علم، فناوری، ثروت، معرفت و نیروی کارآمد ^۵ .	پشتوانه عام برای تنوع کارکردها؛ فاقد دستور فنی وزن‌دهی خروجی‌ها.
اثر اجتماعی	نیاز ملت و کارکرد دانشگاه ^۶ ؛ علم نافع و جهت‌گیری مقاله و پایان‌نامه به مسائل کشور ^۷ ؛ ^۸ .	تصریح موضوعی برای نیازمحوری؛ اثبات اثر هر پژوهش به شواهد مستقل نیاز دارد.
زمینه و مأموریت	نیازهای ملی و تنوع حوزه‌های علم نافع ^۹ ؛ ^{۱۰} .	پشتوانه عام برای زمینه ملی؛ شاهد مستقیم وزن‌دهی رشته‌ای در این منابع وجود ندارد.
مسئولیت و اخلاق	پیوند تخصص و تعهد ^{۱۱} ؛ انصاف در برخورد علمی ^{۱۲} .	تصریح در سطح اخلاق علمی؛ استاندارد اجرایی تمامیت پژوهش باید جداگانه تعریف شود.

۴-۶ هم‌پوشانی هنجاری با ارزیابی مسئولانه

جدول ۱۰ مسیر تبدیل پشتوانه هنجاری به مسئله ارزیابی را خلاصه می‌کند. این پشتوانه برای طراحی وزن شاخص‌ها کافی نیست؛ مثلاً عدالت، لزوم توجه به تفاوت مأموریت‌ها را موجه می‌کند، اما وزن مناسب را تعیین نمی‌کند. همچنین سودمندی اجتماعی نباید به حذف پژوهش بنیادی با افق اثر بلندمدت بینجامد. از این رو قواعد اجرایی همچنان باید با روش‌های تخصصی ارزیابی و مشارکت جامعه علمی طراحی شوند.

۷- بحث

۱-۷ سنتز یافته‌ها

تلفیق یافته‌ها نشان می‌دهد تنوع موضوعی ادبیات، روابط چهار سطح و گسترش اهداف برنامه‌های ایران، همگی بر ناکافی بودن یک سنجه منفرد دلالت دارند. سهم چارچوب حاضر در اتصال این شواهد است: معیار معرفتی مشخص می‌کند چه دستاوردی معتبر است؛ تحلیل نهادی امکان تولید آن را توضیح می‌دهد؛ سنجش میزان و الگوی فعالیت را آشکار می‌کند؛ و تحلیل سیاستی ارزش عمومی مورد انتظار را روشن می‌سازد.

۱ امام خمینی، س. ر. (۱۳۵۹، ۲۷ آذر). سخنرانی در جمع طلاب و دانشجویان. در صحیفه امام (ج ۱۳، ص ۴۱۷). مؤسسه تنظیم و نشر آثار امام خمینی.

۲ امام خامنه‌ای، س. ع. (۱۳۹۰، ۱۳ مهر). بیانات در دیدار جمعی از نخبگان و برگزیدگان علمی.

۳ امام خامنه‌ای، س. ع. (۱۴۰۱، ۲۷ مهر). بیانات در دیدار نخبگان و استعدادهای برتر تحصیلی.

۴ امام خمینی، س. ر. (۱۳۶۰، ۴ خرداد). سخنرانی در جمع دانشجویان دانشگاه علم و صنعت. در صحیفه امام (ج ۱۴، صص ۳۵۷-۳۶۰). مؤسسه تنظیم و نشر آثار امام خمینی.

۵ امام خامنه‌ای، س. ع. (۱۳۸۵، ۱ فروردین). بیانات در دیدار زائران و مجاوران حرم مطهر رضوی.

۶ امام خمینی، س. ر. (۱۳۶۰، ۴ خرداد). سخنرانی در جمع دانشجویان دانشگاه علم و صنعت. در صحیفه امام (ج ۱۴، صص ۳۵۷-۳۶۰). مؤسسه تنظیم و نشر آثار امام خمینی.

