



Elaborating and Empirical Validation of the STI Monitoring and Evaluation Framework for Iran

Mahdieh Farazkish ♦

Article Type:	Research Paper
Vol. 35 No. 4 Serial 100	Jan. 2026
Received:	2025.10.11
Revised:	2026.02.27
Accepted:	2026.04.22
Published Online:	2026.09.09
Pages:	65-86
P-ISSN:	1027-2690
E-ISSN:	2783-4514

Abstract

Effective governance of Science, Technology, and Innovation (STI) requires the design and deployment of systematic monitoring and evaluation (M&E) frameworks. Such frameworks enable continuous performance tracking, resource allocation optimization, and evidence-based policy adjustment. However, most existing studies focus on the theoretical design phase, with limited attention to the empirical validation of these frameworks in real-world contexts. This gap is particularly critical in developing countries with evolving STI governance systems, including Iran. The Supreme Council of Science, Research and Technology (ATF) of Iran, as a key high-level policymaking body, has developed three successive generations of national M&E frameworks (referred to as “periods” in this study) over the past two decades. The first period (2009-2016) focused on intervention logic and budget-based indicators, leading to the establishment of the national SAMAT system. The second period (2017-2024) introduced a functional analysis approach with 112 indicators for analyzing the National Innovation System. The third period, recently designed based on a value chain approach and a three-quadrant evaluation model (efficiency, effectiveness, impact), comprises 40 selected indicators. A distinctive feature of this third-period framework, grounded in the design science paradigm, is the integration of the “Beta Test” – a continuous design-test-redesign cycle activated before formal deployment to ensure

Keywords

Science, Technology and Innovation (STI) Monitoring and Evaluation, Evaluation Framework, Empirical Validation, Supreme Council of Science, Research and Technology (ATF), Beta Test, Data Governance, Iran.

♦ Assistant Professor of Science & Technology Policy, National Research Institute for Science Policy, Tehran, Iran (Corresponding Author)
farazkish@nrsp.ac.ir
ORCID: 0000-0003-2564-1747

Cite This Paper: Farazkish, M. (2025). Elaborating and Empirical Validation of the STI Monitoring and Evaluation Framework for Iran. *Rahyaft*, 35 (4), 65-86. (Persian).

DOI: 10.22034/RAHYAFT.2026.11926.1583



© The Author(s)
Publisher: National Research Institute for Science Policy (N.R.I.S.P)

that the framework is not only theoretical but also operational and learning-oriented. Despite its theoretical richness, the third-period framework lacks empirical analysis based on real-world data regarding its adaptability, implementability, and associated challenges. This study aims to fill this gap by empirically validating the framework using actual data from national and international sources during Iran's Sixth Five-Year Development Plan (2017-2022). The findings aim to provide a foundation for establishing a "smart national STI M&E system" in Iran and to transform ATF's structured experience into an influential model for STI policymaking at both national and regional levels.

This research is grounded in a pragmatic, mixed-methods design combining quantitative trend analysis with qualitative expert judgment, framed within the design-science research paradigm. This methodology is uniquely suited for creating and evaluating "artifacts" (here, the M&E framework) intended to solve complex practical problems. The validation involved an indicator-by-indicator analysis of real-world performance from 2017 to 2022. Data for all 40 indicators were systematically extracted from authoritative national sources (including the Plan and Budget Organization, Ministry of ATF, and the Statistical Center of Iran) and international databases (including UNESCO, WIPO, World Bank, OECD). For each indicator, the Compound Annual Growth Rate (CAGR) was calculated using the formula: $CAGR = ((\text{End Value} / \text{Start Value})^{(1/\text{Number of Years})}) - 1$ (CFI, 2023). A dual-criteria evaluation matrix was then applied: (1) comparison of calculated CAGR against quantitative targets in the national development plan; and (2) alignment with international standards (e.g., Frascati Manual, Oslo Manual) using a compliance score and the DAB index (Accessibility, Analyzability, and Being up-to-date). A panel of five senior experts – including STI policy academics, former ATF officials, and national STI reporting practitioners – was convened. Through semi-structured discussions, the panel examined quantitative results to diagnose root causes of poor performance, identify field-level data governance challenges, and assess the perceived relevance and feasibility of proposed indicators from an end-user perspective. Content validity was ensured by

directly adopting the indicator framework from the original design report. Reliability was addressed through transparent documentation of all calculation steps and structured, documented expert elicitation processes. The empirical validation revealed a dualistic outcome with several key findings:

- ◆ **Theoretical Strength:** The average alignment of indicators with international standards (OECD, Frascati, and Oslo) was 85%. This alignment was strategically distributed: 97% for the efficiency quadrant, 67% for the effectiveness quadrant, and 92% for the impact (utility) quadrant. This distribution indicates a deliberate design focus on measuring final socio-economic outcomes.
- ◆ **Empirical Performance Deficiency:** The actual performance of Iran's STI system, as measured by the framework, was predominantly "weak." The weighted average CAGR across all 34 indicators (out of 40 with sufficient data) was 1.7%. Performance across quadrants showed: efficiency (CAGR: 1.5%), effectiveness (CAGR: 1.8%), and impact (CAGR: 1.7%), all rated as "weak." Out of 34 indicators with available data, 21 fell into the "weak" or "very weak" categories. Only one indicator (number of knowledge-based firms relative to targets) was rated "excellent," and two were rated "good."
- ◆ **Critical Implementation Challenges:** Major operational barriers were identified, centered on data governance deficiencies. Key issues included low DAB scores (particularly in the impact quadrant), fragmentation of information systems, and lack of institutional coordination in data collection. Consequently, 6 out of 40 proposed indicators could not be analyzed due to the complete absence of reliable time-series data.
- ◆ **Key Innovations:** The integration of the "Beta Test" (design-test-redesign cycle) was highlighted as a crucial mechanism for iterative learning and adaptation. The strategic reduction of indicators from 112 (second period) to 40 (third period) was seen as a pragmatic response to data capacity constraints and a move toward a more focused, impact-oriented M&E system.

This study concludes that Iran's third-period STI M&E framework is a well-designed, internationally-aligned "design artifact" with significant potential to serve as a model for evidence-based STI governance. The average 85% alignment with international standards, particularly the 92% alignment in the impact quadrant, demonstrates theoretical robustness comparable to third-generation frameworks in countries like Turkey (TÜBİTAK) and South Korea (NRF). The strategic reduction from 112 to 40 indicators and the innovative integration of the Beta Test represent important advances.

However, successful implementation and policy impact are highly conditional. The actual performance of Iran's STI system remains "weak" (1.7% weighted average CAGR), with only marginal improvements in effectiveness and impact quadrants compared to efficiency. The most critical bottleneck is not framework design but national data governance infrastructure. Low DAB scores, fragmented information systems, and a lack of institutional coordination severely undermine the framework's operational potential.

Therefore, three interdependent recommendations are proposed for short-term implementation:

- ◆ Operationalize the Beta Test as a genuine learning and adaptation cycle, focusing on overcoming data accessibility barriers in the effectiveness and impact quadrants.
- ◆ Develop and mandate a "National STI Data Governance Policy" establishing unified standards, legal obligations, and transparent data-sharing protocols across all relevant institutions (including the Statistical Center, Central Bank, specialized ministries, and universities).
- ◆ Create explicit links between M&E results and resource allocation mechanisms, even experimentally during the pilot phase, to demonstrate practical value for stakeholders and incentivize data quality improvement.

Without these parallel institutional and process-oriented reforms, the framework risks remaining a sophisticated but underutilized document, unable to break the cycle of "good design, weak implementation, and negligible impact" on Iran's innovation system. The modest but encouraging growth in effectiveness (1.8%) and impact (1.7%)

relative to efficiency (1.5%) suggests the system has capacity for longer-term outcomes, but this trajectory requires decisive acceleration through data infrastructure and institutional strengthening.



توسعه و اعتبارسنجی تجربی چهارچوب پایش و ارزیابی علم، فناوری و نوآوری ایران

مهدیه فرازکیش

• نوع مقاله: پژوهشی

• دوره ۳۵ | شماره ۴ | پیاپی ۱۰۰ | دی ۱۴۰۴

• تاریخ دریافت: ۱۴۰۴/۰۷/۱۹

• تاریخ بازنگری: ۱۴۰۴/۱۲/۰۸

• تاریخ پذیرش: ۱۴۰۵/۰۲/۰۲

• تاریخ انتشار برخط: ۱۴۰۵/۰۵/۱۸

• صفحات: ۶۷-۸۸

• شاپای چاپی: ۱۰۲۷-۲۶۹۰

• شاپای الکترونیکی: ۲۷۸۳-۴۵۱۴

چکیده

این پژوهش با هدف اعتبارسنجی تجربی چهارچوب دوره سوم پایش و ارزیابی علم، فناوری و نوآوری (STI) شورای عالی علوم، تحقیقات و فناوری (عتف) انجام شده است. چهارچوب یادشده مبتنی بر مدل یکپارچه ارزیابی (چارک‌های کارایی، اثربخشی، اثرگذاری) و شامل ۴۰ شاخص منتخب است. هدف اصلی، سنجش میزان انطباق، قابلیت اجرا و ارزش افزوده این چهارچوب پیش از استقرار رسمی آن در نظام حکمرانی علم و فناوری ایران بوده است.

پژوهش حاضر با رویکرد ترکیبی (تحلیل کمی روندهای زمانی و تحلیل کیفی مبتنی بر نظر خبرگی) انجام شد. داده‌های واقعی ۴۰ شاخص از گزارش‌های ملی و بین‌المللی در بازه زمانی برنامه ششم پیشرفت کشور (۱۳۹۶-۱۴۰۱) گردآوری و تحلیل شده و روایی و پایایی ابزارها تأیید گردیده است.

یافته‌های پژوهش نشان می‌دهد میانگین انطباق شاخص‌ها با استانداردهای بین‌المللی ۸۵ درصد است. این انطباق در چارک‌های کارایی ۹۷ درصد، اثربخشی ۶۷ درصد و اثرگذاری ۹۲ درصد بوده است. مضاف بر اینکه عملکرد کلی پژوهش و فناوری کشور براساس چهارچوب پیشنهادی در سطح «ضعیف» با میانگین وزنی نرخ رشد سالانه مرکب (CAGR) معادل ۱/۷ درصد ارزیابی شدند و چارک‌های کارایی، اثربخشی و اثرگذاری همگی در سطح «ضعیف» قرار گرفتند. از ۳۴ شاخص دارای داده، ۲۱ شاخص در سطوح «ضعیف» یا «بسیار ضعیف» بودند. همچنین براساس یافته‌های پژوهش، مهم‌ترین چالش‌های اجرایی، ضعف در دسترسی، تحلیل‌پذیری و به‌روز بودن داده‌ها (به‌ویژه در حوزه اثرگذاری)، پراکندگی سامانه‌های اطلاعاتی و عدم یکپارچگی در گردآوری داده‌ها شناسایی شده‌اند. با وجود این، ارزش افزوده چهارچوب ارزیابی متأخر در مقایسه با موارد پیشین، در یکپارچگی زنجیره ارزش، کاهش شاخص‌ها از ۱۱۲

کلیدواژه‌ها

پایش و ارزیابی علم، فناوری و نوآوری، چهارچوب ارزیابی، اعتبارسنجی تجربی، شورای عالی عتف، آزمون بتا.

♦ استادیار سیاست‌گذاری علم و فناوری، مؤسسه تحقیقات سیاست علمی کشور، تهران، ایران (پدیده‌آور رابط)

farazkish@nrisp.ac.ir

ORCID: 0000-0003-2564-1747

استناد به این مقاله: فرازکیش، م. (۱۴۰۴). توسعه و اعتبارسنجی تجربی چهارچوب پایش و ارزیابی علم، فناوری و نوآوری ایران. *راهیافت*، ۳۵ (۴)، صص. ۶۷-۸۸.

DOI: 10.22034/RAHYAFT.2026.11926.1583

ناشر: مؤسسه تحقیقات سیاست علمی کشور
نویسندگان: © حق مؤلف



به ۴۰ مورد و ادغام «آزمون بتا» برای یادگیری سیاستی پیش از اجرا متجلی است.

نتیجه‌گیری کلی حاکی از آن است که چهارچوب یادشده از نظر ساختاری و نظری آمادۀ استقرار است، اما مستلزم اصلاحاتی در شاخص‌های اثرگذاری و تقویت زیرساخت داده‌ای است. اجرای آزمایشی این چهارچوب با پایش مستمر و الگوبرداری از تجربیات موفق کشوری و بین‌المللی در افق کوتاه‌مدت پیشنهاد می‌شود. این چهارچوب می‌تواند به الگویی مؤثر برای نظام پایش ملی هوشمند STI در ایران و کشورهای منطقه تبدیل شود.

مقدمه

حکمرانی مؤثر در حوزه علم، فناوری و نوآوری (STI)، مستلزم طراحی و استقرار چهارچوب‌های نظام‌مند پایش و ارزیابی است. این چهارچوب‌ها با فراهم آوردن امکان رصد مستمر عملکرد و اثربخشی برنامه‌های کلان، زمینه‌ساز تخصیص بهینه منابع و تنظیم سیاست‌های هوشمندانه می‌شوند (Ghazinoory, Farazkish, Montazer, & Soltani, 2017; Harris & Albury, 2009). بااین‌حال، عمده مطالعات موجود بر مرحله طراحی نظری این‌گونه چهارچوب‌ها متمرکز بوده و کمتر به تحلیل تجربی پیاده‌سازی و اعتبارسنجی عملی آن‌ها در بسترهای واقعی پرداخته‌اند (Rodriguez & Soeparwata, 2015). این شکاف پژوهشی، به‌ویژه در کشورهای در حال توسعه، با نظام‌های حکمرانی STI در حال تکامل (kosari & alizadeh, 2021; Bagheri Moghaddam & Mosleh, 2016)، ضرورت بررسی تجربی استقرار چنین چهارچوب‌هایی را دوچندان می‌کند.

در ایران، نهادهای حکمرانی متعددی در موضوع پایش و ارزیابی STI مشارکت داشته‌اند (Ghazinoory & Farazkish, 2018). از جمله، شورای عالی انقلاب فرهنگی، طی چند دوره ارزیابی‌های کلان علم و فناوری (سال‌های ۱۳۸۱، ۱۳۸۴)، ارزیابی آموزش عالی با ۴۸ شاخص (سال ۱۳۸۳)، سند تحول آموزش عالی با ۳۵ معیار (سال ۱۳۹۵)، و ارزیابی شاخص‌های ۶۴ گانه نقشه جامع علمی کشور (به‌صورت سالیانه) را در دستور کار داشته است. همچنین، مرکز بررسی‌های استراتژیک ریاست جمهوری گزارش‌های کلان علم و فناوری در سال ۱۴۰۰ و معاونت علمی، فناوری و اقتصاد دانش‌بنیان ریاست جمهوری گزارش‌های موردی ارزیابی فناوری و نوآوری را از سال ۱۳۹۲ در کارنامه خود داشته‌اند (Montazer, Sharanji, Moradipor, & Farazkish, 2019)، اما از میان این تجربیات پایش و ارزیابی ملی، تجربه شورای عتف منحصر به فرد است.

