Archive of SID

The Bible of STI Policy
Volume 11, Number 2, Summer 2019

Journal of

Science & Technology Policy

Selecting and Applying Science, Technology and Innovation Evaluation Indices

Mahdieh Farazkish[1]*, Nasrin Dastranj[2]

1- Postdoctoral Researcher of Science and Technology Policy, Tarbiat Modares University, Tehran, Iran
2- Assistant Professor of Research Institute for Information and Communication Technology, Tehran, Iran

Abstract

Recently, with the advent of knowledge-based economies, the importance of science, technology and innovation (STI) have grown in different countries development. Therefore, the performance evaluation of STI activities recognized as a necessity. In this paper, the STI evaluation concepts and models introduce and a new evaluation model has been proposed based on the integration of performance-based and key stakeholders-based evaluation approaches.

The pool of input, output, outcome and impact indicators have classified into five specialized fields including natural sciences, engineering sciences, human sciences, social sciences, and health sciences. Then, performance evaluation is feasible in terms of relevance, efficiency, effectiveness, and usefulness for the four main stakeholders including policymakers, researchers and technicians, industrial beneficiaries, and the public. The final results presented in the form of a radar chart which we showed one sample performance evaluation of universities and research institutes.

Keywords: Science, Technology and Innovation Evaluation; Evaluation Indicators; Performance Evaluation of Universities and Research Institutions

* Corresponding author: mfarazkish@gmail.com

www.SID.ir

Archive of SID

فصلنامه علمی-پژوهشی
سیاست علم و فناوری

ویژه‌نامه جامع سیاست علم، فناوری و نوآوری
سال یازدهم، شماره ۲، تابستان ۱۳۹۸

انتخاب و بکارگیری شاخص‌های ارزیابی علم، فناوری و نوآوری

مهدیه فرازکیش[1*]، نسرین دسترنج[2]

۱- پژوهشگر پسادکتری سیاست‌گذاری علم و فناوری، دانشگاه تربیت مدرس، تهران
۲- عضو هیأت‌علمی پژوهشگاه ارتباطات و فناوری اطلاعات، تهران

چکیده

طی دهه‌های اخیر، با ظهور اقتصادهای دانش‌بنیان، اهمیت علم، فناوری و نوآوری (STI) در توسعه کشورهای مختلف افزون شده و با توجه به ماهیت این حوزه، ارزیابی عملکرد و تخصیص بودجه در توسعه STI از طرف دولت‌ها ضروری تشخیص داده شده است. در این مقاله، مبانی و مدل‌های ارزیابی STI معرفی و بر اساس تلفیق مدل‌های ارزیابی عملکردمحور و ذینفعان کلیدی، مدل جدیدی برای ارزیابی STI به کار گرفته شده است. مجموعه شاخص‌های درون‌دادی، برون‌دادی، پیامدی و اثرگذاری مدل به تفکیک پنج حوزه تخصصی علوم (طبیعی، مهندسی، انسانی، اجتماعی و سلامت) تدوین و ارزیابی عملکرد برای چهار گروه ذینفعان اصلی مشتمل بر سیاست‌گذاران، پژوهشگران و فناوران، بهره‌برداران صنعتی و عموم مردم، به ترتیب در قالب سنجش تناسب، کارایی، اثربخشی و سودمندی، امکان‌پذیر می‌شود. نتایج نهایی نیز به صورت نمودار راداری در سطح ارزیابی مورد نظر نمایش داده می‌شود که در بخش پایانی مقاله نمونه‌ای از آن برای ارزیابی عملکرد پژوهشی دانشگاه‌ها و پژوهشگاه‌ها ارائه شده است.

کلیدواژه‌ها: ارزیابی علم، فناوری و نوآوری (STI)؛ شاخص‌های ارزیابی STI؛ ارزیابی عملکرد دانشگاه‌ها و پژوهشگاه‌ها

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

Farazkish, M., & Dastranj, N. (2019). **Selecting and Applying Science, Technology and Innovation Evaluation Indices.** *Journal of Science & Technology Policy*, 11(2), 579-598. {In Persian}.

۱- مقدمه

از اولین سال انتشار مجلات علمی دنیا، ارزیابی بخش جدایی‌ناپذیری از نظام علمی، به ویژه به شیوه ارزیابی همتایان[1] بوده و از آن در طیف وسیعی از فرآیندهای انتخاب، از جمله: تقاضا برای رتبه دانشگاهی، انتشارات علمی، استفاده از کمک‌هزینه‌های پژوهشی، جوایز و درجات علمی، استفاده شده؛ اما در طول ٤٠ سال گذشته، انواع جدیدی از ارزیابی‌های علمی در تکمیل ارزیابی همتایان، هم از نظر مقیاس و هم از نظر دامنه، توسعه یافته است. این توسعه را می‌توان به عنوان نتیجه درک جدیدی از تعامل علم و جامعه (تغییر از سبک ۱ تولید دانش به سبک ۲ آن[2]) در میان سایر عوامل، در نظر گرفت. هنگامی که پژوهش به صورت کاربردی (مسئله‌محور) و در بافتار الزامات جامعه انجام می‌شود، انتظار می‌رود اهداف دوگانه‌ای (هم علمی و هم عملی) تحقق یابد و طبعاً زمانی که این پژوهش توسط

* نویسنده عهده‌دار مکاتبات: mfarazkish@gmail.com

1- Peer Review

2- Knowledge mode1 to mode 2
شکل جدیدی از تولید دانش از اواسط قرن ۲۰ در حال ظهور است که بافتارگرا، متمرکز بر حل مسئله و میان رشته‌ای است. این نوع دانش شامل تیم‌های چندرشته‌ای است که برای مدت زمان کوتاهی بر حل مشکلات خاصی از دنیای واقعی متمرکز می‌شوند.گیبونز و همکارانش این نوع تولید دانش را "سبک ۲" نامیده‌اند. او و همکارانش تحقیقات سنتی، که آنها برچسب "سبک ۱"را به آن می‌زنند، را تحقیقاتی دانشگاهی می‌دانند که محقق آن را آغاز می‌کند و رشته‌محور است. بنابراین سبک (۱) تولید دانش محقق‌محور و مبتنی بر رشته خاصی است، در حالی که سبک (۲) متمرکز بر حل مسئله و میان‌رشته‌ای است [۲].

www.SID.ir

Archive of SID

نهادهای مختلف انجام می‌گیرد و تنها در نهاد دانشگاهی متمرکز نیست، باید نوع جدیدی از ارزیابی نیز به کار گرفته شود. بر این اساس می‌توان رویکردهای ارزیابی STI را در دوره‌های قبل و بعد از دهه میلادی ۱۹۸۰ مجزا دانست.

در واقع پیش از دهه ۱۹۸۰، ارزیابی پژوهش کلاسیک وجود داشت که تمرکز آن روی ارزیابی تقاضا برای جایگاه‌های شغلی، مقاله و پایان‌نامه‌ها و درخواست‌های حمایت از پروژه‌ها و نتیجه این ارزیابی‌ها تصمیم‌گیری در مورد استخدام، رتبه‌بندی، انتشار و تأمین مالی در زیست‌بوم علمی بود و لذا صرفاً اهداف علمی-پژوهشی دنبال می‌شد؛ اما از دهه ۱۹۸۰ به بعد، ارزیابی STI در سطوح دیگر گسترش یافته و در ارزیابی‌گروه‌های پژوهش و فناوری، نهادها، حوزه‌ها و برنامه‌های پژوهشی و فناورانه، دانشگاه‌ها و نظام‌های مشاوره STI به کار گرفته شد. از سوی دیگر ارزیابی STI در نظام‌های مالی برای توزیع منابع و در رتبه‌بندی جهانی دانشگاه‌ها، و نیز به عنوان یک ابزار مدیریتی برای پایش، یادگیری و مستندسازی نتایج و قابل اعتماد بودن[1] آنها کاربرد داشته است؛ لذا در این دوران خروجی‌ها، پیامدها و تأثیرات علمی-پژوهشی و به دنبال آن توسعه فناوری و حصول نوآوری، به عنوان عوامل اثرگذار بر جامعه و اقتصاد مورد ردیابی قرار گرفته‌اند [۱].

در شکل ۱ سیر تکامل روند ارزیابی STI، از ارزیابی همتایان کلاسیک تا مفهوم اجتماعی گسترده‌تر "آگورا[2]"، تحت عنوان «پیوستار پیچیدگی ارزیابی STI»، نشان داده شده است.