۷ امام خامنه‌ای، س. ع. (۱۳۹۸، ۸ خرداد). بیانات در دیدار جمعی از استادان، نخبگان و پژوهشگران دانشگاه‌ها.

۸ امام خامنه‌ای، س. ع. (۱۴۰۲، ۲۵ مهر). بیانات در دیدار نخبگان و استعدادهای برتر علمی.

۹ امام خمینی، س. ر. (۱۳۶۰، ۴ خرداد). سخنرانی در جمع دانشجویان دانشگاه علم و صنعت. در صحیفه امام (ج ۱۴، صص ۳۵۷-۳۶۰). مؤسسه تنظیم و نشر آثار امام خمینی.

۱۰ امام خامنه‌ای، س. ع. (۱۳۹۸، ۸ خرداد). بیانات در دیدار جمعی از استادان، نخبگان و پژوهشگران دانشگاه‌ها.

۱۱ امام خمینی، س. ر. (۱۳۵۹، ۲۷ آذر). سخنرانی در جمع طلاب و دانشجویان. در صحیفه امام (ج ۱۳، ص ۴۱۷). مؤسسه تنظیم و نشر آثار امام خمینی.

۱۲ امام خامنه‌ای، س. ع. (۱۳۸۹، ۲ آبان). بیانات در دیدار جمعی از استادان، فضلا و طلاب نخبه حوزه علمیه قم.

جدول ۱۰) ماتریس هم‌پوشانی هنجاری برای بومی‌سازی اصلاح ارزیابی پژوهش

محور	جهت‌گیری ارزیابی مسئولانه	ظرفیت هنجاری بومی	دلالت برای ایران
اعتبار و کیفیت	ارزیابی کیفی و تخصصی؛ عدم استفاده جانشین‌وار از شاخص‌ها	علم معتبر، برهان و دقت	کاهش اتکای صرف به شمارش برونداد
استفاده مسئولانه از سنج‌ها	CoARA و Leiden: زمینه‌مندی و شفافیت	عدالت و میزان	پرهیز از کاربرد یکسان شاخص‌ها در همه زمینه‌ها
تنوع دستاوردها	CoARA: شناسایی انواع فعالیت و خروجی	نگاه چندبعدی به علم و کارکرد آن	شناسایی فناوری، آموزش، حل مسئله و سایر خروجی‌ها
اثر اجتماعی	توجه به Societal Impact	سودمندی و رفع نیاز	افزودن اثر اجتماعی و حل مسئله به ارزیابی
زمینه‌مندی	تفاوت رشته‌ها و مأموریت‌ها	توجه به شرایط نهادی و اجتماعی	ارزیابی متناسب با مأموریت و رشته
مسئولیت و اخلاق	Research Integrity و ارزیابی مسئولانه	مسئولیت، عدالت و صدق	افزودن تمامیت و مسئولیت پژوهشی

۷-۲ «شکاف مفهوم – سنجش» در ایران

«شکاف مفهوم – سنجش» به ناهم‌زمانی گسترش اهداف و تنوع معیارها اشاره دارد؛ «شکاف هدف – عملکرد» فاصله یک هدف مشخص با عملکرد هم‌تعریف آن است. ممکن است شاخصی محقق شود اما بخش مهمی از کیفیت یا اثر علم را نسنجد؛ بنابراین رفع شکاف دوم لزوماً شکاف نخست را برطرف نمی‌کند.