شورای عالی علوم، تحقیقات و فناوری (عتف) به‌عنوان یکی از مهم‌ترین نهادهای سیاست‌گذار سطح کلان کشور در این حوزه، که سه دوره چهارچوب پایش و ارزیابی را طی قریب به دو دهه

گذشته تجربه کرده است، تجربه زیسته ارزشمندی دارد. دوره اول طی سال‌های ۱۳۸۸-۱۳۹۵ با تمرکز بر منطق مداخله و به‌ویژه شاخص‌های بودجه‌محور، به راه‌اندازی سامانه سمات ملی برای ارزیابی سالیانه هزینه‌کرد دستگاه‌های اجرایی در پژوهش و فناوری منجر شد (Ghazinoory, Divsalar, & Ghazinoory, 2012; Abbasi & Ashrafi, 2011)؛ دوره دوم طی سال‌های ۱۳۹۶ تاکنون با تصویب آیین‌نامه پایش و ارزیابی شورای عالی عتف (Supreme Council for Science, Research and Technology, 2021) و متشکل از ۱۱۲ شاخص به‌منظور تحلیل کارکردی نظام ملی نوآوری مستقر شده است (Khayyatian, Fartash, & Pourasgari, 2020) و دوره سوم که به‌تازگی مبتنی بر ارزیابی چارک‌های ارزیابی تناسب، کارایی، اثربخشی، اثرگذاری و شامل ۴۰ شاخص منتخب در امتداد زنجیره ارزش STI طراحی شده است (Aghaei, 2021). ویژگی خاص چهارچوب دوره سوم مبتنی بر پارادایم علم طراحی، ادغام «آزمون بتا»^۱ در چرخه طراحی است که چرخه مستمر «طراحی/آزمون/بازخورد/بازطراحی» را پیش از استقرار رسمی فعال می‌سازد و تضمین می‌کند چهارچوب نه‌تنها نظری، بلکه عملیاتی و یادگیرنده باشد (Hevner, March, Park, & Ram, 2004; Van Aken, 2004; Farazkish, 2016).

با توجه به اجرا نشدن این آزمون، به‌رغم غنای نظری چهارچوب دوره سوم، فقدان تحلیل تجربی مبتنی بر داده‌های واقعی درباره انطباق‌پذیری، قابلیت اجرا و چالش‌های پیش روی آن، یک خلأ پژوهشی مشهود است. این مقاله در پی پر کردن این خلأ از طریق اعتبارسنجی تجربی چهارچوب یادشده، با استفاده از داده‌های واقعی مستخرج از اسناد ملی و بین‌المللی و آمار نهادهای مربوط در داخل کشور، در بازه زمانی برنامه ششم پیشرفت (۱۳۹۶-۱۴۰۱) است. یافته‌های این پژوهش می‌تواند مبنایی برای استقرار «نظام هوشمند پایش و ارزیابی STI»^۲ در ایران فراهم آورد و تجربه ساختاریافته عتف را به الگویی اثرگذار در سیاست‌گذاری علم و فناوری در سطح ملی و

۱. در روش‌شناسی علم طراحی، آزمون بتا فرایندی است که در آن کاربران نهایی در محیط واقعی و عملی، یک اثر طراحی‌شده (artifact) را با هدف افزایش عینیت یافته‌های تحقیق، آزمایش قابلیت انتقال روش و اعتبارسنجی مفهوم اثر پیش از کاربری، ارزیابی می‌کنند. به عبارت دیگر، درحالی‌که آزمون آلفا به‌صورت داخلی و در محیط آزمایشگاهی انجام می‌شود، آزمون بتا، شکلی از آزمون پذیرش خارجی است که بازخورد ارزشمندی برای بهبود نهایی اثر فراهم می‌کند.

۲. در این مقاله، منظور از «نظام هوشمند ملی پایش و ارزیابی STI» نظامی است که: (۱) داده‌های پراکنده از نهادهای مختلف (مرکز آمار، وزارت عتف، معاونت علمی، بانک مرکزی و ...) را یکپارچه می‌کند؛ (۲) شاخص‌های کلیدی (۴۰ شاخص دوره سوم) را به‌صورت منظم و سالانه اندازه‌گیری و رصد می‌کند (عملکرد پایشی)؛ (۳) یافته‌ها را تحلیل و تفسیر کرده، به توصیه‌های سیاستی تبدیل می‌کند (عملکرد ارزیابی)؛ (۴) دارای حلقه بازخورد یادگیرنده است؛ نتایج ارزیابی به بازطراحی شاخص‌ها و اصلاح برنامه‌ها و سیاست‌های آتی منجر می‌شود (همان‌طور که در «آزمون بتا» پیش‌بینی شده است)؛ و (۵) در بلندمدت، با بهره‌گیری از داده‌های تاریخی و ابزارهای تحلیلی پیشرفته، به‌سمت پیش‌بینی روندها و هشدار زودهنگام حرکت می‌کند (افق آینده).

منطقه‌ای تبدیل کند.

با توجه به خلأ پژوهشی یادشده و هدف اصلی این مطالعه،

پرسش‌های پژوهش به شرح زیر تدوین شده‌اند:

۱. درجه انطباق ۴۰ شاخص دوره سوم چهارچوب پایش و ارزیابی STI با داده‌های موجود در اسناد ملی و بین‌المللی تا چه اندازه است؟
۲. عملکرد هریک از چارک‌های ارزیابی منتخب (کارایی، اثربخشی و اثرگذاری) بر اساس داده‌های واقعی چگونه است؟
۳. مهم‌ترین چالش‌های اجرایی در زمینه دسترسی به داده‌ها، محاسبه‌پذیری شاخص‌ها و اجرای عملی چهارچوب دوره سوم کدام‌اند؟
۴. ارزش افزوده چهارچوب دوره سوم در مقایسه با دوره‌های پیشین (فرایندمحور، کارکردی و سیاستی) در چه ابعادی نمود می‌یابد؟
۵. راهکارهای بهبود انطباق‌پذیری، استانداردسازی و استقرار موفق نظام پایش ملی STI مبتنی بر یافته‌های پژوهش چیست؟

پس از این مقدمه، در ادامه مقاله در پنج بخش اصلی سازمان‌دهی شده است: بخش دوم، پیشینه نظری و تجربی، شامل تعیین محدوده پژوهش، مرور چهارچوب‌های ملی و بین‌المللی و تبیین سیر تکاملی سه‌دوره‌ای عتف را دربردارد. سپس در بخش سوم، روش‌شناسی پژوهش، شامل تشریح جامعه آماری، مراحل گردآوری و تحلیل داده‌های واقعی و ابزارهای ارزیابی، ارائه و در بخش چهارم، یافته‌های پژوهش، شامل نتایج تحلیل روند، وضعیت عملکرد و بررسی انطباق، تبیین شده است. در بخش پنجم، بحث و تفسیر یافته‌ها شامل پاسخ به پرسش‌های پژوهش، مقایسه با مطالعات پیشین و تدقیق چالش‌های اجرایی و سیاستی، مدنظر قرار گرفته و در بخش پایانی، نتیجه‌گیری و پیشنهادهای پژوهش با تأکید بر رفع گلوگاه‌های شناسایی شده، ارائه شده است.

پیشینه نظری و تجربیات ملی و بین‌المللی

تعاریف مفهومی و تمایزات محوری

در مطالعات مربوط به حوزه علم، فناوری و نوآوری (STI)، چهارچوب^۱ به مجموعه‌ای ساختاریافته از شاخص‌ها، معیارها و سازوکارهای عملیاتی اطلاق می‌شود که جهت‌گیری نظام‌مند پایش و ارزیابی عملکرد نظام STI را فراهم می‌آورد. در مقابل، مدل^۲ عمدتاً به‌مثابه ابزاری نظری و انتزاعی برای تبیین روابط علی میان متغیرها به کار می‌رود (Partelow, 2023; OECD, 2015; Markiewicz & Patrick, 2015; Contandriopoulos & Brousselle, 2012; Van Aken, 2004). بنابراین، تمایز اصلی چهارچوب از مدل در عملیاتی

بودن نهفته است؛ چهارچوب‌ها قابلیت اجرا، پایش و اصلاح مستمر دارند، اما مدل‌ها عمدتاً توصیفی و پیش‌بینی‌کننده هستند.

پایش^۳ فرایند گردآوری منظم داده‌ها، رصد روندها و هشدار درباره انحرافات از اهداف سیاستی است، درحالی که ارزیابی^۴ شامل تحلیل عمیق اثربخشی، قضاوت ارزشی و ارائه توصیه‌های اصلاحی می‌شود (Eurostat, 2018; Patton, 2017; OECD, 2015; Rodriguez & Soeparwata, 2015; Markiewicz & Patrick, 2015). بدین ترتیب، تمایز پایش داده‌محوری و منظم بودن است، درحالی که ارزیابی دورهای و تحلیل‌محور است. همچنین از مطالعات بالا به نظر می‌رسد نظام یکپارچه پایش و ارزیابی، ترکیبی هوشمندانه از این دو رویکرد است که در گام نخست داده‌های خام را به‌صورت منظم گردآوری می‌کند و در گام بعد به تحلیل کمی و کیفی دورهای آن‌ها می‌پردازد.

بر این اساس، تمرکز این مقاله بر «چهارچوب پایش و ارزیابی» است؛ چنانچه علاوه بر اصطلاح رسمی مصوبه هیئت وزیران در خصوص «آیین‌نامه پایش و ارزیابی علم، فناوری و نوآوری» (مصوب ۱۳۹۵/۱۱/۱۲)، عمدتاً با رویکرد پایش سالیانه گزارش‌های شورا تدوین و ارائه شده است، اما در هر سه دوره، چهارچوب طراحی شده ناظر به ابعاد عملکردی نیز بوده است که این موضوع در بخش بعدی تدقیق خواهد شد. همچنین جایگاه کاربردی مصنوع طراحی شده مبتنی بر شاخص‌های ارزیابی نیز مفهوم «چهارچوب» را بر ذهن متبادر می‌سازد.

تجربیات بین‌المللی و ملی منتخب

در سطح بین‌المللی، دستنامه‌هایی همچون فراسکاتی (ویراست هفتم، ۲۰۱۵) و اسلو (ویراست پنجم، ۲۰۱۸) به‌ترتیب بر سنجش نهادهای پژوهش و توسعه (مانند بودجه و نیروی انسانی) و بروندهای نوآوری متمرکز بوده و مبانی اقتصاددانش‌بنیان را شکل داده‌اند. تکامل این‌گونه چهارچوب‌ها در سال‌های اخیر به‌سمت ادغام ابعاد محیط‌زیستی و اجتماعی اثرگذاری حرکت کرده است (Pinar & Horne, 2022; OECD, 2021; Orozco, Ordóñez-Matamoros, García-Estévez, Sierra-González, & Bortagaray, 2021; UNEP, 2020).

در سطح ملی، کشورهای مختلف الگوهای متنوعی را در پیش گرفته‌اند که تحلیل تطبیقی این تجربه‌های زیسته نشان می‌دهد. در واقع، مسیر تکاملی چهارچوب‌های پایش و ارزیابی این کشورها، از الگویی نظام‌مند و تکاملی تبعیت می‌کند که حاکی از گذار تدریجی از رویکردهای ساده و توصیفی به چهارچوب‌های پیچیده، یکپارچه و

3. Monitoring

4. Evaluation

1. Framework

2. Model

سیاست‌محور است. به‌طور مشخص، چهار نهاد برجست جهانی در این حوزه عبارت‌اند از

♦ مرکز ملی پژوهش‌های علمی فرانسه^۱ (CNRS): این نهاد نیز از یک گذار دودسته‌ای از دوره اول (دهه ۱۹۹۰، مبتنی بر شاخص‌های پژوهشی) به دوره دوم (دهه ۲۰۲۰، مبتنی بر تحلیل اثرگذاری اجتماعی و بین‌رشته‌ای) عبور کرده است (Vetterli et al., 2023; Acciai, 2023).

♦ سازمان پژوهش‌های علمی و فناوری ترکیه (TÜBİTAK): این سازمان مسیری سه‌دوره‌ای را طی کرده است: دوره اول (حوالی ۲۰۰۵) با تمرکز بر شاخص‌های کمی پژوهشی، دوره دوم (حدود ۲۰۱۵) با حرکت به سمت ارزیابی کارکردی و اثربخشی برنامه‌ها، و دوره سوم (پیش‌بینی شده برای ۲۰۲۵ به بعد) که بر پایه فرایندهای نوآوری و زنجیره ارزش طراحی شده است (Çubuk, 2023; Gemici & Gemici, 2021).

♦ بنیاد پژوهش ملی کره جنوبی^۲ (NRF): این نهاد در کره جنوبی نیز تحول سه‌دوره‌ای قابل توجهی را از دهه ۱۹۹۰ تجربه کرده است. دوره اول (دهه ۱۹۹۰) با تمرکز بر شاخص‌های کمی پژوهش، دوره دوم (دهه ۲۰۰۰) با تأکید بر ارزیابی کارکردی و همکاری دانشگاه - صنعت، و دوره سوم (از ۲۰۱۰ به بعد) با رویکرد فرایندمحور و سامانه‌های پایش یکپارچه توسعه یافته است (Cho, Ahn, & Park, 2024).

بنیاد ملی علوم آمریکا (NSF): این بنیاد پیشرو، مسیر تکامل چهاردوره‌ای را طی کرده که الگویی جامع‌تر ارائه می‌دهد: دوره توصیفی (دهه ۱۹۷۰)، دوره تحلیلی (دهه ۱۹۹۰)، دوره سیاست‌محور (حدود ۲۰۱۰) و دوره داده‌محور و آینده‌نگر (پیش‌بینی شده برای ۲۰۲۵ به بعد) (Falkenheim & Alexander, 2023).

سیر تکاملی چهارچوب‌های پایش و ارزیابی در شورای عالی عتف

شورای عالی عتف از زمان تأسیس رسمی خود در سال ۱۳۸۳ براساس قانون اهداف، وظایف و تشکیلات وزارت علوم، تحقیقات و فناوری، نقش نهاد فرابخشی سیاست‌گذار و ناظر کلان در توسعه ملی علم و فناوری کشور را داشته است (Supreme Council of Science, Research & Technology, 2021). این شورا با استناد به مصوبات قانونی مجلس شورای اسلامی، مأموریت دارد شاخص‌های ملی علم، فناوری و نوآوری را برای ارزیابی عملکرد دستگاه‌های اجرایی و تدوین گزارش‌های ملی عملکرد تدوین و پایش کند. در مسیر ۱۵ سال فعالیت شورا در این حوزه، سه دوره عمده تحول در چهارچوب‌های

پایش و ارزیابی علم و فناوری کشور شکل گرفته است:

الف) دوره نخست: رویکرد تحلیل منطق مداخله در ارزیابی عملکرد تخصیص اعتبارات پژوهش و فناوری (۱۳۸۸-۱۳۹۵) نخستین چهارچوب ملی ارزیابی علم و فناوری در قالب طرح کلان ملی «تدوین چهارچوب ارزیابی عملکرد تخصیص اعتبارات پژوهش و فناوری» بر پایه منطق مداخله طراحی شد (Ghazinoory et al., 2012). این چهارچوب تأکید داشت که داده‌های بودجه‌ای، برون‌دادی و پیامدی نظام پژوهش و فناوری کشور از طریق سامانه‌های متمرکز (سمات ملی) گردآوری و تحلیل شود. اگرچه این دوره در تکوین فرهنگ ارزیابی ملی نقش مؤثری داشت، اما به دلیل نبود سازوکارهای انگیزشی و زیرساختی و ضعف در حلقه بازخورد یادگیرنده، کاربست سیاستی محدودی پیدا کرد (Farazkish, 2016).