با وجود تکامل رویکردهای ارزیابی STI، همواره این سؤالات مطرح بوده که: «چگونه می‌توان ارزیابی STI را برای دستیابی به خواسته‌ها و موقعیت‌های متناقض، توسعه داد و کیفیت و کمیت هم‌زمان را در ارزیابی‌ها تأمین نمود؟»

به طور مشخص، در پاسخ به پرسش اول رویکردهای یکپارچه با در نظر گرفتن ذینفعان کلیدی حوزه مورد ارزیابی، و در پاسخ به پرسش دوم، پیاده‌سازی هم‌زمان روش‌های ارزیابی کمی و کیفی، در مطالعات ارزیابی STI پیشنهاد شده است. این مهم را می‌توان در گونه‌شناسی‌های مختلف نظام‌های ارزیابی STI به وضوح مشاهده کرد [۳]. اگر چه در نظر گرفتن هر دو پاسخ مطرح شده در طراحی یک نظام ارزیابی STI واحد، متضمن دیدگاه جدیدی است.

با بهره‌گیری از این دیدگاه جدید و چندبعدی، در این مقاله ابتدا به چارچوب مفهومی ارزیابی علم، فناوری و نوآوری پرداخته شده و سپس مدل‌های ارزیابی STI و معیارهای انتخاب مدل مناسب معرفی می‌شود. در بخش سوم مقاله، ارزیابی عملکرد و مدل جدیدی در ارزیابی STI تبیین گردیده است. نحوه انتخاب و بکارگیری شاخص‌های ارزیابی در این مدل در بخش بعدی به صورت مشروح مورد بحث قرار گرفته و در بخش پایانی، مطالعه موردی ارزیابی عملکرد STI در دانشگاه‌ها و پژوهشگاه‌های ایران ارائه شده است.

۲- چارچوب مفهومی ارزیابی علم، فناوری و نوآوری

پایش[3]، ارزیابی و مقایسه شاخص‌های مرتبط با علم، فناوری و نوآوری، اکنون یکی از فعالیت‌های معمول در نظام‌های ملی و بین‌المللی و معیاری برای اندازه‌گیری توسعه‌یافتگی، رقابت‌پذیری و نوآوری شناخته می‌شود. با این وجود شناسایی و طبقه‌بندی فعالیت‌های زیرمجموعه STI و نتایج این فعالیت‌ها موضوعی محل چالش است. بدین منظور، سازمان همکاری‌های اقتصادی و توسعه (OECD) فعالیت‌های علم و فناوری[4] (STA) را مشتمل بر «تمامی فعالیت‌های نظام‌مندی که با هدف تولید، توسعه، انتشار و کاربرد دانش علمی و فنّاورانه، انجام شده و در تمامی حوزه‌های علم و فناوری قابل استفاده است»، می‌داند [٤]. بر مبنای این تعریف، یونسکو مجموعه فعالیت‌های علم و فناوری را به پنج گروه تقسیم کرده است [۵]:

1- Accountability

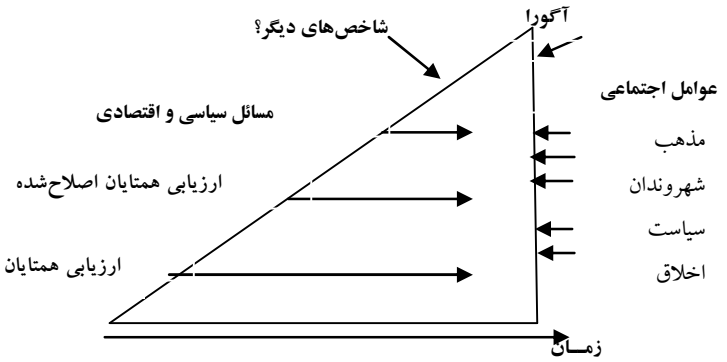
۲- مفهوم «آگورا» از تاریخ کلاسیک یونان نشأت گرفته که بیانگر میدان شهر بود که در آن، افراد در قالب ملاقات با یکدیگر، نقطه نظرات خود را تبادل می‌کردند. این مفهوم اولین‌بار در اوایل قرن ۲۱ به منظور بحث در خصوص تغییرات ارتباط علم و جامعه مطرح شد. در حقیقت، "آگورا" یک مکان فیزیکی یا یک سازمان جدید سیاستی ویژه نیست، بلکه یک عنوان معمول برای افراد، شبکه‌ها و مؤسسات علاقمند یا تأثیرگذار در تنازعات بر سر تصدیق ادعاهای دانش علمی، می‌باشد. این بدان معنا است که "آگورا" بسیاری از نهادهای شناخته شده علم و فناوری همچون مجلات، کنفرانس‌های علمی، دانشگاه‌ها، انجمن‌های تحقیقاتی، صنعت، دستگاه‌های وزارتی، پارلمان‌ها، رسانه‌ها، عقائد عمومی، جریانهای اجتماعی و ... را شامل می‌شود. در واقع، این مفهوم بر اهمیت روابط، فرآیندها و تعاملات میان بازیگران مختلف در عرصه پژوهش و فناوری تأکید دارد؛ پدیده‌ای که نویسندگان آن را "بافتارگرایی" نیز نامیده‌اند [۲].

3- Monitoring

4- Science and Technology Activities

www.SID.ir

Archive of SID



شکل ۱) تحول رویکردهای ارزیابی STI در طی زمان [۲]

الف) آموزش و مهارت‌آموزی علمی و فنی[1]

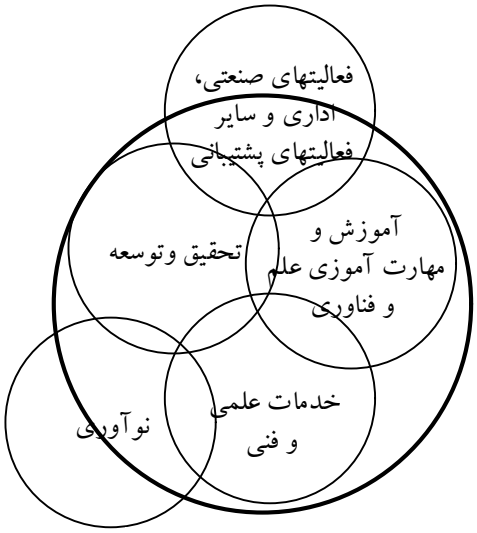
ب) پژوهش و توسعه آزمایش‌محور[2]

ج) نوآوری و سایر فعالیت‌های صنعتی[3]

د) خدمات علمی و فناورانه[4]

هـ) فعالیت‌های اداری و سایر فعالیت‌های پشتیبان[5]

طبعاً برای ارزیابی این طیف فعالیت‌ها نیاز به تنوعی از شاخص‌ها مطابق شکل ۲ نمایش خواهیم داشت.



شکل ۲) چارچوب شاخص‌های گونه‌شناسی فعالیت‌های علم و فناوری [۵]

مطابق با این گونه‌شناسی، فعالیت‌های علم و فناوری طیف وسیعی را در بر دارد. بنابراین اولین بُعد چارچوب مفهومی ارزیابی STI، تعیین حوزه فعالیت‌های مورد ارزیابی است که به تناسب آن انتخاب شاخص‌های مناسب برای ارزیابی به تفکیک نوع فعالیت زیرمجموعه آن، امکان‌پذیر خواهد بود. به عنوان مثال، در ارزیابی عملکرد پژوهشی یک دانشگاه، عمدتاً فعالیت‌های پژوهش و توسعه آن (گروه دوم) ارزیابی می‌شود و فعالیت‌های آموزشی (گروه اول) مدنظر نیست.

دومین بُعد چارچوب مفهومی ارزیابی STI، تعیین سطح ارزیابی است. یک یا چند گروه فعالیت‌های علم و فناوری می‌تواند در یکی از سطوح بین‌المللی، ملی، دستگاهی، منطقه‌ای یا سازمانی ارزیابی شود. در شکل ۳ سطوح مختلف ارزیابی فعالیت‌های علم و فناوری نشان داده شده است.

هر چه از ارزیابی سطوح کلان‌تر به سطوح خُردتر حرکت کنیم، شاخص‌های ارزیابی فعالیت‌های علم و فناوری باید متمرکز بر «استانداردسازی عملکرد» معطوف شود و به عکس، هر چه سطح ارزیابی به سمت کلان‌تر پیش رود، شاخص‌های «سطح مشترک و مورد توافق بین سطوح پائین‌تر» می‌باید مدنظر قرار گیرد [٦]. به عنوان مثال، ارزیابی فعالیت‌های پژوهشی دانشگاه‌ها در سطح سازمانی (خرد) معنا پیدا می‌کند، لذا معطوف به ارزیابی استانداردهای عملکرد دانشگاه در حوزه فعالیت‌های پژوهش و توسعه است.