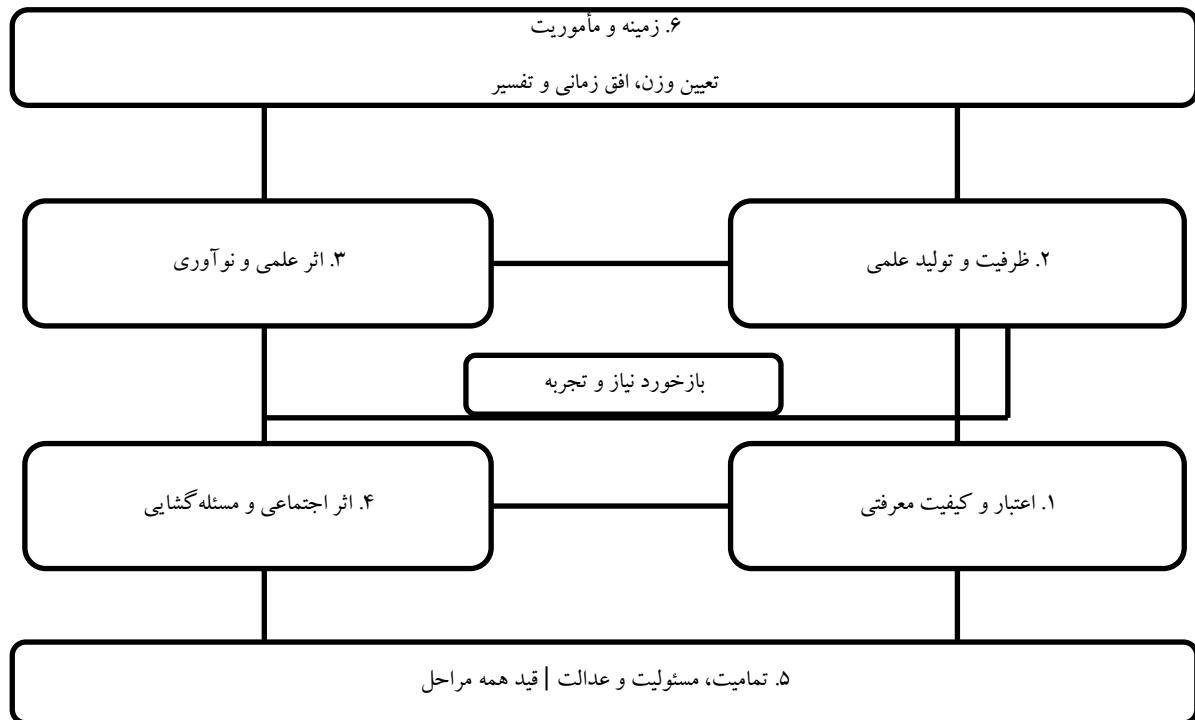
۷-۳ چارچوب چندبعدی پیشنهادی برای سنجش پیشرفت علمی

ابعاد جدول ۱۱ هم‌وزن و قابل جایگزینی کامل نیستند. ظرفیت

و تولید، امکان فعالیت علمی را فراهم می‌کنند؛ اعتبار معرفتی شرط اعتماد به دستاورد است؛ اثر علمی و نوآوری می‌توانند یکی از مسیرهای اثر اجتماعی باشند، اما اثر اجتماعی از مسیر آموزش یا مشارکت مستقیم در سیاست نیز ممکن است. تمامیت و عدالت بر همه مراحل قید می‌گذارند و زمینه و مأموریت، وزن، افق زمانی و تفسیر شواهد را تعدیل می‌کنند. بنابراین افزایش تعداد مقاله نمی‌تواند ضعف اعتبار یا نقض اخلاق را جبران کند. شکل ۴ این روابط پیشنهادی و بازخورد نیازهای جامعه به ظرفیت و اولویت پژوهش را نشان می‌دهد؛ پیکان‌ها فرض مفهومی‌اند، نه ضرایب علی برآوردشده.