ب) دوره دوم: رویکرد تحلیل کارکردهای نظام ملی نوآوری در ارزیابی و پایش علم، فناوری و نوآوری (۱۳۹۶-۱۴۰۳) در دومین مرحله تحول، شورای عالی عتف با تصویب «آیین‌نامه اجرایی نظام پایش و ارزیابی علم، فناوری و نوآوری کشور» در سال ۱۳۹۵، رویکرد خود را به «تحلیل کارکردی نظام ملی نوآوری» تغییر داد. در این چهارچوب، ۱۱۲ شاخص در ۱۱ معیار اصلی تدوین شد و دستگاه‌های اجرایی مکلف شدند سالیانه داده‌های مرتبط را به دبیرخانه شورا ارائه کنند (Khayyatian et al., 2020). این اقدام، نخستین تلاش ساختاریافته برای مستندسازی گزارش‌های ملی بر مبنای شاخص‌های متوازن بود. باوجود این، مشکلاتی همچون تأخیر در تهیه داده‌ها، ضعف در خوشه‌بندی مقایسه‌ای و عدم انطباق برخی شاخص‌ها با استانداردهای جهانی باعث شد نیاز به بازنگری در مدل ارزیابی نمایان شود (Aghaie, 2021).

ج) دوره سوم: رویکرد تحلیل زنجیره ارزش در ارزیابی و پایش علم، فناوری و نوآوری

دبیرخانه شورای عالی عتف با بهره‌گیری از یافته‌های آسیب‌شناسی مدل‌های گذشته، از سال ۱۴۰۰ بازطراحی نظام پایش و ارزیابی را با تکیه بر «رویکرد زنجیره ارزش فرایندی علم، فناوری و نوآوری» آغاز کرد. دوره سوم نه به مثابه تکاملی تدریجی، بلکه همچون چهارچوبی مستقل و جدید طراحی شد. این دوره از پارادایم علم طراحی (Van Aken, 2004) بهره برده و به جای پیوستگی ساختاری با گذشته، براساس مدل یکپارچه چهارچارک (با تمرکز بر سه چارک کارایی، اثربخشی، اثرگذاری) (Ghazinoory et al., 2023) و انتخاب هدفمند ۴۰ شاخص کلیدی بنا نهاده شد. این رویکرد با هدف تقویت حلقه پیوستگی میان تولید دانش، توسعه فناوری، بهره‌وری اقتصادی و ارزش اجتماعی طراحی شد و اکنون در مرحله پیش‌آزمایی (آزمون بتا)

1. French National Centre for Scientific Research (Centre national de la recherche scientifique)
2. National Research Foundation of Korea

و ارزش افزایی مستمر در زنجیره علم، فناوری و نوآوری صورت پذیرد (Aghaie, 2021).
در جدول ۱ مقایسه و سیر تحول سه دوره چهارچوب‌های پایش و ارزیابی STI شورای عالی عتف ارائه شده است.

قرار دارد تا در صورت موفقیت، به نظام رسمی ارزیابی کشور تبدیل شود. به نظر می‌رسد هدف این دوره جدید از چهارچوب‌های ارزیابی، ایجاد بازخوردهای یادگیرنده میان سیاست‌گذاری، داده‌کاوی و بهبود مستمر شاخص‌هاست تا تصمیم‌سازی در سطح کلان مبتنی بر شواهد

جدول ۱. مقایسه تحولات سه دوره‌ای چهارچوب‌های پایش و ارزیابی شورای عالی عتف

دوره	رویکرد غالب	تعداد شاخص	حوزه تمرکز ارزیابی	سازوکار اجرایی
نخست	منطق مداخله در تخصیص اعتبارات	۵۷	درونداد - برونداد - پیامد	سامانه سمات
دوم	کارکردهای نظام ملی نوآوری	۱۱۲	کارکرد	گزارش‌دهی سالیانه
سوم	زنجیره ارزش STI و ذی‌نفعان کلیدی	۴۰	چارک‌های ارزیابی کارایی - اثربخشی - سودمندی	آزمون بتا (چرخه طراحی - آزمون - بازطراحی)

و افزایش احتمال موفقیت چهارچوب در استقرار رسمی فراهم آورد (Peffer, Tuunanen, Rothenberger, & Chatterjee, 2007). این رویکرد، پژوهش حاضر را از مطالعات صرفاً توصیفی یا ارزیابی‌های پسینی متمایز می‌سازد.

بدین منظور، این پژوهش باهدف اعتبارسنجی تجربی چهارچوب دوره سوم، از روش ترکیبی ساختاریافته شامل تحلیل کمی روندهای زمانی^۴ و ترکیب کیفی خبرگی^۵ استفاده کرده است. چنانچه اشاره شد، منطق روش‌شناختی، دنبال کردن چرخه کاملی از گردآوری داده‌های واقعی، تحلیل روند، ارزیابی عملکرد و تجویز سیاستی است که امکان سنجش عینی و نیز تفسیر عمیق چالش‌ها و فرصت‌ها را فراهم می‌آورد.

جامعه و منابع آماری پژوهش

جامعه آماری این پژوهش، همه گزارش‌ها و داده‌های رسمی مرتبط با ۴۰ شاخص چهارچوب دوره سوم (پیوست ۱) در بازه زمانی برنامه ششم پیشرفت کشور (۱۳۹۶ تا ۱۴۰۰) است. داده‌های پژوهش از منابع ثانویه، شامل گزارش‌های ملی و بین‌المللی مرتبط با شاخص‌های موضوعه و اسناد و سوابق سازمان‌های متولی انتشار آمار و ارقام ملی هریک از شاخص‌های چهارچوب استخراج شده‌اند. در برخی موارد، اطلاعات میدانی پژوهش شامل مصاحبه‌های موضوعی در خصوص شاخص‌های پیشنهادی کمیسیون‌های تخصصی شورای عالی عتف و نیز برگزاری جلسه‌های هم‌فکری و اعتبارسنجی داده‌های حاصل در کارگروه پایش و ارزیابی شورای عالی عتف انجام شده است. در نهایت نتایج اعتبارسنجی و تحلیل روند شاخص‌ها نیز در همین کارگروه مشورتی مطرح گردیده و جمع‌بندی سیاستی براساس نظرات نمایندگان دستگاهی در این کارگروه تدوین شده است. از تلفیق این

بر این اساس، مسیر از دوره اول و دوم به دوره سوم، یک «پرش کیفی» مبتنی بر یادگیری از تجربیات گذشته محسوب می‌شود، نه صرفاً تکامل خطی. این چهارچوب جدید با هدف ارائه نظامی هوشمند، یکپارچه و کاملاً عملیاتی برای پایش ملی STI طراحی شده است که در ادامه امکان‌پذیری، قابلیت اجرایی و ارزش افزوده سیاستی آن بررسی و تحلیل خواهد شد.

روش پژوهش

این پژوهش در چهارچوب روش‌شناسی علم طراحی^۱ (Hevner et al., 2004; Van Aken, 2004) و با راهبرد پژوهشی اعتبارسنجی تجربی انجام شده است. هسته مرکزی این رویکرد، خلق و ارزیابی «مصنوعات»^۲ — در اینجا، چهارچوب دوره سوم پایش و ارزیابی STI — برای حل مسائل عملی پیچیده است. در این روش‌شناسی، پژوهش نه صرفاً برای توصیف یا تبیین، بلکه برای طراحی و اعتبارسنجی یک راه‌حل در بستر واقعی انجام می‌شود (Gregor & Hevner, 2013). راهبرد اصلی این مطالعه، کاربست و آزمون چرخه تکرارشونده «طراحی - ساخت - ارزیابی» است که در قالب «آزمون بتا»^۳ متجلی می‌شود. آزمون بتا به‌عنوان مرحله نهایی اعتبارسنجی پیش از استقرار رسمی، در این پژوهش به حوزه حکمرانی علم و فناوری تعمیم یافته و به‌عنوان سازوکاری برای ارزیابی کاربردپذیری، قابلیت اطمینان و اثربخشی بالقوه چهارچوب در یک محیط نیمه‌واقعی با استفاده از داده‌های تاریخی به کار رفته است. بنابراین، روش‌شناسی ترکیبی کمی — کیفی به‌کاررفته (تحلیل روند و خبرگی)، نه به‌مثابه هدف، بلکه ابزاری در خدمت این چرخه یادگیری طراحی‌محور است تا از طریق ارائه بازخورد پیشینی، زمینه را برای بازطراحی هوشمندانه

1. Design Science Research Paradigm
2. Artifacts
3. Beta Testing

4. Quantitative Trend Analysis
5. Expert Synthesis

منابع، در نهایت ۲۰۰ سری زمانی کامل (۴۰ شاخص × ۵ سال) برای تحلیل کمی آماده شد.

در خصوص تعیین استاندارد شاخص‌های مطالعه‌شده نیز علاوه بر تحلیل کیفی مقایسه‌ای چهارچوب‌های ارزیابی شورای عالی عتف، از استانداردهای تدوین شناسنامه شاخص ملی (مرکز آمار ایران) و بین‌المللی (یونسکو) نیز استفاده شده است تا معیارهای اعتبارسنجی هر شاخص تعیین شود. جمع‌بندی نهایی این بخش نیز براساس کارگروه مشورتی پایش و ارزیابی شورای عالی عتف انجام گرفته است.

مراحل اجرای پژوهش

فرایند پژوهش در چهار مرحله متوالی و مکمل طراحی و از مرداد ۱۴۰۲ تا اردیبهشت ۱۴۰۳ اجرا شد:

و در یک ماتریس ۵×۵ (از بسیار ضعیف تا عالی) سطح‌بندی شدند:

◇ مقایسه روند محاسبه‌شده (CAGR) و وضعیت نهایی شاخص با اهداف کمی برنامه ششم پیشرفت؛

◇ مقایسه با استانداردهای بین‌المللی مرتبط در قالب نمره انطباق^۱ (مثلاً با میانگین کشورهای مطالعه‌شده در گزارش‌های OECD) و نمره ترکیبی DAB^۲ (OECD, 2021).

۳. ترکیب خبرگی و تحلیل چالش‌ها: به‌منظور تفسیر یافته‌های کمی و شناسایی عوامل زمینه‌ای، گروهی پنج‌نفره از خبرگان و متخصصان حوزه عتف تشکیل شد. سپس، داده‌های خام عملکرد و روند شاخص‌ها در اختیار خبرگان قرار گرفت و نظرات آنان در قالب جلسات ساختاریافته خبرگی گردآوری شد. برونداد این مرحله، شناسایی چالش‌های اصلی اجرایی (مانند دسترسی به داده یا محاسبه‌پذیری) و اولویت‌های اصلاحی بود.

۴. تجویز راهکارهای سیاستی: بر مبنای یافته‌های مراحل ۲ و ۳، تحلیل نقاط قوت و ضعف چهارچوب دوره سوم انجام شد. همچنین، راهکارهای بهبود در سه افق زمانی کوتاه‌مدت، میان‌مدت و بلندمدت دسته‌بندی برای استقرار موفق نظام پایش ملی STI پیشنهاد شد.

ملاحظات روایی، پایایی و اخلاق پژوهش

برای تضمین روایی محتوا (اعتبار)، چهارچوب شاخص‌ها عیناً از گزارش طراحی چهارچوب پیشنهادی (Aghaie, 2021) اخذ شد. انتخاب منابع داده نیز مبتنی بر رسمی و معتبر بودن آن‌ها بوده است. در بخش کمی، تمامی مراحل و فرمول‌های محاسبه به شکلی ارائه شده‌اند که مطالعه قابل تکرار باشد و پایایی (قابلیت اطمینان) پژوهش

۱. گردآوری و یکپارچه‌سازی داده‌های واقعی: در این مرحله داده‌های عددی هریک از ۴۰ شاخص از گزارش‌های شناسایی‌شده استخراج شد. سپس، داده‌ها در یک ماتریس واحد با ساختار «شاخص × سال» سازماندهی و یکپارچه شدند تا امکان تحلیل مقایسه‌ای و روندیابی فراهم شود؛

۲. تحلیل کمی روند و ارزیابی عملکرد: این مرحله هسته اعتبارسنجی تجربی را تشکیل می‌دهد. ابتدا برای هر شاخص، نرخ رشد سالانه مرکب (CAGR) به‌عنوان شاخص اصلی سنجش روند محاسبه شد. فرمول به‌کارگرفته‌شده به‌صورت ذیل است (CFI, 2023):

$$CAGR = \left[\frac{\text{ارزش سال پایانی}}{\text{ارزش سال}} \right]^{\frac{1}{\text{تعداد سال‌ها}}} - 1 \quad (1)$$

در بخش بعدی، عملکرد هر شاخص براساس دو معیار ارزیابی

لحاظ شده است. به همین منظور، در بخش کیفی نیز فرایند گردآوری نظرات خبرگان به‌صورت ساختاریافته و مستند انجام پذیرفته است. همچنین، با رعایت اصول اخلاقی، از داده‌های ثانویه و عمومی استفاده شده و در گزارش‌دهی به حریم خصوصی و محرمانگی داده‌ها توجه کامل شده است.

یافته‌های پژوهش: اعتبارسنجی چهارچوب از طریق آزمون شاخص‌ها

در این بخش نتایج پژوهش به‌طور عینی و بدون تفسیر، مشتمل بر داده‌های خام و یافته‌های کمی حاصل از فرایند تحلیل ارائه می‌شود. از این‌رو، ابتدا با توجه به میزان حصول داده‌های آماری ملی و بین‌المللی شاخص‌های چهارچوب پیشنهادی و میزان استاندارد بودن تعاریف و اصطلاحات آماری به‌کارگرفته‌شده، نتایج به‌کارگیری چهارچوب، به تفکیک سه چارک ارزیابی کارایی، اثربخشی و سودمندی ارائه می‌شود. سپس تحلیل عملکرد چارک‌های ارزیابی و درجه انطباق چهارچوب جمع‌بندی شده است.