سومین بُعد چارچوب مفهومی ارزیابی STI، تعیین کارکرد ارزیابی است. گالارت[6] [۷] سه کارکرد اصلی از ارزیابی STI STI را به شرح ذیل معرفی نموده است:

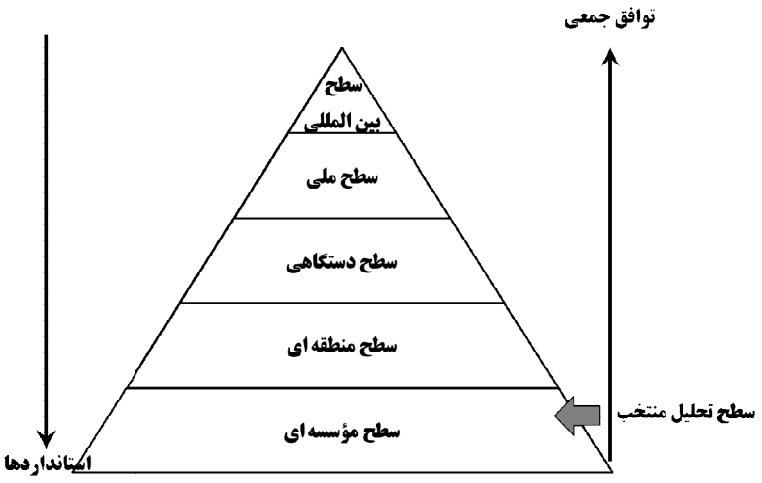
- کارکرد کنترلی که به بررسی دقیق نحوه استفاده از منابع برای انجام فعالیت‌های علم و فناوری می‌پردازد؛
- کارکرد توزیعی که با هدف تخصیص منابع میان بازیگران بالقوه و ذینفعان مختلف فعالیت‌های علم و فناوری انجام می‌شود؛

1- Scientific and technical education and training (STET)
2- Research and experimental Development (R&D)
3- Innovation and other industrial activities (I&I)
4- Scientific and technological services (STS)
5- Administration and other supporting activities (A&S)

6- Gallart

www.SID.ir

Archive of SID



شکل ۳) سطوح ارزیابی فعالیت‌های علم و فناوری [٦]

– کارکرد بهبودی که به دنبال یادگیری از تجارب گذشته در جهت پیدا کردن بهترین شیوه توسعه فعالیت‌های علم و فناوری و تکرار آن است.

به عنوان مثال، ارزیابی عملکرد پژوهشی دانشگاه‌ها در انگلیس مبتنی بر کارکرد بهبودی است [۸و۹]، در حالی که در اسپانیا عموماً از نوع کنترل عملکرد [۱۰]، و در سوئد بر اساس تخصیص بودجه به آنها شکل گرفته است [۱۱].

مشخص کردن گزینه مورد نظر ما در هریک از این ابعاد پیش‌نیاز طراحی مدل‌های ارزیابی STI است که در بخش بعدی مورد بحث قرار می‌گیرد.

۳- مدل‌های ارزیابی علم، فناوری و نوآوری

ارزیابی[1]، فرآیند تعیین، تهیه و گردآوری داده‌ها و اطلاعات به منظور قضاوت و تصمیم‌گیری در جهت بهبود مداوم است [۱۲]. در تعریفی دیگر، ارزیابی به عنوان «سنجش نظام‌مند شایستگی و ارزش سازمان، خروجی و پیامد مداخلات سیاست‌های عمومی، مطرح است که این سنجش در شرایط آینده فعالیت‌های اجرایی نقش خواهد داشت» [۱۳]. با توجه به این تعاریف، می‌توان ارزیابی STI را به عنوان مفهومی ترکیبی به دو صورت تعریف کرد: «سنجش نظام‌مند فعالیت‌های علم و فناوری در تمامی اشکال آن» و «سنجش نظام‌مند شایستگی‌ها و ارزش‌های سازمان، خروجی و پیامد سیاست‌های حوزه STI» [١٤].

مدل‌های مختلفی برای ارزیابی وجود دارد (جدول ۱) که هر یک از این مدل‌ها، از جنبه خاصی به موضوع ارزیابی می‌نگرد و دارای مزایا و محدودیت‌های خاص خود هستند.اگر چه وجود مدل‌های متعدد رقیب از دیدگاه نظری و تحلیلی جالب توجه است، اما از موضع عملی می‌تواند گیج‌کننده به نظر برسد. چگونه باید ارزیابان از میان مدل‌های ارزیابی موجود در جعبه ابزار، یکی را انتخاب کنند؟ در متون نظری، سه نوع منطق مختلف برای انتخاب مدل مناسب ارزیابی پیشنهاد شده است: اول این که طراحی مدل ارزیابی با توجه به هدف ارزیابی انجام شود؛ دوم، بر مبنای ویژگی‌های ارزیابی شونده[2] صورت پذیرد؛ و سوم، با توجه به ویژگی‌های مسئله‌ای که سیاست، برنامه و یا سازمان تحت ارزیابی به دنبال حل و فصل آن است، صورت پذیرد.

۳-۱ توجه به هدف ارزیابی

پژوهشگران معتقدند طرح ارزیابی[3] باید متناسب با هدف ارزیابی انتخاب شود. در ادبیات ارزیابی برنامه [١٦] توصیه شده که اگر ارزیابی برای ایجاد یادگیری و در نتیجه بهبود برنامه انجام می‌شود، ارزیابی تکوینی[4] مناسب است. در مقابل، مقابل، ارزیابی تجمیعی[5] زمانی توصیه می‌شود که ارزیابی به منظور کنترل عملکرد و پاسخگویی، هدایت می‌شود. همانطور که در جدول ۲ نیز مشخص است، هدف ارزیابی، تعیین کننده سازماندهی ارزیابی، گردآوری داده‌ها، و انتشار و استفاده از نتایج آن است. اگر هدف کنترل باشد، طرح ارزیابی مبتنی بر مدل نتیجه‌محور و به ویژه مدل مبتنی بر هدف، مناسب است.

1- Evaluation

2- Evaluand
3- Evaluation design
4- Formative evaluation
5- Summative evaluation

www.SID.ir

Archive of SID

جدول ۱) انواع مدل‌های ارزیابی [۱۵]

نوع مدل‌ها	مدل ارزیابی	سؤالات	معیارهای ارزیابی
نتیجه‌محور	مدل‌های هدف‌محور[1]	میزان دستیابی به اهداف چقدر بوده است؟	برگرفته از اهداف
	مدل‌های اثرات[2]	چه اثراتی قابل حصول است؟	تمامی اثرات در نظر گرفته شود
فرآیندی	مدل‌های فرآیندی توصیفی[3]	– سطح فعالیت رضایت‌بخش است؟ – مسائل اجرایی وجود دارد؟	عملکرد از ایده به تصمیم و اجرا و بازخورد تحلیل شود
سیستمی	مدل‌های سیستمی[4]	عملکرد کلی چگونه است؟	سنجش ورودی، خروجی، پیامد و اثر نسبت به اهداف
اقتصادی	مدل‌های هزینه–کارایی	آیا بهره‌وری رضایت‌بخش است؟	سنجش خروجی نسبت به هزینه
	مدل‌های هزینه–اثربخشی	آیا اثربخشی رضایت‌بخش است؟	سنجش اثر نسبت به هزینه
	مدل‌های هزینه–منفعت	آیا سودمندی رضایت‌بخش است؟	سنجش سودمندی نسبت به هزینه
بازیگر[5]	مدل‌های مشتری‌مدار[6]	آیا مشتریان راضی هستند؟	توسط مشتری پیشنهاد می‌شود
	مدل‌های ذینفعان[7]	آیا ذینفعان راضی هستند؟	توسط ذینفع پیشنهاد می‌شود
	مدل‌های بررسی هم‌تراز	آیا کیفیت حرفه‌ای تامین شده است؟	توسط هم‌ترازان پیشنهاد می‌شود
نظریه برنامه (ارزیابی نظریه‌محور)[8]	مدل‌های نظری	– چه کارهایی برای چه کسانی در چه بافتاری؟ – آیا تعیین خطاها در نظریه برنامه امکان‌پذیر است؟	نظریه برنامه از طریق تحلیل تجربی بازسازی و ارزیابی می‌شود

اگر هدف یادگیری است، بهترین طرح ارزیابی، بر اساس مدل ذینفعان است. البته لازم به ذکر است که که در اکثر موارد، مطلوبترین حالت انتخاب ترکیبی از مدل‌های ارزیابی باشد، چرا که انتخاب یک مدل خاص معمولا ریسک نگرش تک‌بُعدی به هدف ارزیابی را دارد.