جدول ۱۱) چارچوب شش‌بعدی پیشنهادی برای سنجش پیشرفت علمی در ایران

بعد	پرسش اصلی	نمونه حوزه‌های سنجش	نوع شواهد
۱. اعتبار و کیفیت معرفتی	آیا دانش تولیدشده از اعتبار و کیفیت علمی برخوردار است؟	کیفیت روش، استحکام شواهد، داوری تخصصی، بازبینی و قابلیت اعتماد	کیفی + شواهد کمی مکمل
۲. ظرفیت و تولید علمی	نظام علمی چه ظرفیت و بروندادی دارد؟	منابع انسانی، R&D، مقاله، کتاب، داده، زیرساخت	کمی؛ با تفسیر زمینه‌ای
۳. اثر علمی و نوآوری	پژوهش چه تغییری در علم و فناوری ایجاد می‌کند؟	استاد، نوآوری، فناوری، اختراع، مسیرهای جدید دانش	کمی + کیفی
۴. اثر اجتماعی و مسئله‌گشایی	علم چه مسئله‌ای از جامعه، اقتصاد یا حکمرانی حل می‌کند؟	اثر اجتماعی، سیاستی، اقتصادی، سلامت، محیط‌زیست، نیازهای ملی	شواهد موردی + شاخص‌های اثر
۵. تمامیت، مسئولیت و عدالت	علم چگونه تولید و ارزیابی می‌شود؟	تمامیت علمی، شفافیت، اخلاق، عدالت و مسئولیت	شاخص‌های فرایندی + ارزیابی کیفی
۶. زمینه و مأموریت	آیا معیارها با رشته، مؤسسه و مأموریت آن متناسب‌اند؟	تفاوت رشته‌ها، مأموریت سازمان، شرایط منطقه‌ای و بومی	تعدیل وزن و تفسیر شاخص‌ها



شکل ۴) روابط پیشنهادی ابعاد ارزیابی پیشرفت علمی

۴-۷ معماری اجرایی پیشنهادی

اجرای چارچوب در سه لایه پیشنهاد می‌شود: (۱) لایه پایش پایه، شواهد ظرفیت و تولید را ثبت می‌کند؛ (۲) لایه داور، کیفیت و آثار علمی و اجتماعی را بر پایه نمونه دستاوردها و شواهد اثر بررسی می‌کند؛ و (۳) لایه زمینه‌ای و هنجاری بر هر دو نظارت دارد. لایه سوم مرحله‌ای پس از محاسبه امتیاز نیست: مأموریت و افق زمانی باید پیش از انتخاب شاخص مشخص شوند و تمامیت و عدالت در گردآوری داده و داور نیز رعایت شوند.

برای مثال، پژوهشگاه مأموریت‌گرا می‌تواند سهم بیشتری برای شواهد حل مسئله در نظر بگیرد، درحالی‌که در علوم پایه باید افق بلندتر اثر و ارزش پیشرفت معرفتی لحاظ شود. وزن‌ها در این مقاله تعیین نشده‌اند؛ تعیین آن‌ها نیازمند مشارکت رشته‌ها، تحلیل حساسیت و آزمون پیامدهای رفتاری است [۵].

۵-۷ دلالت‌های سیاستی

(۱) سبد شاخص‌ها: واحد ارزیابی هر مؤسسه، شناسنامه‌ای برای تعریف، منبع، سال و محدودیت هر شاخص تهیه کند و شمارش‌ها را همراه با نمونه دستاوردها گزارش دهد.

(۲) تناسب با مأموریت: کمیته‌ای با نمایندگان رشته‌ها و مدیریت

پژوهش، مأموریت، افق اثر و قواعد تفسیر را پیش از ارزیابی اعلام کند؛ مقایسه میان مؤسسات هم‌مأموریت انجام شود. (۳) کیفیت و اثر: هیئت تخصصی، مجموعه محدودی از دستاوردهای منتخب و شواهد استفاده از آن‌ها را بررسی کند؛ برای ادعای اثر اجتماعی، تأیید بهره‌بردار یا سند قابل بررسی مطالبه شود.

(۴) تمامیت و عدالت: تعارض منافع داوران ثبت، امکان اعتراض فراهم و داده‌ها از نظر پوشش رشته‌ای و سوگیری بررسی شوند؛ سرعت و حجم انتشار جانشین سلامت فرایند نباشد.

(۵) بازبینی: اجرا ابتدا در چند مؤسسه داوطلب و یک چرخه ارزیابی آزمایش شود؛ هزینه، زمان داور، اختلاف نظر و تغییر رفتار پژوهشی سنجیده و سپس درباره گسترش تصمیم گرفته شود.