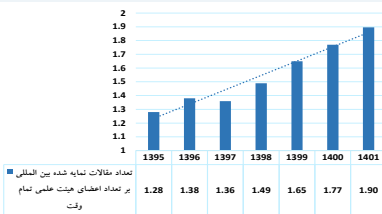
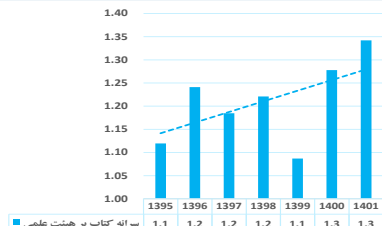
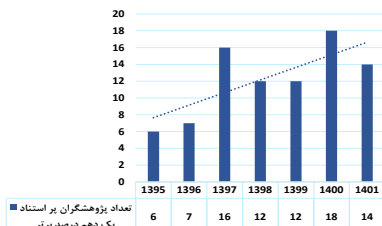
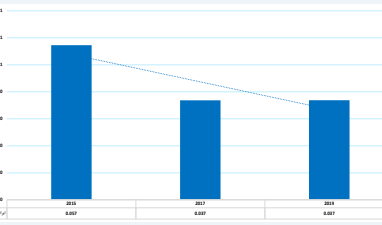
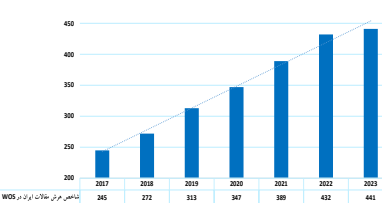
شاخص‌های چارک کارایی

شاخص‌های کارایی، نسبت بروندادها به دروندادها را می‌سنجند و ناظر به پاسخگویی در قبال ذی‌نفعان درونی نظام علم و فناوری اعم بر دانشجویان، اعضای هیئت علمی، پژوهشگران، فناوران و نوآوران است. در جدول ۲ نتایج تحلیل عملکرد کارایی نظام علم و فناوری کشور در بازه زمانی برنامه ششم پیشرفت براساس چهارچوب پیشنهادی مشاهده می‌شود.

1. Compliance Score
2. Data Accessibility, Analyzability, Being up-to-date

جدول ۲. نتایج تحلیل شاخص‌های کارایی ارزیابی عملکرد نظام علم و فناوری

شاخص‌های کارایی	معادل بین‌المللی شاخص	آمار داخلی	آمار بین‌المللی	تغییرات در بازه زمانی برنامه ششم پیشرفت
۱ شدت تحقیق و توسعه (نسبت هزینه‌کرد ناخالص داخلی تحقیق و توسعه به تولید ناخالص داخلی)	GERD as percentage of GDP	سازمان برنامه‌وبودجه	Unesco	
۲ سرانه مقالات نمایه‌شده در پایگاه‌های معتبر بین‌المللی به پژوهشگر تمام‌وقت	Number of journal articles (indexed by Scopus/ WOS) per researcher (FTE)	وزارت عتف	World bank/ Unesco	
۳ نسبت تعداد اختراعات، اکتشافات، نشان‌های تجاری یا برندهای به‌ثبت‌رسیده بر میزان هزینه‌کرد تحقیق و توسعه	Total patent, trade mark, design and hearing data registrations/ GERD	کانون پتنت ایران	WIPO/ Orbit	
۴ نسبت تعداد اختراعات، اکتشافات، نشان‌های تجاری یا برندهای به‌ثبت‌رسیده بر تعداد کل نیروی انسانی تحقیق و توسعه تمام‌وقت	Total patent, trade mark, design and hearing data registrations/ Number of R&D personnel (FTE)	کانون پتنت ایران	WIPO/ Orbit	
۵ سرانه بودجه آموزش عالی به‌ازای هر دانش‌آموخته دانشگاهی	Government expenditure on education in US\$ (millions) /Number of graduates from tertiary education	مؤسسه پژوهش و برنامه‌ریزی آموزش عالی / وزارت عتف	-	
۶ سرانه مقالات نمایه‌شده در پایگاه‌های معتبر بین‌المللی به دانشجوی تحصیلات تکمیلی	Number of journal articles (indexed by Scopus/ WOS) per postgraduate students	وزارت عتف / مؤسسه پژوهش و برنامه‌ریزی آموزش عالی	-	

شاخص‌های کارایی	معادل بین‌المللی شاخص	آمار داخلی	آمار بین‌المللی	تغییرات در بازه زمانی برنامه ششم پیشرفت
۷ سرانه مقالات نمایه‌شده در پایگاه‌های معتبر بین‌المللی به عضو هیئت علمی	Number of journal articles (indexed by Scopus/ WOS) per Faculty members (FTE)	وزارت عتف/ مؤسسه پژوهش و برنامه‌ریزی آموزش عالی	-	
۸ سرانه کتب (علمی) منتشرشده به عضو هیئت علمی	Number of published books per Faculty member	خانه کتاب و ادبیات ایران/ مؤسسه پژوهش و برنامه‌ریزی آموزش عالی	-	
۹ پژوهشگران پراستناد یک‌دهم درصد برتر براساس پایگاه‌های معتبر بین‌المللی	Highly Cited Researchers (HCR)	ISC مؤسسه استنادی جهان اسلام	Clarivate	
۱۰ نسبت میزان صادرات محصولات با فناوری‌های متوسط و پیشرفته بر میزان هزینه کرد تحقیق و توسعه	Share of Medium and High-Tech Activities Manufacturing Export in GDP/ GERD as percentage of GDP	وزارت عتف	Unido/ Unesco	
۱۱ شاخص هرش	H-index	ISC مؤسسه استنادی جهان اسلام	Scimago	

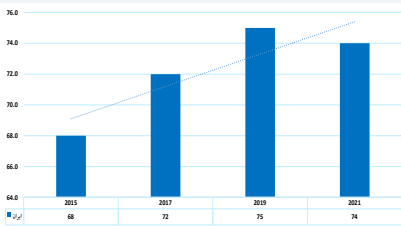
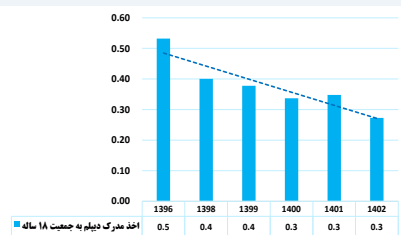
شاخص‌های چارک اثربخشی

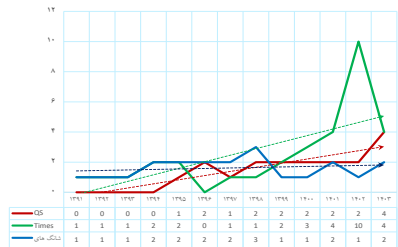
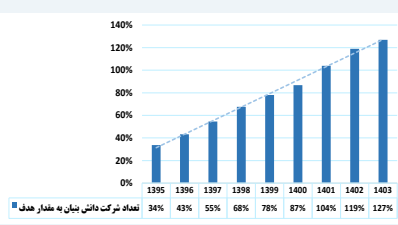
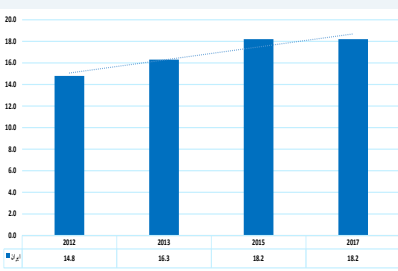
شاخص‌های اثربخشی، نسبت پیامدها (نتایج میان مدت و غیرمستقیم حاصل از فعالیت‌های علم و فناوری) به اهداف برنامه‌ای یا سیاستی را می‌سنجند و عمدتاً ناظر به پاسخگویی در قبال ذی‌نفعان بیرونی نظام علم و فناوری اعم بر بهره‌برداران صنعتی و حاکمیتی است. در جدول ۳ نتایج تحلیل عملکرد اثربخشی نظام علم و فناوری کشور در بازه زمانی برنامه ششم پیشرفت براساس چهارچوب پیشنهادی مشاهده می‌شود.

از جدول ۲ چنین برمی‌آید که از مجموع ۱۱ شاخص کارایی انتخاب‌شده در الگوی یادشده، پنج شاخص به‌طور کامل در هر دو سطح ملی و بین‌المللی (البته با کاستی‌هایی که بعضاً با شاخص جایگزین با موضوع مشابه رفع شده است) تحلیل و مقایسه شده است و شش شاخص فقط در سطح ملی (به‌دلیل نبود آمار بین‌المللی لازم) تحلیل شده است. شایان ذکر است برخی داده‌های جهانی همچون هزینه کرد تحقیق و توسعه یا صادرات محصولات فناوری‌محور برای همه سال‌های تحلیل مورد نظر موجود نبوده و براساس داده‌های موجود تحلیل انجام شده است.

جدول ۳. نتایج تحلیل شاخص‌های اثربخشی عملکرد نظام علم و فناوری

ردیف	شاخص‌های اثربخشی STI	معادل بین‌المللی شاخص	منبع آمار داخلی	منبع آمار بین‌المللی	تغییرات در بازه زمانی برنامه ششم پیشرفت																		
۱۲	بخش اقتصاد	نسبت (افزایش) بهره‌وری کل عوامل تولید بر میزان مطلوب آن	Total factor productivity	Penn World Table	<table border="1"> <thead> <tr> <th>سال</th> <th>1395</th> <th>1396</th> <th>1397</th> <th>1398</th> <th>1399</th> <th>1400</th> <th>1401</th> </tr> </thead> <tbody> <tr> <td>بهره‌وری کل عوامل تولید</td> <td>1.00</td> <td>1.02</td> <td>1.05</td> <td>1.08</td> <td>1.12</td> <td>1.15</td> <td>1.15</td> </tr> </tbody> </table>	سال	1395	1396	1397	1398	1399	1400	1401	بهره‌وری کل عوامل تولید	1.00	1.02	1.05	1.08	1.12	1.15	1.15		
سال	1395	1396	1397	1398	1399	1400	1401																
بهره‌وری کل عوامل تولید	1.00	1.02	1.05	1.08	1.12	1.15	1.15																
۱۳	بخش آب، کشاورزی و محیط‌زیست	سهم میزان انتشار گازهای گلخانه‌ای از مقدار کل جهان	Greenhouse Gas Emission (GHG) (million tones CO2 Eq.)/ Global Total emission	European Union	<table border="1"> <thead> <tr> <th>سال</th> <th>2015</th> <th>2016</th> <th>2017</th> <th>2018</th> <th>2019</th> <th>2020</th> <th>2021</th> </tr> </thead> <tbody> <tr> <td>سهم میزان انتشار گازهای گلخانه‌ای</td> <td>1.70%</td> <td>1.76%</td> <td>1.78%</td> <td>1.80%</td> <td>1.78%</td> <td>1.77%</td> <td>1.77%</td> </tr> </tbody> </table>	سال	2015	2016	2017	2018	2019	2020	2021	سهم میزان انتشار گازهای گلخانه‌ای	1.70%	1.76%	1.78%	1.80%	1.78%	1.77%	1.77%		
سال	2015	2016	2017	2018	2019	2020	2021																
سهم میزان انتشار گازهای گلخانه‌ای	1.70%	1.76%	1.78%	1.80%	1.78%	1.77%	1.77%																
۱۴	بخش آب، کشاورزی و محیط‌زیست	نسبت بهره‌وری کل آب بر مقدار هدف‌گذاری مطلوب آن	Total Water Productivity (constant 2015 US\$ GDP per cubic meter of total freshwater withdrawal)	World Bank	<table border="1"> <thead> <tr> <th>سال</th> <th>2015</th> <th>2016</th> <th>2017</th> <th>2018</th> <th>2019</th> <th>2020</th> </tr> </thead> <tbody> <tr> <td>نسبت بهره‌وری کل آب</td> <td>4</td> <td>5</td> <td>5</td> <td>5</td> <td>5</td> <td>5</td> </tr> </tbody> </table>	سال	2015	2016	2017	2018	2019	2020	نسبت بهره‌وری کل آب	4	5	5	5	5	5				
سال	2015	2016	2017	2018	2019	2020																	
نسبت بهره‌وری کل آب	4	5	5	5	5	5																	
۱۵	بخش تولید، انرژی و صنعت و معدن	نسبت سهم صادرات محصولات و خدمات با فناوری متوسط و پیشرفته از کل صادرات کشور بر مقدار هدف‌گذاری	Share of Medium and High-Tech Activities Manufacturing Export in Total Export	Unido	<table border="1"> <thead> <tr> <th>سال</th> <th>2015</th> <th>2016</th> <th>2017</th> <th>2018</th> <th>2019</th> <th>2020</th> <th>2021</th> </tr> </thead> <tbody> <tr> <td>نسبت سهم صادرات</td> <td>100%</td> <td>100%</td> <td>100%</td> <td>100%</td> <td>100%</td> <td>100%</td> <td>100%</td> </tr> </tbody> </table>	سال	2015	2016	2017	2018	2019	2020	2021	نسبت سهم صادرات	100%	100%	100%	100%	100%	100%	100%		
سال	2015	2016	2017	2018	2019	2020	2021																
نسبت سهم صادرات	100%	100%	100%	100%	100%	100%	100%																
۱۶	بخش تولید، انرژی و صنعت و معدن	نسبت (کاهش) شاخص شدت انرژی بر میزان هدف‌گذاری شده	Energy intensity (Total energy consumption per unit of GDP)	Enerdata	<table border="1"> <thead> <tr> <th>سال</th> <th>2015</th> <th>2016</th> <th>2017</th> <th>2018</th> <th>2019</th> <th>2020</th> <th>2021</th> <th>2022</th> </tr> </thead> <tbody> <tr> <td>شاخص شدت انرژی</td> <td>0.21</td> <td>0.199</td> <td>0.211</td> <td>0.221</td> <td>0.221</td> <td>0.212</td> <td>0.209</td> <td>0.207</td> </tr> </tbody> </table>	سال	2015	2016	2017	2018	2019	2020	2021	2022	شاخص شدت انرژی	0.21	0.199	0.211	0.221	0.221	0.212	0.209	0.207
سال	2015	2016	2017	2018	2019	2020	2021	2022															
شاخص شدت انرژی	0.21	0.199	0.211	0.221	0.221	0.212	0.209	0.207															
۱۷	بخش تولید، انرژی و صنعت و معدن	نسبت سهم اقتصاد ICT از GDP بر سهم مطلوب آن	ICT investment as a share GDP/ ICT industry as a share of GDP	statista	<table border="1"> <thead> <tr> <th>سال</th> <th>1392</th> <th>1393</th> <th>1394</th> <th>1395</th> <th>1396</th> <th>1397</th> <th>1398</th> </tr> </thead> <tbody> <tr> <td>نسبت سهم اقتصاد ICT</td> <td>143.60%</td> <td>133.60%</td> <td>130.40%</td> <td>109.80%</td> <td>92.70%</td> <td>86.00%</td> <td>75.00%</td> </tr> </tbody> </table>	سال	1392	1393	1394	1395	1396	1397	1398	نسبت سهم اقتصاد ICT	143.60%	133.60%	130.40%	109.80%	92.70%	86.00%	75.00%		
سال	1392	1393	1394	1395	1396	1397	1398																
نسبت سهم اقتصاد ICT	143.60%	133.60%	130.40%	109.80%	92.70%	86.00%	75.00%																