۳-۲ توجه به ویژگی‌های ارزیابی‌شونده

سری دوم معیارها مبتنی بر ویژگی‌های ارزیابی‌شونده است. این منطق برگرفته از دیدگاه اقتضائی در نظریه سازمان است که می‌گوید: "ویژگی‌های وظیفه‌ای و محیط یک سازمان، پیش‌فرض لازم برای طراحی ساختار آن است" [۱۷]. این استدلال خود به دو گونه تقسیم می‌شود: ۱) انتخاب مدل (مدل‌ها) بر مبنای ویژگی‌های احتمالی ارزیابی شونده (استدلال احتمال[9])؛ ۲) انتخاب مدل (مدل‌ها) بر مبنای ویژگی‌های مشروع ارزیابی شونده (استدلال مشروعیت[10]). منطق استدلال احتمال این است که برخی مدل‌های ارزیابی قوی‌تر از سایرین هستند و در نتیجه، کاربرد آنها ارجح است.

جدول ۲) انتخاب مدل ارزیابی بر مبنای هدف ارزیابی [۱۵]

طرح ارزیابی	هدف ارزیابی	
	کنترل	یادگیری
مدل ارزیابی	تأکید اصلی بر مدل‌های نتیجه‌محور	تأکید اصلی بر مدل‌های فرآیندمحور
سازماندهی	بالا به پائین	خود ارزیابی یا هم‌ترازان و مشاوران مستقل
معیار	برگرفته از اهداف سطح بالا	برگرفته از درخواست‌های ذینفعان
روش	تأکید اصلی بر اندازه‌گیری کمی اثرات	تأکید اصلی بر مطالعات فرآیندی کیفی
انتشار نتایج	حرکت رو به بالا در سلسله‌مراتب	از طریق تعامل میان گروه‌های علاقمند
استفاده	مبنایی برای ضمانت و کنترل برای سطح بالاتر	مبنایی برای خود ادراکی و توسعه سازمانی محلی

1- Goal-attainment model
2- Effects model
3- Explanatory process model
4- System model
5- Actor model
6- Client-oriented model
7- Stakeholder model
8- Programme theory model (theory-based evaluation)
9- Possibility Reasoning
10- Legitimacy and justice reasoning

www.SID.ir

Archive of SID

در واقع، حالت بهینه، استفاده از مدل‌های اقتصادی ارزیابی است. اگر مقدور نبود، ارزیابی هدف‌محور گزینه بعدی است؛ اگر نه مسیر کاملاً متفاوتی باید دنبال شود. جدول ۳ نشان می‌دهد که دو ویژگی ارزیابی شونده (هدف و دانش نظری برنامه)، انتخاب مدل ارزیابی را ممکن می‌سازد. مطابق با این استدلال، تنها زمانی باید یک مدل اقتصادی را انتخاب کرد که اهداف ارزیابی شونده (یک برنامه یا یک سازمان) شفاف بوده و دانش نظری برنامه[1] فعالیت‌ها کامل است. فقط تحت این شرایط مدل اقتصادی ارزیابی خواهد کرد که آیا عملکرد با موفقیت بهینه شده است؟ اگر هدف روشن اما دانش نظری برنامه ناقص است، قوی‌ترین مدل ارزیابی در دسترس، مدل هدف‌محور است.

جدول ۳) انتخاب مدل ارزیابی بر مبنای استدلال احتمال [۱۵]

هدف	دانش درباره نظریه برنامه	
	کامل	ناقص
شفاف	مدل‌های اقتصادی	مدل‌های هدف‌محور
مبهم	ارزیابی تطبیقی	ارزیابی تطبیقی

معیار دیگر در طراحی مدل ارزیابی، استدلال مشروعیت است. اساس این استدلال، آن است که بخش‌های حاکمیتی ناهمگن هستند. حل مسئله در بافتارهای مختلف و در زمینه‌های سیاستی متعدد، به صورت متفاوتی سازمان یافته است. فعالیت‌ها، برنامه‌ها و سازمان‌های دولتی به شبکه‌ای از ارزش‌ها و بافتارهای ناهمگن پیوند خورده‌اند. در این راستا، یورگسن[2] [۱۸] چهار نوع بافتار ایده‌آل معرفی نمود که هر یک مدل ارزیابی متفاوتی را می‌طلبد (جدول ٤).

جدول ٤) انتخاب مدل ارزیابی بر مبنای استدلال مشروعیت [۱۸]

مدل دولت	بافتار سازمانی	مدل ارزیابی
پارلمانی	فرماندهی زنجیره پارلمانی	مبتنی بر هدف
خودمختار	متخصصان	بررسی هم‌تراز
تعاملی	گروه‌های علاقمند	ذینفعان کلیدی
پاسخگو	ارباب رجوعان	مشتری‌مدار

۳-۳ توجه به ویژگی‌های مسئله

مجموعه سوم معیارها برگرفته از درک مسئله‌ای است که توسط ارزیابی شونده باید تحلیل شود. این نوع طرح ارزیابی، دارای قابلیت توسعه بیشتر نظریه برنامه است و به عنوان "ارزیابی واقع‌گرایانه" نامیده می‌شود. به ویژه، صلاحیت پیاده‌سازی این روش در زمینه‌های سیاستی مانند پیشگیری از جرم و ارتقاء سلامت به اثبات رسیده است؛ اما به نظر می‌رسد در ارزیابی‌های بسیار پیچیده و مداخلات یکپارچه، مانند اصلاحات بخش دولتی و اصلاحات قانونی که به طور هم‌زمان چند بازیگر به طور مشترک و به صورت شبکه‌ای به طور هم‌زمان درگیر یک مسئله خاص هستند، استفاده از آن مشکل است [۱۹]؛ چرا که شناسایی روابط علی میان بافتار، سازوکار و نتایج، زمانی که تعداد متغیرها افزایش می‌یابد، ممکن نیست. به عبارت دیگر، استفاده از رویکرد واقع‌گرایانه غیرواقعی می‌شود، اگر مشخص کردن تمامی متغیرهای مهم غیرممکن باشد.

٤-۳ خروجی‌های مدل

علاوه بر انتخاب طرح مدل ارزیابی، عامل مهم دیگر خروجی‌های مورد نظر آن است. معمولا به منظور اخذ خروجی از یک مدل ارزیابی از شاخص‌های کمی، کیفی یا ترکیبی (کمی و کیفی) استفاده می‌شود. نوع شاخص‌های انتخابی با توجه به مدل ارزیابی متفاوت خواهد بود. به عنوان مثال، در مدل‌های فرآیندی معمولاً از شاخص‌های سنجش فعالیت‌ها استفاده می‌شود، اما در مدل‌های عملکردمحور سنجش هر چهار نوع شاخص درون‌دادی، برون‌دادی، پیامدی و اثرگذاری موردنظر است. در مدل‌های بررسی هم‌تراز یا مدل‌های اثرات، عمدتاً شاخص‌های کیفی با بهره‌گیری از نظر خبرگان حوزه مورد توجه هستند [۱۵].

۵-۳ رویکرد مدل‌سازی منتخب

همانطور که پیش از این نیز در بخش مقدمه بیان شد، رویکردهای متأخر ارزیابی STI به سمت یکپارچگی پیش رفته و به ویژه توجه هم‌زمان به کیفیت و کمیت در ارزیابی را می‌طلبد. با توجه به هدف تلفیق این دو موضوع در مدل ارزیابی موردنظر این مقاله، هر دو بعد کنترل (کمی) و بهبود (کیفی) موردنیاز است (جدول ۲) اگر چه رویکرد غالب یادگیری و بهبود است؛ لذا تلفیقی از مدل‌های نتیجه‌محور

۱- «دانش نظری برنامه» در ارزیابی به معنای کاربرد نظام‌مند دانش واقع‌گرایانه درباره پدیده‌های تحت بررسی است و از روش‌های علمی برای بهبود، پدیدآوری و بازخورد در مورد این دانش استفاده می‌کند تا ارزش‌مندی و اهمیت موضوعات ارزیابی همچون برنامه‌های اجتماعی، آموزشی، سلامتی، عمومی یا سازمانی را تعیین کند [۱۷].

2- Jørgensen

www.SID.ir

Archive of SID

(همچون مدل نظام‌مند) و فرآیندمحور (مانند مدل ذینفعان کلیدی) می‌تواند مناسب باشد. با توجه به بافتار[1] شبکه‌ای و چندبُعدی نظام‌های STI در سطوح مختلف آن نیز لازم است با بهره‌گیری از استدلال تغییر (جدول ٤) مدل ذینفعان کلیدی طراحی شود[2].