۶-۷ چالش‌های اجرایی و راه مواجهه

موانع اصلی عبارت‌اند از هزینه داور، پراکندگی داده، مقاومت در برابر تغییر مشوق‌ها، تعارض منافع و دشواری انتساب اثر اجتماعی به یک پژوهش. پیشنهاد عملی، شروع با نمونه‌ای محدود از دستاوردها، استفاده از داده‌های موجود با شناسنامه مشترک، آموزش داوران و ثبت مستند استدلال آنان

پژوهش‌های بعدی می‌توانند با داده و تنظیمات قابل انتشار، خوشه‌بندی و تحول زمانی ادبیات را بازتولید کنند؛ غربالگری مستقل دو ارزیاب را به کار گیرند؛ با مطالعه طولی بودجه و مشوق‌ها، علل تفاوت تحقق شاخص‌ها را بیازمایند؛ و چارچوب شش‌بعدی را در مؤسسات دارای مأموریت متفاوت از نظر اعتبار، هزینه، عدالت و پیامد رفتاری ارزیابی کنند.

تعارض منافع

نویسندگان تعهد می‌کنند که هیچ تعارض منافی در این مقاله وجود نداشته‌است.

اظهارینامه استفاده از هوش مصنوعی

در ویرایش متن از نسخه ایرانی GapGPT-5.4 صرفاً برای بهبود نگارش و بازخوانی متون استفاده شده‌است.

References

- [1] Godin, B. (2002). **Outline for a History of Science Measurement. Science, Technology, & Human Values**, 27(1), 3–27. <https://doi.org/10.1177/01622439020700101>
- [2] Godin, B. (2007). **Science, Accounting and Statistics: The Input-Output Framework. Research Policy**, 36(9), 1388–1403. <https://doi.org/10.1016/j.respol.2007.06.002>
- [3] American Society for Cell Biology. (2013). **San Francisco Declaration on Research Assessment (DORA). American Society for Cell Biology**. <https://sfdora.org/read/>
- [4] Coalition for Advancing Research Assessment. (2022). **Agreement on Reforming Research Assessment. Coalition for Advancing Research Assessment**. <https://coara.eu/agreement/the-agreement-full-text/>
- [5] Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). **Bibliometrics: The Leiden Manifesto for Research Metrics. Nature**, 520(7548), 429–431. <https://doi.org/10.1038/520429a>
- [6] Ghazinoory, S., Ghazinoori, S., & Azadegan-Mehr, M. (2011). **Iranian Academia: Evolution After Revolution And Plagiarism as A Disorder. Science and engineering ethics**, 17(2), 213–216. <https://doi.org/10.1007/s11948-010-9206-6>
- [7] Nasr, S. H. (1968). **Science and Civilization in Islam. Harvard University Press**.
- [8] Ragep, F. J. (2001). **Freeing Astronomy from Philosophy: An Aspect of Islamic Influence on**

است. برای کاهش رفتار فرصت‌طلبانه، معیارها پیشاپیش اعلام و امتیازدهی مکانیکی به روایت اثر محدود شود. در داوری اثر، سهم پژوهش در کنار عوامل دیگر و فاصله زمانی تولید تا استفاده بررسی شود. تخصیص منابع و بازبینی سازمانی، با منطق اجرایی CoARA سازگار است [۴]؛ ولی امکان‌پذیری این پیشنهادها در ایران باید در اجرای آزمایشی سنجیده شود.

۸- نتیجه‌گیری

پاسخ پژوهش این است که تحول پیشرفت علمی از تعامل معیارهای معرفتی، نهادهای تولید دانش، ابزارهای سنجش و اهداف سیاستی حاصل شده‌است. در ایران، گسترش اهداف به نوآوری و مسئله‌محوری با تنوع متناظر معیارهای ارزیابی همراه نشده‌است. از این رو تحقق شاخص‌های شمارشی به‌تنهایی رفع «شکاف مفهوم-سنجش» را نشان نمی‌دهد.

چارچوب پیشنهادی، پایش ظرفیت و تولید را به داوری کیفیت و اثر پیوند می‌دهد و مأموریت، تمامیت و عدالت را بر هر دو حاکم می‌کند. منابع هنجاری بومی می‌توانند برای توجیه این جهت‌گیری پشتوانه فراهم کنند؛ تعیین وزن‌ها و اثبات کارآمدی، به آزمون تجربی نیاز دارد. اولویت اجرایی، آزمایش محدود و مستند چارچوب پیش از تعمیم آن به نظام ملی است. به عنوان محدودیت پژوهش، پوشش Scopus، زبان انگلیسی و محدودیت‌های موضوعی، تعمیم به کل ادبیات را محدود می‌کنند. غربالگری تک‌نفره نیز احتمال سوگیری انتخاب را افزایش می‌دهد و توافق بین ارزیابان در دسترس نیست.