ردیف	شاخص‌های اثربخشی STI	معادل بین‌المللی شاخص	منبع آمار داخلی	منبع آمار بین‌المللی	تغییرات در بازه زمانی برنامه ششم پیشرفت														
	نسبت ارزش مصالح نوین و سبک استفاده‌شده بر کل هزینه ساختمانی کشور بر میزان هدف‌گذاری‌شده	-	-	-	-														
	نسبت ظرفیت سامانه‌های حمل‌ونقل هوشمند بر مقدار هدف‌گذاری‌شده	-	-	-	-														
	نسبت افزایش شاخص UHC (پوشش همگانی سلامت) بر مقدار مطلوب این شاخص	The Universal Health Coverage (UHC) Service Coverage Index	-	World Health Organization	 <table border="1"><thead><tr><th>سال</th><th>2015</th><th>2017</th><th>2019</th><th>2021</th></tr></thead><tbody><tr><td>شاخص</td><td>68</td><td>72</td><td>75</td><td>74</td></tr></tbody></table>	سال	2015	2017	2019	2021	شاخص	68	72	75	74				
سال	2015	2017	2019	2021															
شاخص	68	72	75	74															
	نسبت مبلغ درآمدهای ناشی از صادرات محصولات نظامی و دفاعی با فناوری‌های پیشرفته بر میزان هدف‌گذاری‌شده	-	-	-	-														
	نسبت سرانه میزان مطالعه بر مقدار مطلوب هدف‌گذاری‌شده	-	-	-	-														
	نسبت تعداد اخذکنندگان مدرک دیپلم بر جمعیت واجد شرایط آن سال (۱۸ ساله)	School enrollment, tertiary (% gross)	پژوهشگاه آموزش و پرورش	یونسکو	 <table border="1"><thead><tr><th>سال</th><th>1396</th><th>1398</th><th>1399</th><th>1400</th><th>1401</th><th>1402</th></tr></thead><tbody><tr><td>نسبت</td><td>0.5</td><td>0.4</td><td>0.4</td><td>0.3</td><td>0.3</td><td>0.3</td></tr></tbody></table>	سال	1396	1398	1399	1400	1401	1402	نسبت	0.5	0.4	0.4	0.3	0.3	0.3
سال	1396	1398	1399	1400	1401	1402													
نسبت	0.5	0.4	0.4	0.3	0.3	0.3													

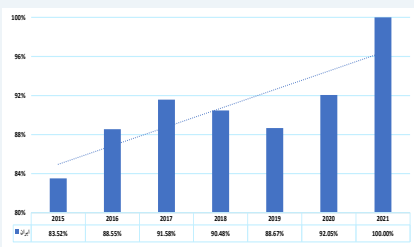
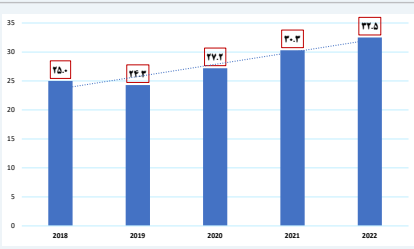
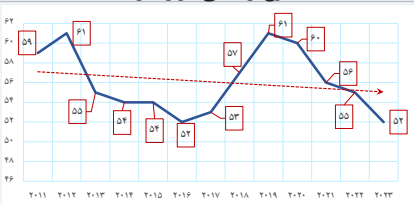
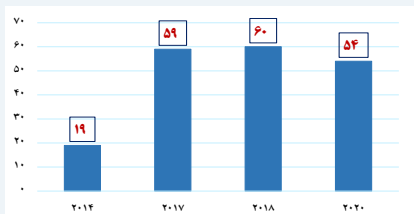
ردیف	شاخص‌های اثربخشی STI	معادل بین‌المللی شاخص	منبع آمار داخلی	منبع آمار بین‌المللی	تغییرات در بازه زمانی برنامه ششم پیشرفت																																																																																												
۲۴	نسبت تعداد دانشگاه‌های در زمره ۵۰۰ دانشگاه برتر جهان بر مقدار مطلوب هدف گذاری شده	The number of universities among the top 500 universities in the world	وزارت عتف	Times ranking/ Shanghai ranking/ QS ranking	 <table border="1"> <thead> <tr> <th>Year</th> <th>QS</th> <th>Times</th> <th>تتک</th> </tr> </thead> <tbody> <tr><td>1996</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>1997</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>1998</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>1999</td><td>0</td><td>2</td><td>2</td></tr> <tr><td>2000</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2001</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2002</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2003</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2004</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2005</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2006</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2007</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2008</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2009</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2010</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2011</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2012</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2013</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2014</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2015</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2016</td><td>1</td><td>2</td><td>2</td></tr> <tr><td>2017</td><td>1</td><td>2</td><td>2</td></tr> </tbody> </table>	Year	QS	Times	تتک	1996	0	1	1	1997	0	1	1	1998	0	1	1	1999	0	2	2	2000	1	2	2	2001	1	2	2	2002	1	2	2	2003	1	2	2	2004	1	2	2	2005	1	2	2	2006	1	2	2	2007	1	2	2	2008	1	2	2	2009	1	2	2	2010	1	2	2	2011	1	2	2	2012	1	2	2	2013	1	2	2	2014	1	2	2	2015	1	2	2	2016	1	2	2	2017	1	2	2
Year	QS	Times	تتک																																																																																														
1996	0	1	1																																																																																														
1997	0	1	1																																																																																														
1998	0	1	1																																																																																														
1999	0	2	2																																																																																														
2000	1	2	2																																																																																														
2001	1	2	2																																																																																														
2002	1	2	2																																																																																														
2003	1	2	2																																																																																														
2004	1	2	2																																																																																														
2005	1	2	2																																																																																														
2006	1	2	2																																																																																														
2007	1	2	2																																																																																														
2008	1	2	2																																																																																														
2009	1	2	2																																																																																														
2010	1	2	2																																																																																														
2011	1	2	2																																																																																														
2012	1	2	2																																																																																														
2013	1	2	2																																																																																														
2014	1	2	2																																																																																														
2015	1	2	2																																																																																														
2016	1	2	2																																																																																														
2017	1	2	2																																																																																														
۲۵	نسبت تعداد شرکت‌های دانش‌بنیان بر مقدار هدف گذاری شده	-	معاونت علمی، فناوری و اقتصاد دانش‌بنیان	-	 <table border="1"> <thead> <tr> <th>Year</th> <th>Percentage</th> </tr> </thead> <tbody> <tr><td>1395</td><td>34%</td></tr> <tr><td>1396</td><td>43%</td></tr> <tr><td>1397</td><td>55%</td></tr> <tr><td>1398</td><td>68%</td></tr> <tr><td>1399</td><td>78%</td></tr> <tr><td>1400</td><td>87%</td></tr> <tr><td>1401</td><td>104%</td></tr> <tr><td>1402</td><td>119%</td></tr> <tr><td>1403</td><td>127%</td></tr> </tbody> </table>	Year	Percentage	1395	34%	1396	43%	1397	55%	1398	68%	1399	78%	1400	87%	1401	104%	1402	119%	1403	127%																																																																								
Year	Percentage																																																																																																
1395	34%																																																																																																
1396	43%																																																																																																
1397	55%																																																																																																
1398	68%																																																																																																
1399	78%																																																																																																
1400	87%																																																																																																
1401	104%																																																																																																
1402	119%																																																																																																
1403	127%																																																																																																
۲۶	نسبت تعداد دارندگان مدرک دانشگاهی (حداقل کارشناسی یا معادل آن) بر جمعیت ۲۵ ساله و بیشتر (تجمعی)	Educational attainment, at least Bachelor's or equivalent, population 25+, total (%) (cumulative)	-	OECD/ World Bank	 <table border="1"> <thead> <tr> <th>Year</th> <th>Percentage</th> </tr> </thead> <tbody> <tr><td>2012</td><td>14.8</td></tr> <tr><td>2013</td><td>16.1</td></tr> <tr><td>2015</td><td>18.2</td></tr> <tr><td>2017</td><td>18.2</td></tr> </tbody> </table>	Year	Percentage	2012	14.8	2013	16.1	2015	18.2	2017	18.2																																																																																		
Year	Percentage																																																																																																
2012	14.8																																																																																																
2013	16.1																																																																																																
2015	18.2																																																																																																
2017	18.2																																																																																																

شاخص‌های چارک سودمندی

شاخص‌های سودمندی یا اثربخشی، نسبت تأثیر اجرای برنامه یا فعالیت (نتایج بلندمدت و دارای مدت اثر زیاد) به دروندادهای استفاده‌شده برای دستیابی به آن‌ها را نشان می‌دهند. آنچه در سودمندی مدنظر است، عمدتاً ناظر به پاسخگویی در قبال ذی‌نفعان بیرونی نهایی نظام علم و فناوری، به‌ویژه عموم مردم جامعه است. در جدول ۴ نتایج تحلیل عملکرد سودمندی نظام علم و فناوری کشور در بازه زمانی برنامه ششم پیشرفت براساس چهارچوب پیشنهادی مشاهده می‌شود.

از جدول بالا چنین برمی‌آید که از مجموع ۱۵ شاخص اثربخشی انتخاب‌شده در الگوی یادشده، شاخص‌های «نسبت ارزش مصالح نوین و سبک استفاده‌شده بر کل هزینه ساختمانی کشور بر میزان هدف‌گذاری‌شده»، «نسبت ظرفیت سامانه‌های حمل‌ونقل هوشمند بر مقدار هدف‌گذاری‌شده»، «نسبت مبلغ درآمدهای ناشی از صادرات محصولات نظامی و دفاعی با فناوری‌های پیشرفته بر میزان هدف‌گذاری‌شده» و «نسبت سرانه میزان مطالعه بر مقدار مطلوب هدف‌گذاری‌شده»، دارای سبقت گردآوری داده در مراجع آماری کشور نبوده (یا به‌دلیل محرمانگی از دسترس پژوهشگران خارج بوده است) و شاخص دقیق متناظر آن نیز در مراجع بین‌المللی وجود ندارد؛ از این‌رو، در تحلیل‌های این بخش وارد نشده است. داده‌های ۱۱ شاخص باقی‌مانده، غالباً در سطح بین‌المللی (البته با کاستی‌هایی که بعضاً با شاخص با موضوع مشابه جایگزین شده است) و بعضاً در سطح ملی (در صورت وجود داده) تحلیل شده است.

جدول ۴. نتایج تحلیل شاخص‌های سودمندی ارزیابی عملکرد نظام علم و فناوری

شاخص‌های سودمندی	معادل بین‌المللی شاخص	منبع آمار داخلی	منبع آمار بین‌المللی	تغییرات در بازه زمانی برنامه ششم پیشرفت
۲۷ نسبت (افزایش) سرانه تولید ناخالص داخلی برحسب برابری قدرت خرید (ppp) بر میزان فعلی آن	GDP per capita in terms of purchasing power parity (PPP)	-	World bank	
۲۸ نسبت (افزایش) شاخص کیفیت هوا به مقدار فعلی این شاخص	Air quality index based on annual average PM2.5 concentration (µg/m3)	سازمان هواشناسی	IQAir	
۲۹ نسبت (افزایش) شاخص امنیت غذایی بر مقدار فعلی این شاخص	Food security index	کمیسیون تخصصی کشاورزی	UNCCD Knowledge Hub	آمار بین‌المللی برای همه کشورها به جز ایران در پایگاه اطلاعاتی Knowledge Hub UNCCD موجود است.
۳۰ نسبت (افزایش) شاخص امنیت آبی بر مقدار فعلی این شاخص	Water security index	کمیسیون تخصصی کشاورزی	The United Nations University Institute for Water Environment and Health	براساس گزارش سال ۲۰۲۳ سازمان ملل متحد، کشور ایران با امتیاز ۴۸ در وضعیت ناامن از نظر شاخص امنیت آبی قرار دارد.
۳۱ نسبت (افزایش) شاخص تنوع زیستی بر مقدار فعلی این شاخص	Biodiversity Habitat Index (BHI)	کمیسیون تخصصی کشاورزی	CSIRO (Australia's national science agency)	براساس گزارش سال ۲۰۲۲ دانشگاه ییل، ایران با رتبه ۱۵۰ از میان ۱۸۰ کشور دنیا از نظر شاخص تنوع زیستی قرار دارد.
۳۲ نسبت (افزایش) شاخص رقابت‌پذیری صنعتی بر میزان فعلی آن	Competitive Industrial Performance Index	-	Unido	
۳۳ نسبت (افزایش) شاخص امنیت سایبری بر میزان فعلی آن	Cybersecurity Index	-	International Telecommunication Union	
۳۴ نسبت (افزایش) میزان شاخص کیفیت زندگی بر مقدار فعلی آن	The quality of life index	-	Economist Intelligence Unit	

شاخص های سودمندی	معادل بین المللی شاخص	منبع آمار داخلی	منبع آمار بین المللی	تغییرات در بازه زمانی برنامه ششم پیشرفت																
۳۵ نسبت (کاهش) شاخص بار بیماری ها بر میزان این شاخص در سال مبدأ	Burden of Disease index	-	Health data organization	<table><tr><th>Year</th><th>Value</th></tr><tr><td>2015</td><td>27153.26</td></tr><tr><td>2016</td><td>26260.4</td></tr><tr><td>2017</td><td>25820.28</td></tr><tr><td>2018</td><td>25418.35</td></tr><tr><td>2019</td><td>25029.07</td></tr></table>	Year	Value	2015	27153.26	2016	26260.4	2017	25820.28	2018	25418.35	2019	25029.07				
Year	Value																			
2015	27153.26																			
2016	26260.4																			
2017	25820.28																			
2018	25418.35																			
2019	25029.07																			
۳۶ نسبت (افزایش) شاخص ایمنی و امنیت لگاتوم (Legatum) بر میزان فعلی آن	THE LEGATUM PROSPERITY INDEX	-	مؤسسه پژوهشی لگاتوم	<table><tr><th>Year</th><th>Value</th></tr><tr><td>2015</td><td>51.59</td></tr><tr><td>2016</td><td>54.77</td></tr><tr><td>2017</td><td>53.94</td></tr><tr><td>2018</td><td>55.74</td></tr><tr><td>2019</td><td>56.8</td></tr><tr><td>2020</td><td>55.1</td></tr><tr><td>2021</td><td>54.9</td></tr></table>	Year	Value	2015	51.59	2016	54.77	2017	53.94	2018	55.74	2019	56.8	2020	55.1	2021	54.9
Year	Value																			
2015	51.59																			
2016	54.77																			
2017	53.94																			
2018	55.74																			
2019	56.8																			
2020	55.1																			
2021	54.9																			
۳۷ نسبت (کاهش) میزان کلی آسیب ها و ناهنجاری های اجتماعی بر میزان فعلی آن	-	=تعداد محکومین + کیفری + آمار اعتیاد + تعداد کودکان بی سرپرست	-	-																
۳۸ نسبت تعداد اشتغال فارغ التحصیلان دانشگاهی بر تعداد کل فارغ التحصیلان	Employment rate of university graduates	مؤسسه پژوهش و برنامه ریزی آموزش عالی	World bank	<table><tr><th>Year</th><th>Value</th></tr><tr><td>1397</td><td>49.09</td></tr><tr><td>1398</td><td>63.67</td></tr><tr><td>1399</td><td>65.21</td></tr></table>	Year	Value	1397	49.09	1398	63.67	1399	65.21								
Year	Value																			
1397	49.09																			
1398	63.67																			
1399	65.21																			
۳۹ نسبت (رشد) شاخص نوآوری بر مقدار فعلی آن	Global Innovation Index	-	Global economy																	
۴۰ نسبت (رشد) شاخص حاکمیت جهانی بر																				

سابقه آماری در داخل و شاخص متناظر بین المللی، در تحلیل این بخش در نظر گرفته نشدند. برای برخی شاخص ها نیز به رغم نبود آمار داخلی، تحلیل براساس آمار بین المللی انجام شد.