به عنوان جمع‌بندی این بخش می‌توان گفت در طراحی مدل ارزیابی مطلوب می‌باید تلفیقی از مدل‌های ارزیابی نظام‌مند یا عملکردمحور با ذینفعان کلیدی را در نظر گرفت تا بحث یکپارچگی هم در زنجیره ارزش STI و هم در دخیل کردن منابع همه ذینفعان کلیدی تأمین شود. مضافاً در مدل‌های ارزیابی نظام‌مند ، طراحی شاخص‌ها به گونه‌ای است که می‌تواند از جنبه‌های مختلف بافتار مورد نظر را ارزیابی کند. این موضوع در بخش بعدی مورد بحث بیشتر قرار گرفته است.

٤- ارزیابی عملکرد علم، فناوری و نوآوری

بر اساس مباحث بخش پیشین، یکی از انواع مدل‌های نتیجه‌محور به ویژه در سطح ارزیابی سازمانی، مدل‌های ارزیابی نظام‌مند یا عملکردمحور است که عملکرد کلی نظام را از طریق سنجش درون‌دادها، نتایج و آثار نسبت به اهداف تعیین شده مورد ارزیابی قرار می‌دهد. درواقع «ارزیابی عملکرد»[3] فرآیندی است که به سنجش و اندازه‌گیری، ارزش‌گذاری و قضاوت درباره عملکرد طی دوره‌ای معّین می‌پردازد [۲۰]. در بُعد سازمانی هم شامل میزان دستیابی به اهداف و برنامه‌ها و هم کارا بودن فعالیت‌ها و عملیات است و اندازه‌گیری عملکرد از طریق مقایسه وضع موجود با وضع مطلوب یا ایده‌آل براساس شاخص‌های از پیش تعیین شده، انجام می‌شود. بنابراین می‌توان نظام ارزیابی عملکرد را چنین تعریف کرد [۲۱]: «فرآیند سنجش، اندازه‌گیری و مقایسه میزان و نحوه دستیابی به وضعیت مطلوب با توجه به شاخص‌های معّین، در دامنه تحت پوشش معّین، در دوره زمانی معّین، با هدف بازنگری، اصلاح و بهبود مستمر آن».

1- Context

۲- طبعا معیار سوم یعنی انتخاب مدل بر مبنای ویژگی‌های مسئله، در رابطه با مسائل چندجانبه و پیچیده چندان جوابگو نیست، زیرا با افزایش تعداد متغیرها از حالت واقع‌گرایانه خود خارج خواهد شد، لذا در این مقاله بحث نمی‌شود.

3- Performance Evaluation

به علاوه، مبحث ارزیابی عملکرد را می‌توان از زوایای متفاوتی مورد بررسی قرار داد؛ توجه به ارزیابی عملکرد از دیدگاه نظام‌مند[4]، در هر سطحی که انجام پذیرد، ناگزیر از آن خواهد بود که فرآیند عملکردی را در نظر بگیرد. براون و استیونسن در تحقیقات خود پیرامون نظام‌های ارزیابی عملکرد STI، پیش‌نیاز درک علت موفقیت یا شکست نظام‌های ارزیابی STI را برخورداری از نگاه فرآیندی به آن دانستند. به همین دلیل زمانی زمانی یک نظام ارزیابی STI مؤثر عمل خواهد کرد که به جای توجه به «رفتار» (فعالیت‌های انجام شده توسط پژوهشگران و فنّاوران)، بر «برون‌دادها» و «پیامدها»ی این فعالیت‌ها متمرکز شود. رفتار در هنگام ارزیابی یک پژوهشگر منفرد مهم است، اما برای ارزیابی عملکرد یک نظام مناسب نیست. در این حالت خروجی‌ها باید از هر دو بُعد کیفی و کمی ارزیابی شوند [۲۲].

اولین بار اتحادیه اروپا مدل ارزیابی عملکرد نظام‌مند را تحت عنوان «منطق مداخله[5]» معرفی کرد. بر اساس این منطق، نقطه شروع هر مداخله عمومی «تحلیل موقعیت موجود و تشخیص مسائل یا نیازهای آن» است. از این منظر، علاوه بر توجه به تعاملات درونی یک نظام، در نظر گرفتن تعامل آن با محیط پیرامون نیز ضروری است. در واقع، ارزیابی نتایج کوتاه‌مدت و بلندمدت هر مداخله بر محیط پیرامون سیستم و تأثیرپذیری آن از محیط، لازم است. در چنین مدلی برای ارزیابی، معمولاً شاخص‌ها منبع اصلی اطلاعات هستند که در مراحل مختلف ارزیابی استفاده می‌شوند و این شاخص‌ها باید تعریف، ارزش و واحد اندازه‌گیری داشته باشند. در این مدل شاخص‌محور (IOOI[6])، «شاخص‌های درون‌دادی» معمولاً مشتمل بر منابع مالی، منابع انسانی و تجهیزات هستند. «شاخص‌های برون‌دادی» مربوط به نتایج مستقیم فعالیت‌های انجام شده در نظام است. «شاخص‌های پیامدی» مربوط به نتایج سطح متوسط (زمانی و عملکردی) برای بهره‌برداران نظام است. این شاخص‌ها اطلاعاتی را درباره تغییرات رفتاری، ظرفیت یا عملکرد، ارائه می‌کنند و می‌توانند به صورت فیزیکی و یا مالی تعریف شوند. «شاخص‌های اثرگذاری» به نتایج سطح

4- Systematic Approach
5- Logic of intervention
6- Input/Output/Outcome/Impact

www.SID.ir

Archive of SID

کلان و معمولاً بلندمدت (که معمولاً در سطح جامعه معنا پیدا می‌کند)، اطلاق می‌شود. شایان ذکر اینکه دو مفهوم مرتبط با «تأثیر» قابل تعریف است [۲۳]:

✍ تأثیر خاصی که پس از گذشت زمان معیّنی رخ می‌دهد، اگر چه به طور مستقیم به اقدامات انجام شده مرتبط نیست.

✍ تأثیر عامی که بلندمدت است و بر جمعیت گسترده‌تری اعمال می‌شود.

با توجه به نسبت شاخص‌های پیش‌گفته می‌توان عملکرد یک سیستم را بر اساس تناسب (تطابق) با اهداف، کارایی، اثربخشی، سودمندی و پایداری اندازه گرفت [۲۳]: (شکل ٤)

✍ «تناسب»[1] یا «تطابق»،[2] حلقه اتصال میان اهداف تعریف شده و مسائل و نیازهای موجود است که باید از طریق مداخله حل شود.

✍ «کارایی»[3] نسبت برون‌دادها به درون‌دادهای مورد استفاده برای دستیابی به آنها (معمولاً منابع مالی) است.

✍ «اثربخشی»[4] نسبت پیامدهای واقعی به درون‌دادهای مورد استفاده برای دستیابی به آنها است.

✍ «سودمندی»[5] نسبت تأثیر حاصل از مداخله مورد استفاده برای دستیابی به آنها را نشان می‌دهد.

✍ «پایداری»[6] توانایی تأثیر آثار میان مدت یا بلندمدت را بررسی می‌کند.

در خصوص شناسایی مؤلفه‌های اصلی ارزیابی عملکرد، علاوه بر منطق مداخله، توجه به نیازهای ذینفعان ارزیابی نیز بسیار مهم است. همانطور که پیش از این نیز اشاره شد، رویکردهای اخیر ارزیابی STI مسیری تکاملی را طی کرده و در نهایت، به سمت اتخاذ رویکردهای جامع با در نظر گرفتن منافع کلیه ذینفعان کلیدی پیش رفته است؛ لذا با تلفیق «منطق مداخله» که ارزیابی عملکردمحور را ارجح می‌داند و «مدل ارزیابی ذینفعان کلیدی[7]»، مدل مفهومی ارزیابی STI در شکل شکل ۵ نشان داده شده [٢٤] که هر یک از ذینفعان کلیدی ارزیابی فرآیند STI به بخش خاصی از این فرآیند مرتبط هستند؛ بنابراین می‌توان متناظر با هر گروه ذینفع، بخش خاصی از نتایج را اختصاص داد:

➢ اولین گروه ذینفعان، سیاست‌گذاران حوزه STI هستند که وظیفه اصلی هدفگذاری و به تناسب آن تخصیص درون‌دادها به فعالیت‌های STI بر اساس نیازهای جامعه را بر عهده دارند. بنابراین، کاربر اصلی ارزیابی «اهداف» و به تبع آن درون‌دادها، سیاست‌گذاران خواهند بود که از آن برای ارزیابی سیاست‌های کلان STI استفاده خواهند کرد.