گزارش فنی کامل خوشه‌بندی و تحلیل زمانی کلیدواژه‌ها در این مطالعه ارائه نشده‌است؛ بنابراین شبکه فقط نقش توصیفی مکمل دارد و مبنای انتساب آماری خوشه‌ها یا دوره‌بندی نیست. در تحلیل ایران، ناپیوستگی داده، تفاوت سال‌ها و تعریف شاخص‌ها، ارزیابی نهایی همه برنامه‌ها را ناممکن می‌سازد. مطالعه اسنادی برای اثبات سازوکارهای علی بودجه و مشوق‌ها کافی نیست.

انتخاب منابع هنجاری هدفمند و غیرجامع است و مقایسه تفسیری آن‌ها، پذیرش اجتماعی یا کارآمدی سیاست را اثبات نمی‌کند. چارچوب شش‌بعدی نیز هنوز اعتبارسنجی تجربی نشده‌است.

- Integrity.** *PLOS Biology*, 18(7), e3000737. <https://doi.org/10.1371/journal.pbio.3000737>
- [24] Dewey, J. (1938). **Logic: The Theory of Inquiry.** Henry Holt.
- [25] James, W. (1907). **Pragmatism: A New Name for Some Old Ways of Thinking.** Longmans, Green and Co.
- [26] Etzkowitz, H., & Leydesdorff, L. (2000). **The Dynamics of Innovation: From National Systems and "Mode 2" to a Triple Helix of University–Industry–Government Relations.** *Research Policy*, 29(2), 109–123. [https://doi.org/10.1016/S0048-7333\(99\)00055-4](https://doi.org/10.1016/S0048-7333(99)00055-4)
- [27] Mazzucato, M. (2018). **Mission-Oriented Innovation Policies: Challenges and Opportunities.** *Industrial and Corporate Change*, 27(5), 803–815. <https://doi.org/10.1093/icc/dty034>
- [28] Nelson, R. R. (Ed.). (1993). **National Innovation Systems: A Comparative Analysis.** Oxford University Press. <https://doi.org/10.1093/oso/9780195076165.001.0001>
- [29] Souzanchi Kashani, E. (2019). **A Historical Review of Science, Technology and Innovation Studies and the Role of the State.** *Journal of Science and Technology Policy*, 12(2), 1–16. https://jstp.nrisp.ac.ir/article_13682.html {In Persian}
- [30] Miremadi, S. I. (2019). **National Innovation System and Its Role in Improving Science, Technology and Innovation Policies.** *Journal of Science and Technology Policy*, 12(2), 135–154. https://jstp.nrisp.ac.ir/article_13691.html {In Persian}
- [31] Supreme Council of the Cultural Revolution. (2003). **Report on the Evaluation of Science and Technology in the Islamic Republic of Iran (Session 528; Registration No. 5124/DSh).** *Supreme Council of the Cultural Revolution.* {In Persian}
- [32] World Bank. (n.d.). **Research and Development Expenditure (% of GDP) – Iran, Islamic Rep.** [Data set; indicator GB.XPD.RSDV.GD.ZS]. *World Development Indicators.* <https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?locations=IR>
- [33] United Nations Conference on Trade and Development. (2016). **Science, Technology and Innovation Policy Review: Islamic Republic of Iran.** *United Nations.*
- [34] Institute for Research and Planning in Higher Education. (n.d.). **Designing an Indicator-Based Evaluation System for Research, Technology and Innovation in the Seventh Development Plan (Summary of Sixth Plan Indicator Performance, pp. 25–26).** *Ministry of Science, Research and Technology.* {In Persian}
- Science.** *Osiris*, 16, 49–71. <https://doi.org/10.1086/649338>
- [9] Strobino, R. (2018). **Ibn Sina's Logic.** In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy.* <https://plato.stanford.edu/entries/ibn-sina-logic/>
- [10] Janavi, E., Moradi, S., & Pakzad, M. (2020). **Assessment Of Iran's Scientific Publications Based On National Master Plan For Science And Education.** *Scientometrics Research Journal*, 6(1, spring & summer), 213–236.
- [11] Farazkish, M. (2016). **Designing an Evaluation Model for Science, Technology and Innovation Process of Iranian Governmental Organizations** (Doctoral dissertation, PhD dissertation). Tehran: Tarbiat Modares University.
- [12] Soofi, A. S., & Ghazinoory, S. (2011). **The Network of The Iranian Techno-Economic System.** *Technological Forecasting and Social Change*, 78(4), 591–609. <https://doi.org/10.1016/j.techfore.2010.11.005>
- [13] Carnap, R. (1937). **The Logical Syntax of Language** (A. Smeaton, Trans.). Kegan Paul, Trench, Trubner & Co.
- [14] Comte, A. (1975). **Cours de Philosophie Positive** (Vols. 1–2). *Hermann.* (Original work published 1830–1842).
- [15] Neurath, O. (1944). **Foundations of the Social Sciences.** In O. Neurath, R. Carnap, & C. Morris (Eds.), *International Encyclopedia of Unified Science* (Vol. 2, No. 1). *University of Chicago Press.*
- [16] Popper, K. (1959). **The Logic of Scientific Discovery.** *Hutchinson.*
- [17] Kuhn, T. S. (2012). **The Structure of Scientific Revolutions** (4th ed.). *University of Chicago Press.* <https://doi.org/10.7208/chicago/9780226458144.001.0001>
- [18] Lakatos, I. (1978). **The Methodology of Scientific Research Programmes.** *Cambridge University Press.* <https://doi.org/10.1017/CBO9780511621123>
- [19] Merton, R. K. (1973). **The Sociology of Science: Theoretical and Empirical Investigations.** *University of Chicago Press.*
- [20] Ziman, J. (2000). **Real Science: What It Is, and What It Means.** *Cambridge University Press.* <https://doi.org/10.1017/CBO9780511541391>
- [21] Jasanoff, S. (Ed.). (2004). **States of Knowledge: The Co-Production of Science and the Social Order.** *Routledge.* <https://doi.org/10.4324/9780203413845>
- [22] Latour, B., & Woolgar, S. (1979). **Laboratory Life: The Social Construction of Scientific Facts.** *Sage.*
- [23] Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., Coriat, A.-M., Foeger, N., & Dirnagl, U. (2020). **The Hong Kong Principles for Assessing Researchers: Fostering Research**