از جدول بالا چنین برمی آید که از مجموع ۱۴ شاخص سودمندی انتخاب شده در الگوی یادشده، شاخص «نسبت (کاهش) میزان کلی آسیب ها و ناهنجاری های اجتماعی بر میزان فعلی آن» به دلیل نبود

نتایج تحلیل روند و عملکرد شاخص‌ها

جمع‌بندی نتایج کلی تحلیل روند شاخص به تفکیک چارک‌های ارزیابی در جدول ۵ آورده شده است.

جدول ۵. مقایسه نتایج ارزیابی شاخص‌های کارایی، اثربخشی و سودمندی نظام علم و فناوری ایران

شماره	چارک ارزیابی	شاخص ارزیابی	وضعیت ایران	
			سطح	روند تغییر
۱	کارایی	شدت تحقیق و توسعه	بسیار ضعیف	نزولی
۲		سرانه مقالات نمایه‌شده در پایگاه‌های معتبر بین‌المللی به پژوهشگر	خوب	صعودی
۳		تعداد اختراعات و نشان‌های تجاری به ثبت بین‌المللی	بسیار ضعیف	نزولی
۴		نسبت تعداد اختراعات و نشان‌های تجاری به ثبت‌رسیده بین‌المللی بر تعداد کل نیروی انسانی تحقیق و توسعه	بسیار ضعیف	نزولی
۵		سرانه بودجه آموزش عالی به‌ازای هر دانش‌آموخته دانشگاهی	ضعیف	صعودی
۶		سرانه مقالات نمایه‌شده در پایگاه‌های معتبر بین‌المللی به دانشجوی تحصیلات تکمیلی	متوسط	صعودی
۷		سرانه مقالات نمایه‌شده در پایگاه‌های معتبر بین‌المللی به عضو هیئت علمی	متوسط	صعودی
۸		سرانه کتب علمی منتشرشده به عضو هیئت علمی	متوسط	صعودی
۹		پژوهشگران پراستناد یک‌دهم درصد برتر براساس پایگاه‌های معتبر بین‌المللی	ضعیف	صعودی
۱۰		نسبت میزان صادرات محصولات با فناوری‌های متوسط و پیشرفته بر میزان هزینه‌کرد تحقیق و توسعه	بسیار ضعیف	نزولی
۱۱		شاخص هرش	ضعیف	صعودی
۱۲	اثربخشی	بهره‌وری کل عوامل تولید	ضعیف	نزولی
۱۳		سهم میزان انتشار گازهای گلخانه‌ای از مقدار کل جهان*	بسیار ضعیف	صعودی
۱۴		بهره‌وری کل آب	بسیار ضعیف	ثابت
۱۵		نسبت سهم صادرات محصولات و خدمات با فناوری متوسط و پیشرفته از کل صادرات کشور	ضعیف	نزولی
۱۶		شاخص شدت انرژی*	بسیار ضعیف	ثابت
۱۷		نسبت سهم اقتصاد ICT از GDP بر سهم مطلوب آن	متوسط	نزولی
۱۸		ارزش مصالح نوین و سبک استفاده‌شده بر کل هزینه ساختمانی کشور	حذف	
۱۹		ظرفیت سامانه‌های حمل‌ونقل هوشمند	حذف	
۲۰		شاخص UHC (پوشش همگانی سلامت)	خوب	صعودی
۲۱		مبلغ درآمدهای ناشی از صادرات محصولات نظامی و دفاعی با فناوری‌های پیشرفته	حذف	
۲۲		سرانه میزان مطالعه	حذف	
۲۳		نسبت تعداد اخذکنندگان مدرک دیپلم بر جمعیت واجد شرایط آن سال (۱۸ساله)	ضعیف	نزولی
۲۴		نسبت تعداد دانشگاه‌های در زمره ۵۰۰ دانشگاه برتر جهان بر مقدار مطلوب هدف‌گذاری‌شده	متوسط	صعودی
۲۵		نسبت تعداد شرکت‌های دانش‌بنیان بر مقدار هدف‌گذاری‌شده	عالی	صعودی
۲۶		نسبت تعداد دارندگان مدرک دانشگاهی (حداقل کارشناسی یا معادل آن) بر جمعیت ۲۵ساله و بیشتر (تجمعی)	متوسط	صعودی

شماره	چارک ارزیابی	شاخص ارزیابی		وضعیت ایران	
				سطح	روند تغییر
۲۷	سودمندی	سراشته تولید ناخالص داخلی برحسب برابری قدرت خرید (ppp)		ضعیف	صعودی
۲۸		شاخص کیفیت هوا*		ضعیف	صعودی
۲۹		شاخص امنیت غذایی		دسترسی نداشتن به آمار دقیق	
۳۰		شاخص امنیت آبی		ضعیف	
۳۱		شاخص تنوع زیستی		ضعیف	
۳۲		شاخص رقابت پذیری صنعتی		ضعیف	نزولی
۳۳		شاخص امنیت سایبری		متوسط	صعودی
۳۴		شاخص کیفیت زندگی		بسیار ضعیف	نزولی
۳۵		شاخص بار بیماری‌ها*		متوسط	نزولی
۳۶		شاخص ایمنی و امنیت لگاتوم		ضعیف	صعودی
۳۷		میزان کلی آسیب‌ها و ناهنجاری‌های اجتماعی		حذف	
۳۸		نسبت تعداد اشتغال فارغ‌التحصیلان دانشگاهی بر تعداد کل فارغ‌التحصیلان		متوسط	ثابت
۳۹		شاخص جهانی نوآوری		متوسط	صعودی
۴۰		شاخص جهانی حکمرانی		بسیار ضعیف	نزولی

* شاخص‌های روند تحلیل آن معکوس است.

هر شاخص در بازه ۱۳۹۶-۱۴۰۰ و مقایسه آن با اهداف برنامه ششم (برای شاخص‌های دارای تناظر)، نتایج جدول ۶ را در سطح چارک‌ها نشان می‌دهد. میانگین وزنی CAGR برای کل ۳۴ شاخص (دارای آمار از مجموع ۴۰ شاخص چهارچوب) معادل ۱/۷ درصد به دست آمده است. شایان ذکر است که شاخص‌های بدون آمار (دقیق) در محاسبه این شاخص لحاظ نمی‌شوند.

مطابق آنچه در جدول ۵ جمع‌بندی شده، فقط یک شاخص (نسبت تعداد شرکت‌های دانش‌بنیان بر مقدار هدف‌گذاری شده) در سطح «عالی» و دو شاخص (سراشته مقالات نمایه‌شده در پایگاه‌های معتبر بین‌المللی به پژوهشگر و شاخص پوشش همگانی سلامت) در سطح «خوب»، برای کشور ایران برآورد شده است و عمدتاً شاخص‌های مطرح‌شده در سطح «ضعیف» و «بسیار ضعیف» هستند. به‌منظور ایجاد تقریب ذهنی، با محاسبه نرخ رشد سالانه مرکب (CAGR) برای

جدول ۶. عملکرد چارک‌های ارزیابی براساس نتایج تحلیل روند

چارک	میانگین CAGR	سطح عملکرد	تعداد شاخص‌های با عملکرد «خوب» و «عالی»	تعداد شاخص‌های با عملکرد «ضعیف» و «بسیار ضعیف»
کارایی	۱/۵ درصد	ضعیف	۱ از ۱۱	۷ از ۱۱
اثر بخشی	۱/۸ درصد	ضعیف	۲ از ۱۱	۹ از ۱۱
اثر گذاری	۱/۷ درصد	ضعیف	۰ از ۱۲	۱۲ از ۱۲
میانگین کل چهارچوب	۱/۷	ضعیف	۳ از ۳۴	۳۱ از ۳۴

میانگین هم‌پوشانی این شاخص‌ها با استانداردهای بین‌المللی همچون (OECD، فراسکاتی، اسلو و ...) معادل ۸۵ درصد است. توزیع این انطباق در چارک‌های ارزیابی چهارچوب به‌ترتیب: چارک کارایی ۹۷ درصد، چارک اثر بخشی ۶۷ درصد و چارک اثر گذاری (سودمندی) ۹۲ درصد بوده است. در جدول ۷ وضعیت انطباق‌پذیری برخی شاخص‌های منتخب در هر چارک ارائه شده است.

به نظر می‌رسد میانگین وضعیت هر سه چارک و میانگین آن‌ها برای ایران در سطح «ضعیف» برآورد شده است. البته شایان ذکر است که برآورد ارائه‌شده در این بخش براساس داده‌های موجود بوده است و در صورت دسترسی به آمار لازم برای تمامی شاخص‌های پیشنهادی، ممکن است نتیجه فعلی تغییر کند.

نتایج انطباق‌سنجی شاخص‌ها

بررسی تطبیقی ۴۰ شاخص چهارچوب دوره سوم نشان داد که

جدول ۷. داده‌های انطباق و قابلیت محاسبه شاخص‌های منتخب

ردیف	شاخص	چارک	نمره انطباق (۱-۰)	نمره DAB (دسترسی / تحلیل پذیری / به روز بودن*)
۱	نسبت هزینه کرد تحقیق و توسعه به تولید ناخالص داخلی	کارایی	۰/۹۲	۰/۰۲
۲	سرانه مقالات نمایه شده در پایگاه‌های معتبر بین‌المللی به عضو هیئت علمی	کارایی	۰/۷۱	۰/۱۲
۳	نسبت سهم صادرات محصولات و خدمات با فناوری متوسط و پیشرفته از کل صادرات کشور بر مقدار هدف‌گذاری	اثر بخشی	۰/۷۹	۰/۱۴
۴	نسبت سرانه میزان مطالعه بر مقدار مطلوب هدف‌گذاری شده	اثر بخشی	۰/۷	۰/۷۴
۵	نسبت (کاهش) میزان کلی آسیب‌ها و ناهنجاری‌های اجتماعی بر میزان فعلی آن	اثر گذاری	۰/۶۵	۰/۷۶
۶	نسبت (رشد) شاخص نوآوری بر مقدار فعلی آن	اثر گذاری	۰/۹۸	۰/۹۵

* مقیاس نمره DAB: صفر (عالی) تا ۱ (بسیار ضعیف)

♦ دسترسی به داده (D): آیا داده‌های خام این شاخص به راحتی و به طور عمومی از یک سامانه متمرکز و معتبر قابل استخراج است؟

♦ تحلیل پذیری (A): آیا داده‌های موجود به فرمت مناسب (مثلاً سری زمانی طولانی، تفکیک استانی/سازمانی) برای تحلیل روند و مقایسه فراهم‌اند؟

♦ به روز بودن (B): آیا داده‌ها با تأخیر زمانی قابل قبول (مثلاً حداکثر یک سال) به روزرسانی می‌شوند؟

جمع‌بندی یافته‌ها

درباره قابلیت اجرا یا ارزش افزوده چهارچوب پیشنهادی چه چیزی را نشان می‌دهد؟» به شرح جدول ۸ ارائه شده است.

در این بخش جمع‌بندی یافته‌ها در پاسخ به این پرسش که «هر یافته

جدول ۸. جمع‌بندی یافته‌ها در ارتباط با پرسش‌های پژوهش

پرسش اعتبارسنجی	روش پاسخ‌دهی	نتیجه اعتبارسنجی
آیا شاخص‌های چهارچوب با استانداردهای جهانی همسو هستند؟	محاسبه نمره انطباق (جدول ۷)	۸۵ درصد میانگین انطباق
آیا داده‌های مورد نیاز شاخص‌ها در دسترس است؟	محاسبه نمره DAB (دسترسی، تحلیل پذیری، به روز بودن)	نمره پایین در حوزه اثر بخشی
آیا چهارچوب قادر به اندازه‌گیری عملکرد واقعی است؟	محاسبه CAGR شاخص‌ها با داده‌های واقعی	عملکرد ضعیف نظام (میانگین ۱/۷ درصد)
آیا چهارچوب در مقایسه با دوره قبل ارزش افزوده دارد؟	مقایسه تعداد شاخص‌ها و رویکردها	کاهش از ۱۱۲ به ۴۰ شاخص

تفسیر یافته‌ها در چهارچوب نظری و بافتار تجربی

یافته‌های این پژوهش که حاصل اعتبارسنجی تجربی چهارچوب دوره سوم پایش و ارزیابی STI است، تصویری چندوجهی از بلوغ طراحی و چالش‌های اجرایی این چهارچوب، و به تبع آن، از نقاط قوت و ضعف نظام حکمرانی علم و فناوری ایران به شرح ادامه ارائه می‌دهد.

سطح انطباق، جایگاه در منظومه جهانی و جهت گیری آینده

میانگین انطباق ۸۵ درصدی شاخص‌های چهارچوب دوره سوم با استانداردهای بین‌المللی پیشرو (همچون OECD، فراسکاتی و اسلو)، نشان‌دهنده همگرایی مناسب طراحی نظری با پارادایم‌های جهانی

است. این سطح از انطباق، جایگاه چهارچوب ایران را در طیف الگوهای دوره سوم کشورهای در حال توسعه، هم‌تراز با تجربیاتی مانند ترکیه (TÜBİTAK) و کره جنوبی (NRF)، تثبیت می‌کند. مقایسه تطبیقی ساختاری این تجربه با نهادهای مرجع جهانی (جدول ۹) نشان می‌دهد که مسیر تکاملی عتف، از منطق مداخله به سمت زنجیره ارزش، در چهارچوب کلی گذار جهانی از رویکردهای کمی و توصیفی به سمت مدل‌های یکپارچه، فرابندمحور و سیاست‌محور قرار دارد.