➢ متخصصین (پژوهشگران، فناوران و نوآوران) به عنوان دومین گروه ذینفعان کلیدی کسانی هستند که معمولاً نتایج مستقیم و بلافاصله فعالیت‌های STI حاصل کار ایشان است، لذا ارزیابی «برون‌دادهای» فعالیت این گروه در فرآیند STI بیانگر میزان کارایی پروژه‌ها خواهد بود.

➢ سومین گروه ذینفعان کلیدی، بهره‌برداران نتایج فعالیت‌های STI هستند که نقش آنها معمولاً مربوط به کاربردی‌سازی نتایج فعالیت‌های علم و فناوری به ویژه در راستای تجاری‌سازی و توسعه فعالیت‌های نوآورانه است. با توجه به اینکه زمانی فعالیت‌های STI اثربخش خواهد بود که بتوان آن را در عرصه عمل به کار بست، ارزیابی «پیامدهای» فعالیت بهره‌برداران، اثربخشی فعالیت‌های STI را نشان خواهد داد.

➢ زمانی فعالیت‌های STI سودمند خواهد بود که منافع نهایی آن به جامعه برسد و نیازهای جامعه را برطرف نماید؛ لذا «آثار» اغلب بلندمدت فعالیت‌های STI ملاک ارزیابی سودمندی خواهد بود که منتفع آن، عموم مردم جامعه هستند.

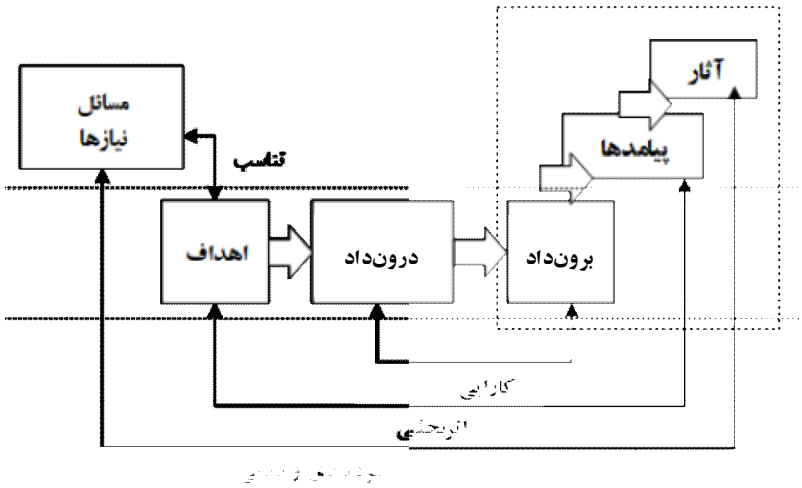
بدین ترتیب با ایجاد تناظر یک به یک میان ذینفعان کلیدی STI با عناصر منطق مداخله، سیری نتیجه‌گرا ایجاد می‌شود و دوباره می‌توان بازخوردی را از آثار عملکرد STI به اهداف سیاست‌گذاری، دریافت کرد که در قالب کمربند حول هسته مرکزی مدل مفهومی به صورت نقطه‌چین مشاهده می‌شود. با توجه به تفاوت ماهوی چهار ربع ارزیابی عملکرد STI در این مدل (تناسب، کارایی، اثربخشی و سودمندی)، شاخص‌های ذیل هر ربع می‌باید با توجه به حوزه فعالیت‌های علم و فناوری مورد ارزیابی و ماهیت درون‌دادها و برون‌دادهای هر ربع مشخص شود [٢٤].

از باب چارچوب کاربردی این مدل مطابق آنچه قبلاً بیان شد می‌توان گفت در ارزیابی همه حوزه‌های علم و فناوری (شکل ۲)، برای تمام سطوح ارزیابی (شکل ۳) در قالب نگاه نظام‌مند

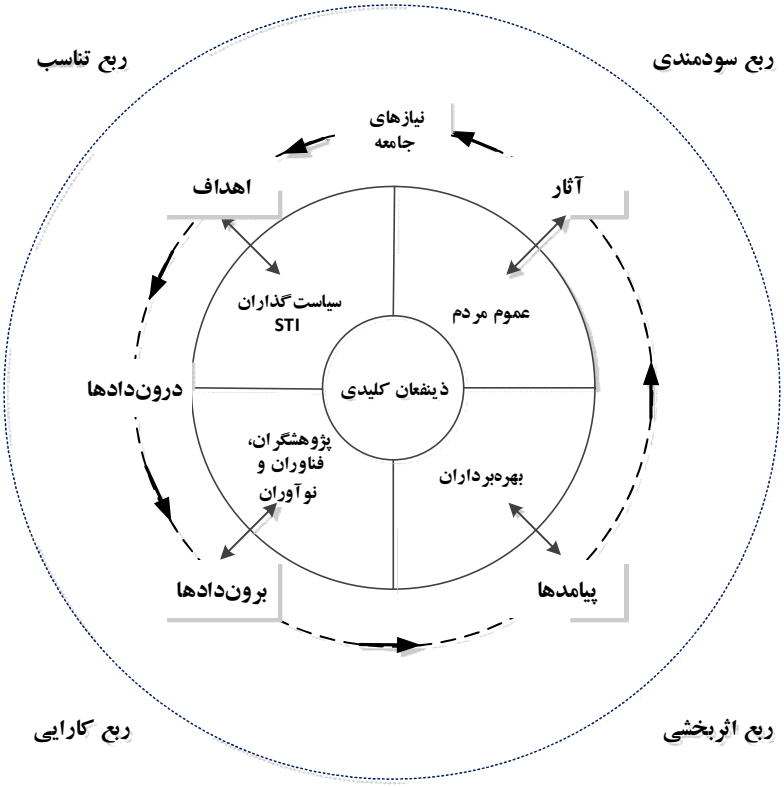
1- Relevance
2- Adaptation
3- Efficiency
4- Effectiveness
5- Utility
6- Sustainability
7- Key Stakeholder Evaluation model

www.SID.ir

Archive of SID



شکل ٤) مفهوم منطق مداخله در نظام ارزیابی عملکرد شاخص‌محور [۲۳]



شکل ٥) مدل مفهومی ارزیابی STI [٢٤]

و در تمامی کارکردهای ارزیابی کنترلی، تخصیصی و بهبودی (با رویکرد غالب یادگیری وبهبود)، قابل استفاده است، صرفاً اگر شاخص‌های ارزیابی هر ربع در تناسب با بافتار ارزیابی مورد نظر، انتخاب و بکارگیری شوند.

٥- نحوه انتخاب و بکارگیری شاخص‌های ارزیابی STI

بر اساس آنچه گفته شد، خروجی‌های مدل ارزیابی از طریق شاخص‌های کمی، کیفی یا ترکیبی قابل استخراج هستند. از آنجا که مفهوم یادگیری و ایجاد حلقه بازخورد در تعریف ارزیابی عملکرد STI مستتر است، نحوه انتخاب شاخص‌ها و چگونگی ارتباط آنها با یکدیگر، اهمیتی اساسی پیدا می‌کند. بدین منظور در انتخاب و بکارگیری شاخص‌های مناسب می‌باید حداقل به دو پرسش کلیدی پاسخ داد [۲۵]:

- گونه‌شناسی فعالیت‌های STI و معیارهای منطبق بر آن کدامند؟

www.SID.ir

Archive of SID

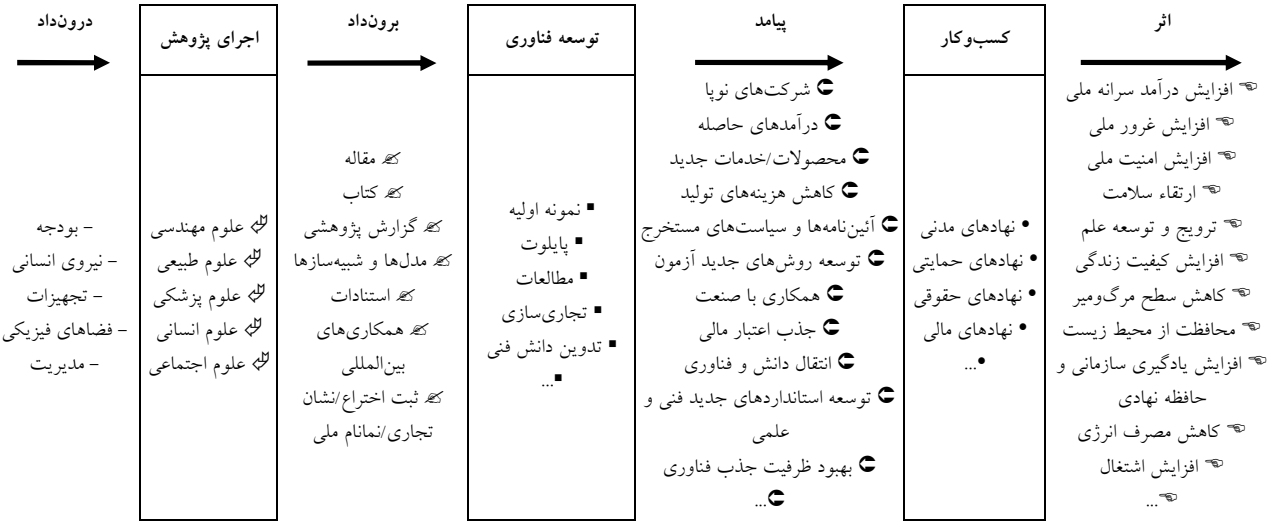
⌖ ارزیابی بر اساس‌کدام‌یک از ربع‌های ارزیابی (تناسب، کارایی، اثربخشی، سودمندی) مدنظر است؟ (انتخاب یک یا چند ربع)