Research System (Higher Education and Research).

Monthly Expert Reports of the Islamic Parliament Research Center, 33(7), Report 21062.

<https://doi.org/10.22034/report.mrc.2025.1404.33.7.21062> {In Persian}

[38] Druart, T.-A. (2016). **Al-Farabi**. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

<https://plato.stanford.edu/entries/al-farabi/>

[39] Gutas, D. (2016). **Ibn Sina [Avicenna]**. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.

<https://plato.stanford.edu/entries/ibn-sina/>

[40] Irwin, R. (2018). **Ibn Khaldun: An Intellectual Biography**. Princeton University Press.

<https://doi.org/10.23943/9781400889549>

[35] Hosseinzadeh Lanjabadi, T., & Hosseini, S. J. (2022). **Evaluation of the Implementation of the Sixth Development Plan Law and the Law on Permanent Provisions of National Development Plans (Report No. 18309)**. *Islamic Parliament Research Center*. {In Persian}

[36] Student Affairs Organization. (n.d.). **Iran's Higher Education Position and Indicators of Scientific Output and H-Index in 2021**. *Ministry of Science, Research and Technology*. {In Persian}

[37] Nasiri, H., & Ghanbarpour, M. (2025). **Evaluation of the Implementation of the Seventh Progress Plan Through September 2025: Chapter 20, Improvement of the Science, Technology and**