جدول ۹. مقایسه تطبیقی تجربه شورای عالی عتف با نهادهای مرجع جهانی در پایش و ارزیابی STI

معیار مقایسه	عتف	NSF	TÜBİTAK	NRF	CNRS
مسیر تکامل (دوره‌ها)	سه دوره (منطق مداخله کارکردی ۱۳۸۸/ زنجیره ارزش فرایندی) ۱۳۹۵	چهار دوره (توصیفی ۱۹۷۰/ تحلیلی ۱۹۹۰/ سیاست‌محور ۲۰۱۰/ داده‌محور ۲۰۲۵)	سه دوره (شاخص‌محور ۲۰۰۵/ کارکردی ۲۰۱۵/ فرایندمحور ۲۰۲۵)	سه دوره (کمی پژوهش ۱۹۹۰/ کارکردی- همکاری ۲۰۰۰/ فرایندمحور- یکپارچه ۲۰۱۰)	دو دوره (شاخص پژوهشی ۱۹۹۰/ اثرگذاری اجتماعی- بین‌رشته‌ای ۲۰۲۰)
تعداد شاخص‌های شاخص‌های دوره اخیر	۴۰ شاخص (کاهش هدفمند از ۱۱۲ شاخص دوره قبل)	۲۰۰+ شاخص (نظام‌های گسترده)	۹۵ شاخص	۸۵ شاخص	۷۲ شاخص
رویکرد محوری دوره اخیر	چارک‌های ارزیابی مبتنی بر زنجیره ارزش و ذی‌نفعان کلیدی	داده‌محوری و هوشمندی پیش‌بینی‌کننده ^۱	فرایند نوآوری و بومی‌سازی چهارچوب‌های OECD	پیوند دانشگاه- صنعت و نظام‌های پایش یکپارچه	ارزیابی بین‌رشته‌ای و اثرگذاری اجتماعی
ابزار/ سازوکار اجرایی متمایز	آزمون بتا (چرخه طراحی- آزمون- بازخورد- بازطراحی پیش از استقرار رسمی)	«GPS Dashboard» (داشبوردهای تعاملی و پویا برای رصد بی‌درنگ و پیش‌بینی روندها)	«سامانه پایشگر ملی نوآوری» (پایگاه داده متمرکز برای پایش زیست‌بوم نوآوری)	نظام‌های ارزیابی همکاری دانشگاه- صنعت (شاخص‌های مشترک و ارزیابی اثرگذاری اقتصادی)	ارزیابی‌های بین‌المللی و هم‌تاسازی (تأکید بر جایگاه بین‌المللی و مقایسه با مراکز برتر جهانی)
نقاط قوت برجسته	یادگیری از آسیب‌شناسی و طراحی مستقل دوره جدید/ تأکید بر اعتبارسنجی پیشینی با آزمون بتا/ پوشش کامل چارک اثرگذاری در زنجیره ارزش	تاریخچه طولانی و تکامل پیوسته/ داده‌های غنی و زیرساخت‌های تحلیل پیشرفته/ تأثیرگذاری گسترده بر سیاست‌گذاری علم جهانی	الگوبرداری مؤثر و بومی‌سازی کارآمد استانداردهای OECD/ هم‌سویی نزدیک با اهداف توسعه ملی و منطقه‌ای	موفقیت عملی در پیوند پژوهش و صنعت/ نظام انعطاف‌پذیر و سازگار با تحولات سریع	تأکید بسیار بر کیفیت پژوهش و برتری علمی/ ادغام ارزیابی پژوهش با اهداف اجتماعی کلان
چالش‌های اصلی (برگرفته از اسناد موجود)	ضعف در دسترسی به‌موقع و باکیفیت داده/ پراکندگی سامانه‌های اطلاعاتی موجود/ نیاز به هماهنگی فراگیر در سطح حکمرانی برای استقرار موفق	پیچیدگی فرایند نظام ارزیابی و چالش در حفظ تمرکز بر اولویت‌های سیاستی در حجم انبوه داده	حفظ تعادل بین استانداردهای بین‌المللی و الزامات خاص توسعه ملی	اندازه‌گیری کیفی و بلندمدت اثرات همکاری‌ها فراتر از شاخص‌های کمی کوتاه‌مدت	سازگاری ارزیابی‌های بین‌المللی و کیفی با فشارهای سیاسی برای نشان دادن نتایج ملموس و سریع
جهت‌گیری آینده	استقرار نظام هوشمند پایش و ارزیابی ملی یکپارچه، مبتنی بر داده‌های واقعی و دارای حلقه بازخورد یادگیرنده	حکمرانی مبتنی بر پیش‌بینی و سامانه‌های هوش مصنوعی برای الگوسازی سیاست	رهبری منطقه‌ای در پایش نوآوری و تبدیل شدن به الگو برای کشورهای همسایه	نوآوری باز و پایش شبکه‌های جهانی دانش و فناوری	پژوهش مسئله‌محور با تمرکز بر چالش‌های اجتماعی پیچیده مانند تغییرات اقلیمی

بر اساس تحلیل تطبیقی جدول ۹، سه گزاره محوری می‌توان استنتاج کرد:

الف) چهارچوب دوره سوم عتف از نظر طراحی مفهومی، ساختار چارک‌ها و تأکید بر زنجیره ارزش در سطحی قابل قبول و هم‌تراز با دوره‌های سوم نهادهای مرجعی چون TÜBİTAK و NRF قرار دارد. ویژگی ممتاز آن، یعنی «آزمون بتا» به‌عنوان سازوکاری برای اعتبارسنجی پیشینی و یادگیری پیش از استقرار، حتی گامی فراتر از بسیاری از این نهادهای محسوب می‌شود. بنابراین، از منظر «چه چیزی باید سنجیده شود»، چهارچوب فعلی کافی و پیشرو است؛

ب) نقطه افتراق جدی عتف از نهادهایی مانند NSF، نه در طراحی، بلکه در ضعف زیرساخت داده‌ای و پراکندگی سامانه‌های اطلاعاتی است. NSF با بیش از ۲۰۰ شاخص و داشبوردهای تعاملی بی‌درنگ^۱ در «داده‌محوری و هوشمندی پیش‌بینی‌کننده» عمل می‌کند، درحالی‌که عتف هنوز با چالش دسترسی به داده‌های به‌موقع و باکیفیت مواجه است. این شکاف نشان می‌دهد که «چگونه داده‌ها گردآوری و تحلیل می‌شوند»، نه «چند شاخص داریم»، گلوگاه اصلی تحول است؛

ج) چهارچوب دوره سوم عتف برای تثبیت جایگاه خود و گام نهادن در مسیر نهادهای پیشرو، نیازمند تغییر تمرکز از «تکمیل طراحی نظری» به «ایجاد زیرساخت اجرایی یکپارچه» است. جهت حرکت آینده باید به‌سمت سه اقدام راهبردی معطوف شود (۱) استقرار یک سامانه ملی یکپارچه داده‌های STI با استانداردهای مشترک و سازوکار به‌اشتراک‌گذاری الزام‌آور؛ (۲) گذار از «آزمون بتا» به‌عنوان یک مرحله آزمایشی به حلقه بازخورد مستمر و نهادینه‌شده در چرخه سیاست‌گذاری؛ و (۳) توسعه داشبوردهای تعاملی و تحلیل‌گر (همسو با تجربه NSF) برای تبدیل داده‌های خام به بینش‌های سیاستی قابل استفاده.

تفسیر توزیع انطباق: اولویت‌دهی راهبردی به سنجش اثرگذاری

تحلیل عمیق‌تر، توزیع نامتوازن اما معنادار انطباق با استانداردهای جهانی در چارک‌های مختلف را آشکار می‌سازد: ۹۷ درصد در چارک کارایی، ۶۷ درصد در چارک اثربخشی و ۹۲ درصد در چارک اثرگذاری (سودمندی). این شکاف قابل توجه، حاکی از یک انتخاب آگاهانه و راهبردی در طراحی چهارچوب است. به نظر می‌رسد تمرکز اصلی طراحان، بر استانداردسازی و همسوسازی شاخص‌هایی قرار گرفته که آثار نهایی علم و فناوری را می‌سنجند. این جهت‌گیری، با پارادایم نوین جهانی که بر اندازه‌گیری «ارزش‌آفرینی»، «تأثیر اجتماعی»

اقتصادی» و پاسخگویی در قبال جامعه تأکید دارد، همسو است. این انتخاب را می‌توان پاسخی هدفمند به ضعف تاریخی نظام نوآوری ایران در تبدیل دانش به اثر ملموس تفسیر کرد. با این حال، انطباق نسبتاً پایین‌تر در حوزه اثربخشی، ممکن است نشان‌دهنده دو امکان باشد: نخست، چالش ذاتی در تطبیق کامل شاخص‌های میان‌مدت (مانند بهره‌وری عوامل تولید) با استانداردهای جهانی که گاه مستلزم تنظیمات بومی‌اند؛ و دوم، اولویت‌دهی کمتر آگاهانه به این حوزه در فرایند طراحی، با این فرض که زیرساخت نسبتاً بهتری برای سنجش آن‌ها وجود دارد یا تمرکز بر اثرات نهایی از اهمیت سیاستی بیشتری برخوردار است. این توزیع، یک نقطه کانونی برای بازنگری آینده (تعادل بخشی به شاخص‌های اثربخشی) را نشان می‌دهد.

چارک تناسب: مغفول‌مانده در اعتبارسنجی تجربی

یکی از خلأهای اساسی این پژوهش — که ناشی از بی‌توجهی به آن در طراحی چهارچوب پیشنهادی بوده است — تحلیل نکردن چارک تناسب^۲ است. براساس مدل چهارچارک در ارزیابی علم و فناوری، چارک تناسب به سنجش این پرسش اختصاص دارد که «آیا اهداف و اولویت‌های علم و فناوری کشور به‌درستی انتخاب شده‌اند و با نیازهای واقعی جامعه، صنعت و حاکمیت همسو هستند؟»

آنچه باعث شده است چارک تناسب در چهارچوب پیشنهادی ارائه نشود، عمدتاً ماهیت کیفی و قضاوتی شاخص‌های این چارک است؛ برخلاف سه چارک دیگر که شاخص‌های آن‌ها عمدتاً کمی و مبتنی بر داده‌های ثانویه هستند (مانند تعداد مقالات، نرخ رشد بهره‌وری، شاخص کیفیت زندگی)، شاخص‌های چارک تناسب ذاتاً کیفی، قضاوتی و نیازمند فرایندهای خبرگی ساختاریافته هستند. همچنین به‌دلیل وابستگی این شاخص‌ها به اسناد فرادستی، سنجش تناسب مستلزم تحلیل میزان همسویی شاخص‌ها با اهداف مندرج در این اسناد است که بعضاً هدف‌گذاری نیز نشده است. این تحلیل فراتر از داده‌های کمی صرف بوده و مستلزم روش‌شناسی ترکیبی متفاوتی است. ضمن اینکه ماهیت «هنجاری» شاخص‌های تناسب در مقابل ماهیت اغلب «توصیفی-تحلیلی» شاخص‌های کارایی، اثربخشی و اثرگذاری، پاسخ به این پرسش را به همراه دارد که «آیا اهداف درستی انتخاب شده‌اند؟» و نمی‌توان صرفاً با داده‌های کمی و روش‌های رایج علم‌سنجی به آن پاسخ داد و ورود خبرگان و ذی‌نفعان مهم نظام STI ضرورت دارد.

بنابراین، اگرچه این پژوهش به‌دلیل محدودیت‌های روش‌شناختی و داده‌ای نتوانست به تحلیل تجربی چارک تناسب بپردازد، اما تأکید می‌شود که این چارک از حیث سیاستی و حکمرانی، شاید مهم‌ترین

چارک در میان چهار چارک باشد. پرسش از «درستی انتخابها» بر پرسش از «کارایی در اجرا» و حتی «اثرگذاری نهایی» تقدم منطقی دارد. چهارچوبی که فقط به «چگونگی انجام کارها» بپردازد و از «چرایی انتخاب آنها» غافل بماند، نمی‌تواند به یک نظام هوشمند پایش و ارزیابی تبدیل شود. بنابراین، تکمیل اعتبارسنجی چهارچوب دوره سوم، مستلزم طراحی و اجرای پژوهشی مستقل و مکمل درباره چارک تناسب است.

عملکرد ضعیف نظام در مقابل تأکید طراحی؛ آشکارسازی تناقض ساختاری

بهرغم تأکید بالای چهارچوب بر سنجش اثرگذاری (با ۹۲ درصد انطباق)، یافته‌ها نشان می‌دهد که عملکرد نظام در این حوزه ضعیف است (میانگین CAGR پایین و روند نزولی یا ثابت شاخص‌هایی مانند کیفیت زندگی، رقابت‌پذیری صنعتی، امنیت آبی). این تناقض، یک شکاف عمیق و ساختاری را آشکار می‌سازد: طراحی چهارچوب به‌درستی بر «چه چیزی که باید سنجیده شود» (سنجش اثرات) متمرکز است، اما نظام نوآوری کشور از نظر «توانایی تولید آن نتایج» دچار ضعف مفرط است. این یافته به‌وضوح نشان می‌دهد که بهبود شاخص‌های اثرگذاری، مستلزم تحول در خود نظام حکمرانی، مشوق‌ها و نهادهای واسطه است، نه صرفاً اصلاح ابزارهای اندازه‌گیری. چهارچوب در حال نقشه‌برداری از قلمروی (اثرگذاری) است که اقتصاد و جامعه ایران هنوز نتوانسته‌اند به‌طور ملموس آن را تصرف کنند.

چالش حکمرانی داده: مغفول‌مانده در سایه انطباق بالا

با وجود انطباق بالای شاخص‌های اثرگذاری با استانداردهای جهانی (۹۲ درصد)، سطح پایین «دسترسی، تحلیل‌پذیری و به‌روز بودن» (DAB) برای این شاخص‌ها و شاخص‌های اثربخشی (مطابق با جدول ۹)، تعارض جدیدی را نشان می‌دهد: «شاخص‌های درستی برای سنجش وجود دارد، اما داده‌ای برای پر کردن آنها وجود ندارد». این امر حاکی از آن است که بزرگ‌ترین مانع، دیگر طراحی نظری نیست، بلکه ضعف در زیرساخت و حکمرانی داده‌های ملی است. پراکندگی، محرمانگی، ناهمگونی و تأخیر در داده‌های مربوط به پیامدهای اجتماعی-اقتصادی پژوهش و فناوری، هرگونه تلاش برای پایش هوشمند را با چالش مواجه می‌کند. موفقیت چهارچوب منوط به پذیرش این واقعیت است که «حکمرانی داده» پیش‌نیاز «حکمرانی مبتنی بر شواهد» است.

آزمون بتا: پل عبور از شکاف بین طراحی ایستا و اجرای پویا

ادغام «آزمون بتا» در هسته چهارچوب، نوآوری تعیین‌کننده برای پر کردن شکاف میان طراحی مبتنی بر استانداردها و واقعیت‌های اجرایی

پیچیده ایران است. این سازوکار، منطق علم طراحی را وارد عرصه سیاست‌گذاری می‌کند و چهارچوب را از یک مجموعه شاخص ایستا به سامانه یادگیری و تطبیق پویایی تبدیل می‌نماید. در شرایطی که داده‌های پایه ناقص‌اند، آزمون بتا این فرصت حیاتی را ایجاد می‌کند تا موانع واقعی گردآوری، پردازش و تفسیر داده برای شاخص‌های باله‌میت اما پیچیده (به‌ویژه در چارک‌های اثربخشی و سودمندی)، در مقیاسی کوچک و کنترل‌شده شناسایی و راه‌حل‌های عملی برای آنها یافت شود. این چرخه تکرارشونده، تنها راه برای تبدیل چهارچوبی با انطباق نظری بالا به یک ابزار عملیاتی مؤثر است.

کاهش شاخص‌ها: همگرایی با جریان جهانی و پاسخ به محدودیت داخلی

کاهش هدفمند شاخص‌ها از ۱۱۲ مورد در دوره دوم به ۴۰ شاخص اصلی در دوره سوم، حرکتی دوجبه‌ای و هوشمندانه است. از یک سو، این کاهش با جریان جهانی به‌سمت «چهارچوب‌های متمرکز و اثرمحور» همسو است و امکان تمرکز منابع مدیریتی و تحلیلی را فراهم می‌آورد. از سوی دیگر، این امر می‌تواند پاسخی واقع‌بینانه به محدودیت شدید ظرفیت گردآوری و پردازش داده در کشور باشد. تمرکز بر ۴۰ شاخص، شانس تولید داده‌هایی با کیفیت و قابلیت اطمینان بالاتر را افزایش می‌دهد. کلید موفقیت این کاهش در آن است که سبد ۴۰ شاخص نهایی، پس از آزمون بتا، به‌طور متوازی هم عوامل محرک (کارایی) و هم نتایج نهایی (اثربخشی و سودمندی) نظام نوآوری را دربرگیرد و از تمرکز صرف بر یکی از این دو جنبه اجتناب شود.