در پاسخ به پرسش اول، در محدوده فعالیت‌های علم و فناوری مورد ارزیابی، باید گونه‌شناسی مناسب انتخاب شود. تخصصی شدن دانش و گسترش دامنه علوم در طول زمان باعث پیدایش رشته‌های علمی شده است. با توجه به تفاوت‌های موجود در ماهیت این حوزه‌ها، سیاست‌گذاری علمی مطلوب در هر کدام از آنها بدون در نظر داشتن این تفاوت‌ها امکان‌پذیر نیست [۲۶]؛ به عنوان نمونه، در ارزیابی حوزه تحقیق و توسعه، کشورهای مختلف طبقه‌بندی‌های مختلفی را از حوزه‌های علمی-پژوهشی برگزیده‌اند که یکی از کاربردی‌ترین آنها طبقه‌بندی بانک جهانی به شرح ذیل است [۲۷]:

- علوم طبیعی شامل: ریاضی، فیزیک، شیمی، زمین، زیست، محیط زیست، کشاورزی و دامپزشکی، کامپیوتر و اطلاعات
- علوم مهندسی و فناوری شامل: مهندسی، فناوری، طراحی و ساخت
- علوم انسانی شامل: هنر و نویسندگی، زبان، ارتباطات، فرهنگ، تاریخ و باستان‌شناسی، فلسفه و مطالعات دینی
- علوم اجتماعی شامل: آموزش، اقتصاد، تجارت، مدیریت، گردشگری و خدمات، روانشناسی و علوم شناختی، مطالعات حقوقی و قانونی، مطالعات جوامع انسانی
- علوم سلامت شامل: بیوشیمی پزشکی، علوم کلینیکی، دندانپزشکی، علوم ورزشی و حرکت، ایمنی‌شناسی، علوم عصبی، پرستاری، تغذیه، داروسازی، سلامت عمومی و خدمات سلامت، فیزیولوژی پزشکی

در پاسخ به پرسش دوم، با اتخاذ رویکرد نظام‌مند و یکپارچه (درنظر گرفتن تمامی ذینفعان کلیدی) در طراحی مدل ارزیابی (شکل ۵)، می‌توان بخشی یا تمامی نسبت‌های ارزیابی عملکرد اعم بر: تناسب، کارایی، اثربخشی و سودمندی به کار را به کار گرفت. در خصوص نسبت «پایداری» می‌باید ارزیابی نسبت سودمندی به صورت دوره‌ای انجام و نسبت به دوره پیشین مقایسه شود. در شکل ٦ تصویری ساده شده از شاخص‌های درون‌دادی، برون‌دادی، پیامدی و اثرگذاری در ارزیابی STI، به منظور ایجاد تقریب ذهنی بیشتر ارائه شده است [۲۸].

به منظور مشخص نمودن شاخص‌های سنجش (درون‌دادی، برون‌دادی، پیامدی و اثرگذاری) در هر حوزه تخصصی (علوم طبیعی، علوم مهندسی، علوم سلامت، علوم انسانی و علوم اجتماعی) - که در واقع تلفیق گونه‌شناسی حوزه‌های تخصصی با شاخص‌های سنجش است - مجموعه (استخر) شاخص‌های هر حوزه تخصصی به تفکیک نوع شاخص تدوین می‌شود. به طور معمول شاخص‌های درون‌دادی برای تمامی حوزه‌های تخصصی مشابه هستند و اغلب تفاوت شاخص‌ها در برون‌دادها، پیامدها و آثار قرار می‌گیرد. به ویژه، ارزیابی علوم انسانی و اجتماعی با شاخص‌های برون‌دادی، پیامدی و اثرگذاری مشترک با سایر حوزه‌های تخصصی به دلیل تفاوت‌های سرشتی این علوم در جهت‌گیری محلی و ملی قوی‌تر، رفتارهای استنادی و انتشاراتی و نیز فرهنگ



شکل ٦) نمایی از رویکرد فرآیندی به شاخص‌های ارزیابی STI [۲۸]

www.SID.ir

Archive of SID

پژوهش، محل انتقاد بسیاری است [۲۹]. با توجه به تفاوت‌های ماهوی مطرح‌شده در جدول‌های ۵ تا ۹ مجموعه شاخص‌های هر حوزه تخصصی به تفکیک ارائه شده است.

جدول ۵) مجموعه شاخص‌های علوم طبیعی [۳۰-۳۳]

شاخص‌های درون‌دادی	شاخص‌های برون‌دادی	شاخص‌های پیامدی	شاخص‌های اثرگذاری
– منابع مالی – منابع انسانی – تجهیزات – فضای فیزیکی	مقاله (مجله /کنفرانس)؛ ارجاعات؛ کتاب (درسی)/ فصل کتاب؛ دست‌نامه‌ها؛ بهبود کیفیت؛ مدل‌ها و شبیه‌سازها؛ سازماندهی کنفرانس (دانشگاهی)/ کارگاه آموزشی/ سمینار آموزشی؛ ساخت نمونه اولیه (پایلوت)؛ مربیگری/داوری پایان‌نامه‌ها؛ فعالیت‌های بررسی هم‌تراز/ عضویت در شوراهای تخصصی/ مشارکت در جوامع یادگیری؛ ویراستاری مجلات؛ جوایز معتبر؛ کمک‌هزینه تحصیلی و پژوهشی؛ همکاری‌های بین‌المللی و مطرح شدن در سطح بین‌المللی؛ ثبت اختراع/ نشان تجاری/ برند ملی؛ محصولات، مواد و فرآیندهای جدید یا بهبودیافته؛ کاهش هزینه‌های تولید	توسعه روش‌شناسی جدید تست؛ توسعه استانداردهای جدید علمی؛ درآمدهای پژوهشی؛ ایجاد شرکت‌های نوپا؛ همکاری با صنعت؛ جذب اعتبار از شخص ثالث؛ برپایی نمایشگاه؛ تعداد یا مبلغ پروژه‌های بعدی در راستای پروژه حاضر؛ انتقال دانش و فناوری	کاهش سطح مرگ و میر؛ ترویج و توسعه علمی؛ افزایش سرمایه‌گذاری در بخش علوم طبیعی؛ افزایش غرور ملی؛ افزایش یادگیری یا حافظه نهادی؛ افزایش امنیت ملی؛ افزایش کیفیت زندگی عموم افراد جامعه؛ افزایش امنیت ملی؛ محافظت از محیط زیست

جدول ٦) مجموعه شاخص‌های علوم مهندسی [۳۱و۳٤و۳۵]