نتیجه‌گیری و راهکارهای سیاستی

این پژوهش با هدف اعتبارسنجی تجربی چهارچوب دوره سوم پایش STI شورای عالی عتف و براساس تحلیل داده‌های واقعی دوره برنامه ششم پیشرفت (۱۳۹۶-۱۴۰۰) انجام شد. یافته‌ها نشان‌دهنده وضعیتی دوگانه است؛ از یک سو، چهارچوب از اعتبار ساختاری و انطباق نظری مطلوبی برخوردار است (میانگین انطباق ۸۵ درصد با استانداردهای بین‌المللی، با تمرکز قابل توجه بر حوزه اثرگذاری که ۹۲ درصد انطباق را نشان می‌دهد). از سوی دیگر، عملکرد واقعی نظام در چارک‌های سه‌گانه با میانگین وزنی رشد مرکب حدود ۱/۷ درصد در سطح «ضعیف» ارزیابی شد. نکته ظریف‌تر آنکه، نرخ رشد در چارک‌های اثربخشی (۱/۸ درصد) و اثرگذاری (۱/۷ درصد) اندکی از چارک کارایی (۱/۵ درصد) فراتر رفته، که می‌تواند حاکی از تأخیر در ظهور آثار یا تغییر آرام اولویت‌ها باشد. افزون‌براین، چالش‌های بنیادین اجرا — شامل ضعف جدی در دسترسی، تحلیل‌پذیری و به‌روز بودن داده‌ها (نمره DAB پایین)، پراکندگی سامانه‌های اطلاعاتی و نبود

و ایجاد پیوندی قانون‌مند و شفاف میان یافته‌های پایش و چرخه برنامه‌ریزی، بودجه‌ریزی و ارزیابی نهادی، این چهارچوب پیشرفته و رشد نسبی مشاهده‌شده در چارک پیامدها و آثار نیز در معرض توقف یا بازگشت قرار دارد.

بر این اساس پیشنهادهای این پژوهش در افق کوتاه‌مدت ناظر بر آن است که موفقیت چهارچوب دوره سوم مستلزم تبدیل آن از یک سند نظری به ابزاری یادگیرنده و مؤثر برای تصمیم‌گیری است. این امر از طریق سه اقدام محوری و به‌هم‌پیوسته می‌تواند محقق شود: نخست، اجرای عملیاتی و تحلیلی «آزمون بتا» با تمرکز بر رفع موانع دسترسی به داده‌های چارک اثرگذاری و اثربخشی، به‌گونه‌ای که این مرحله به فرایند واقعی یادگیری و بازطراحی تبدیل شود. دوم، تدوین و تصویب «سند حکمرانی داده‌های ملی علم و فناوری» به‌عنوان پیش‌نیازی اجتناب‌ناپذیر، که با تعریف تکالیف قانونی، استانداردهای یکپارچه و سازوکارهای شفاف برای به‌اشتراک‌گذاری داده میان تمامی نهادهای ذی‌ربط (از مرکز آمار و بانک مرکزی تا وزارتخانه‌های تخصصی و دانشگاه‌ها)، زیرساخت اطلاعاتی لازم برای چهارچوب را فراهم آورد. سوم، ایجاد پیوند عملی و شفاف میان نتایج پایش و سازوکارهای تخصیص منابع ارزیابی نهادها، حتی به‌صورت آزمایشی در مرحله مقدماتی، تا ارزش عملیاتی چهارچوب برای ذی‌نفعان اثبات شود و انگیزه لازم برای بهبود کیفیت داده‌ها و گزارش‌دهی فراهم آید. بدون این اصلاحات نهادی و فرایندی موازی، خطر آن وجود دارد که این چهارچوب نیز نتواند بر چرخه معیوب «طراحی خوب، اجرای ضعیف و تأثیر ناچیز» غلبه کند.

یکپارچگی نهادی در فرایند گردآوری و پردازش داده — به‌عنوان موانع اصلی پیش روی استقرار مؤثر شناسایی شدند. بنابراین، می‌توان جمع‌بندی کرد که چهارچوب دوره سوم به‌عنوان یک «مصنوع طراحی» پیشرفته و هم‌تراز با جریان جهانی تکمیل شده است، اما «بستر نهادی، داده‌ای و اجرایی» کشور هنوز برای بهره‌برداری کامل از آن آماده نیست و شتاب کنونی تغییر در نظام برای ایجاد تحول ناکافی است.

بنا بر آنچه گفته شد، نتیجه‌گیری نهایی این پژوهش دوجبهی است: نخست تأیید ظرفیت الگوسازی چهارچوب دوره سوم؛ با تکیه بر پارادایم علم طراحی، الگوی ارزیابی عملکرد یکپارچه و رویکرد زنجیره ارزش، به نظر می‌رسد این چهارچوب ظرفیت تبدیل شدن به ستون اصلی یک نظام پایش ملی هوشمند، یادگیرنده و مبتنی بر شواهد را دارد. سطح انطباق و ساختار آن نشان می‌دهد این چهارچوب می‌تواند الگوی مناسبی برای دیگر کشورهای در حال توسعه با چالش‌های مشابه در حکمرانی پیچیده STI باشد. دوم مشروط بودن موفقیت به تحول در حکمرانی داده و نهادهای اجرایی؛ موفقیت عملیاتی و تأثیرگذاری سیاستی این چهارچوب کاملاً مشروط و وابسته به اجرای هم‌زمان تحولات نهادی موازی است. نرخ رشد اندک، اما امیدوارکننده در حوزه‌های اثربخشی و اثرگذاری در مقایسه با کارایی (۱/۸ و ۱/۷ درصد در مقابل ۱/۵ درصد) نشان می‌دهد که نظام ظرفیت حرکت به‌سمت پیامدهای بلندمدت‌تر را دارد، اما این حرکت مستلزم شتابدهی قاطع از طریق اصلاح زیرساخت داده و نهادها است. بدون سرمایه‌گذاری جدی و فوری برای ایجاد زیرساخت یکپارچه حکمرانی داده‌های ملی STI، تقویت نهادهای واسط و سازوکارهای همکاری،

References

- Abbasi, M., & Ashrafi, M. (2011). *Development of a framework for evaluating the performance of research credit in the country*. *Rahyaf*, 21(48), 5-17. (Persian)
- Acciai, C. (2023). Research and innovation policy design in France. In C. Acciai (Ed.), *Policy design for research and innovation: Politics, institutions and interest intermediation practices* (pp. 105-152). Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-031-36628-4>
- Acharya, P., & Rahman, A. (2021). Compound Annual Growth Rate (CAGR) of select financial variables in Indian four-wheeler automobile companies. *Wesleyan Journal of Research*, 14(15).
- Aghaei, P. (2021). *Amendment of the executive regulations of the science, technology and innovation monitoring and evaluation system*. Tehran: Research Project Report, National Research Institution for Science Policy. (Persian)
- Bagheri Moghaddam, N., & Mosleh, E. (2016). Fostering sustainable technologies: A framework for analyzing the governance of innovation systems. *The Journal of Science and Technology Policy Letters*, 6(2), 117-139. (Persian)
- Cho, H., Ahn, H., & Park, E. (2024). Data-driven analysis on the performance evaluation of national R&D projects in Korea. *Evaluation and Program Planning*, 102, 102383. <https://doi.org/10.1016/j.evalprogplan.2023.102383>
- Contandriopoulos, D., & Brousselle, A. (2012). Evaluation models and evaluation use. *Evaluation*, 18(1), 61-77. <https://doi.org/10.1177/1356389011430371>

- Corporate Finance Institute (CFI). (2023). Compound Annual Growth Rate (CAGR). Retrieved from: <https://corporatefinanceinstitute.com/resources/wealth-management/compound-growth-rate/>
- Çubuk, M. (2023). R&D and innovation map of Turkey: Hybrid model approach. *Turkish Journal of Science and Technology*, 18(2), 487-502. <https://doi.org/10.55525/tjst.1340408>
- Eurostat. (2018). *The measurement of scientific, technological and innovation activities Oslo manual 2018 guidelines for collecting, reporting and using data on innovation: Guidelines for collecting, reporting and using data on innovation*. Paris: OECD publishing.
- Falkenheim, J. C., & Alexander, J. M. (2023). Academic research and development. Science & engineering indicators 2024. NSB-2023-26. National Science Foundation. Retrieved from: <https://ncses.nsf.gov/pubs/nsb202326/assets/nsb202326.pdf>
- Farazkish, M. (2016). Designing an evaluation model for science, technology and innovation process of Iranian governmental organizations [PhD dissertation]. Tarbiat Modares University, Tehran. (Persian)
- Gemici, E., & Gemici, Z. (2021). A comparative study on Turkey's Science and Technology (S&T) indicators. *Economics and Business Quarterly Reviews*, 4(3), 126-143. <https://doi.org/10.31014/aior.1992.04.03.376>
- Ghazinoory, R., Divsalar, A., & Ghazinoory, S. (2012). *Evaluation of national R&D projects: Structures and methods*. Tehran: Institute for International Energy Studies Press. (Persian)
- Ghazinoory, S. & Farazkish, M. (2018). A modal for STI national evaluation-based efficiency, effectiveness and Utility index. *Strategic Studies of Public Policy*, 8(27), 205-229. (Persian)
- Ghazinoory, S., Farazkish, M., Montazer, G. A., & Soltani, B. (2017). Designing a national science and technology evaluation system based on a new typology of international practices. *Technological Forecasting and Social Change*, 122, 119-127. <https://doi.org/10.1016/j.techfore.2017.04.012>
- Ghazinoory, S., Farazkish, M., Nasri, S., & Mardani, A. (2023). Designing a Science, Technology, and Innovation (STI) evaluation dashboard: A comprehensive and multidimensional approach. *Technology Analysis & Strategic Management*, 35(8), 1005-1023. <https://doi.org/10.1080/09537325.2021.1990877>
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337-355. <https://doi.org/10.25300/MISQ/2013/37.2.01>
- Harris, M., & Albury, D. (2009). *Why radical innovation is needed to reinvent public services for the recession and beyond: The innovation imperative*. London: The Lab Discussion Paper, NESTA.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- Khayyatian, M. S., Fartash, K. & Poursagari, P. (2020). Development of a framework for monitoring and evaluation of Iran's national system of science, technology and innovation. *Strategy for Culture*, 13(49), 119-154. (Persian) <https://doi.org/10.22034/jsfc.2020.109868>
- Kosari, S., & Alizadeh, P. (2021). A comparative study of the governance of science, technology and innovation in Iran and selected countries. *Rahyaft*, 31(2), 1-22. (Persian) <https://doi.org/10.22034/rahyaft.2021.10737.1223>
- Markiewicz, A., & Patrick, I. (2015). *Developing monitoring and evaluation frameworks*. Thousand Oaks: Sage Publications.
- Montazer, G. A., Sharanji, M., Moradipor, H., & Farazkish, M. (2019). *Sanandaj manual: The national evaluation model for research institutions*. Tehran: University Publication Center. (Persian)
- OECD. (2015). *Frascati manual 2015: Guidelines for collecting and reporting data on research and experimental development*. Paris: OECD Publishing.
- OECD. (2021). *OECD science, technology and innovation outlook 2021: Times of crisis and opportunity*. Paris: OECD Publishing. <https://doi.org/10.1787/75f79015-en>
- Orozco, L. A., Ordóñez-Matamoros, G., García-Estévez, J., Sierra-González, J. H., & Bortagaray, I. (2021). Science, technology, and innovation governance for social inclusion and sustainable development in Latin America. In L. A. Orozco, G. Ordóñez-Matamoros, J. H. Sierra-González, J. García-Estévez, I. Bortagaray (Eds.), *Science, technology, and higher education: Governance approaches on social inclusion and sustainability in Latin America* (pp. 1-18). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-80720-7_1
- Partelow, S. (2023). What is a framework? Understanding their purpose, value, development and use. *Journal of*

- Environmental Studies and Sciences*, 13(3), 510-519.
<https://doi.org/10.1007/s13412-023-00833-w>
- Patton, M. (2017). Developmental evaluation. In J. Pokorski, Z. Popis, T. Wyszynska, & K. Hermann-Pawlowska (Eds.), *Theory-based evaluation in complex environments* (pp. 7-19). Warsaw: Polish Agency for Enterprise Development.
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45-77.
<https://doi.org/10.2753/MIS0742-1222240302>
- Pinar, M., & Horne, T. J. (2022). Assessing research excellence: evaluating the research excellence framework. *Research Evaluation*, 31(2), 173-187.
<https://doi.org/10.1093/reseval/rvab042>
- Rodriguez, V., & Soeparwata, A. (2015). The governance of science, technology and innovation in ASEAN and its member states. *Journal of the Knowledge Economy*, 6(2), 228-249.
<https://doi.org/10.1007/s13132-012-0111-x>
- Stam, C. D. (2007). Making sense of knowledge productivity: Beta testing the KP-enhancer. *Journal of Intellectual Capital*, 8(4), 628-640.
<https://doi.org/10.1108/14691930710830792>
- Supreme Council for Science, Research and Technology. (2021). *Executive regulations for the national science, technology, and innovation monitoring and evaluation system*. Retrieved from:
<https://www.atf.gov.ir/Content/media/digitalibrary/2021/10/book200/200.pdf> (Persian)
- Van Aken, J. E. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of Management Studies*, 41(2), 219-246.
<https://doi.org/10.1111/j.1467-6486.2004.00430.x>
- Vetterli, M., D'Amours, S., Bronstein, M., Daston, L., Díaz, S., Faist, J., ... & Stubbs, C. (2023). *Assessment report of the CNRS (Centre National de la Recherche Scientifique)* [Report]. High Council for Evaluation of Research and Higher Education. Retrieved from:
<https://art.torvergata.it/retrieve/3cdc5810-0bab-4dec-8235-3da93b138592/report-cnrs-2023.pdf>



مهدیه فراز کیش

او دانش‌آموخته دکتری سیاست‌گذاری علم و فناوری از دانشگاه تربیت مدرس در سال ۱۳۹۵ است. وی در حدود بیست سال فعالیت علمی به‌عنوان پژوهشگر، مقالات متعددی در مجلات و کنفرانس‌های علمی داخلی و خارجی منتشر کرده و در طرح‌های کاربردی متعددی نیز مشارکت داشته است. او هم‌اکنون به‌عنوان عضو هیئت علمی گروه پژوهشی ارزیابی سیاست‌ها و پایش علم، فناوری و نوآوری مؤسسه تحقیقات سیاست علمی کشور، فعالیت‌های پژوهشی خود را ادامه می‌دهد.

زمینه‌های تخصصی او عبارت‌اند از ارزیابی نظام‌های پژوهش و فناوری، ارزیابی نظام‌های یادگیری الکترونیکی، توسعه فناوری‌های نوظهور، راهبردهای توسعه فناوری و سیاست‌گذاری پژوهش و فناوری.