شاخص‌های درون‌دادی	شاخص‌های برون‌دادی	شاخص‌های پیامدی	شاخص‌های اثرگذاری
منابع مالی؛ منابع انسانی؛ تجهیزات؛ فضای فیزیکی	مقاله (مجله /کنفرانس)؛ ارجاعات؛ کتاب (درسی)/ فصل کتاب؛ دست‌نامه‌ها؛ آموزش فنی و حرفه‌ای؛ بهبود کیفیت؛ گزارش‌های فنی؛ سازماندهی کنفرانس (دانشگاهی)/ کارگاه آموزشی/ سمینار آموزشی؛ ساخت نمونه اولیه (پایلوت)؛ مربیگری/ داوری پایان‌نامه‌ها؛ فعالیت‌های بررسی هم‌تراز/ عضویت در شوراهای تخصصی/ مشارکت در جوامع یادگیری؛ ویراستاری مجلات؛ جوایز معتبر؛ کمک‌هزینه تحصیلی و پژوهشی؛ همکاری‌های بین‌المللی و مطرح شدن در سطح بین‌المللی؛ ثبت اختراع/ نشان تجاری/ برند ملی؛ محصولات، مواد و فرآیندهای جدید یا بهبود یافته؛ کاهش هزینه‌های تولید؛ ادغام فناوری	توسعه روش‌شناسی جدید تست؛ توسعه استانداردهای جدید فنی؛ درآمدهای پژوهشی؛ ایجاد شرکت‌های نوپا؛ همکاری با صنعت؛ جذب اعتبار از شخص ثالث؛ برپایی نمایشگاه؛ تعداد یا مبلغ پروژه‌های بعدی در راستای پروژه حاضر؛ توسعه اتحادهای راهبردی فناوری؛ بهبود ظرفیت جذب فناوری؛ انتقال دانش و فناوری	کاهش سطح مرگ و میر؛ ترویج و توسعه علمی؛ افزایش سرمایه‌گذاری در بخش علوم مهندسی؛ افزایش پرستیژ ملی؛ افزایش یادگیری یا حافظه نهادی؛ افزایش امنیت ملی؛ افزایش کیفیت زندگی عموم افراد جامعه؛ افزایش امنیت ملی؛ بهبود حمل و نقل؛ افزایش صادرات؛ محافظت از محیط زیست؛ کاهش مصرف انرژی

جدول ۷) مجموعه شاخص‌های علوم انسانی [۳۲و۳٦-۳۸]

شاخص‌های درون‌دادی	شاخص‌های برون‌دادی	شاخص‌های پیامدی	شاخص‌های اثرگذاری
منابع مالی؛ منابع انسانی؛ تجهیزات؛ فضای فیزیکی	مقاله (مجله /کنفرانس)؛ مقاله مباحثه‌ای؛ ارجاعات؛ کتاب (درسی)/ فصل کتاب؛ کتاب‌های مورد علاقه عموم؛ نقد کتاب؛ دست‌نامه‌ها؛ ترجمه؛ روزنامه/ مجله/ بولتن/ ماهنامه/ مجلات آنلاین؛ کارهای کتاب‌شناسی عمده/ کارهای مروری سالیانه در یک حوزه؛ ایجاد پایگاه داده و مجموعه‌های داده؛ خروجی‌های مستقیم مانند نت موسیقی/ فیلمنامه (تئاتر/ مستند)/ نمونه کارهای هنری/ گزارش کاوش یا حفاری/ گزارش حفاظت از هنر؛	راه‌اندازی کارهای خلاقانه؛ اجرای موسیقی/ تئاتر/ ساخت فیلم؛ درآمدهای پژوهشی؛ ایجاد شرکت‌های نوپا؛ همکاری با صنعت و بازار؛ جذب اعتبار از شخص ثالث؛ برپایی نمایشگاه؛ آئین‌نامه‌ها و سیاست‌های جدید یا تغییر یافته؛ تولید محصولات جدید فرهنگی/ ادبی/ هنری؛ همکاری در سیاست‌گذاری؛ تعداد یا مبلغ پروژه‌های بعدی در راستای پروژه حاضر	تغییر فهم عمومی مردم؛ ترویج و توسعه علمی؛ افزایش سرمایه‌گذاری در بخش علوم انسانی؛ افزایش پرستیژ ملی؛ افزایش یادگیری یا حافظه نهادی؛ تغییر سیاست‌های عمومی؛ افزایش کیفیت زندگی عموم افراد جامعه؛ افزایش امنیت ملی؛ حفظ میراث فرهنگی

www.SID.ir

Archive of SID

جدول ۷) مجموعه شاخص‌های علوم انسانی [۳۲و۳۶-۳۸]

شاخص‌های درون‌دادی	شاخص‌های برون‌دادی	شاخص‌های پیامدی	شاخص‌های اثرگذاری
	تک‌نگاشت یا رساله؛ سازماندهی کنفرانس (دانشگاهی)/ کارگاه آموزشی/ سمینار آموزشی؛ مربیگری/ داوری پایان‌نامه‌ها؛ فعالیت‌های بررسی هم‌تراز/ عضویت در شوراهای تخصصی/ مشارکت در جوامع یادگیری؛ ویراستاری مجلات؛ جوایز معتبر؛ کمک‌هزینه تحصیلی و پژوهشی؛ همکاری‌های بین‌المللی و مطرح شدن در سطح بین‌المللی؛ حضور در رسانه‌ها/ سخنرانی‌ها؛ ثبت اختراع/ نشان تجاری/ برند ملی		

جدول ۸) مجموعه شاخص‌های علوم اجتماعی [۳۳و۳۶]

شاخص‌های درون‌دادی	شاخص‌های برون‌دادی	شاخص‌های پیامدی	شاخص‌های اثرگذاری
منابع مالی؛ منابع انسانی؛ تجهیزات؛ فضای فیزیکی	مقاله (مجله /کنفرانس)؛ مقاله مباحثه‌ای؛ ارجاعات؛ کتاب (درسی)/ فصل کتاب؛ کتاب‌های مورد علاقه عموم؛ نقد کتاب؛ دست‌نامه‌ها؛ ترجمه؛ روزنامه/ مجله / بولتن / ماهنامه /مجلات آنلاین؛ کارهای کتاب‌شناسی عمده/ کارهای مروری سالیانه در یک حوزه؛ ایجاد پایگاه داده و مجموعه‌های داده؛ جوایز معتبر؛ کمک‌هزینه تحصیلی و پژوهشی؛ همکاری‌های بین‌المللی و مطرح شدن در سطح بین‌المللی؛ حضور در رسانه‌ها/ سخنرانی‌ها؛ تک‌نگاشت یا رساله؛ ثبت اختراع/ نشان تجاری/ برند ملی؛ سازماندهی کنفرانس (دانشگاهی)/ کارگاه آموزشی/ سمینار آموزشی؛ مربیگری/ داوری پایان‌نامه‌ها؛ فعالیت‌های بررسی هم‌تراز/ عضویت در شوراهای تخصصی/ مشارکت در جوامع یادگیری؛ ویراستاری مجلات	مدیریت و توسعه تیم‌های میان‌رشته‌ای؛ ایجاد مکتب فکری جدید؛ درآمدهای پژوهشی؛ ایجاد شرکت‌های نوپا؛ همکاری با صنعت و بازار؛ جذب اعتبار از شخص ثالث؛ برپایی نمایشگاه؛ کاربرد روش‌شناسی به دست آمده؛ آئین‌نامه‌ها و سیاست‌های جدید یا تغییر یافته؛ تعداد یا مبلغ پروژه‌های بعدی در راستای پروژه حاضر؛ همکاری در سیاست‌گذاری	تغییر فهم عمومی مردم؛ ترویج و توسعه علمی؛ افزایش سرمایه‌گذاری در بخش علوم اجتماعی؛ افزایش پرستیژ ملی؛ افزایش یادگیری یا حافظه نهادی؛ تغییر سیاست‌های عمومی؛ افزایش کیفیت زندگی عموم افراد جامعه؛ افزایش امنیت ملی؛ افزایش سلامت اداری؛ افزایش کارآمدی سیستم؛ افزایش اشتغال؛ کنترل تورم

جدول ۹) مجموعه شاخص‌های علوم سلامت [۳۹و۴۰]

شاخص‌های درون‌دادی	شاخص‌های برون‌دادی	شاخص‌های پیامدی	شاخص‌های اثرگذاری
منابع مالی؛ منابع انسانی؛ تجهیزات؛ فضای فیزیکی	مقاله (مجله /کنفرانس)؛ ارجاعات؛ کتاب (درسی)؛ دست‌نامه‌ها؛ آموزش بالینی؛ راهنماها، رویه‌ها و دستورالعمل‌ها؛ انتشارات حرفه‌ای؛ سازماندهی کنفرانس (دانشگاهی)/ کارگاه آموزشی؛ ساخت نمونه اولیه (پایلوت)؛ عضویت در شوراهای تخصصی/ مشارکت در جوامع یادگیری؛ ویراستاری مجلات؛ جوایز معتبر؛ کمک‌هزینه تحصیلی و پژوهشی؛ همکاری‌های بین‌المللی و مطرح شدن در سطح بین‌المللی؛ ثبت اختراع/ نشان تجاری/ برند ملی؛ محصولات، مواد و فرآیندهای جدید یا بهبودیافته	استفاده حرفه‌ای از رویه‌ها، دستورالعمل‌ها و راهنماها؛ توسعه استانداردهای جدید سلامت؛ درآمدهای پژوهشی؛ ایجاد شرکت‌های نوپا؛ همکاری با صنعت؛ جذب اعتبار از شخص ثالث؛ برپایی نمایشگاه؛ تعداد یا مبلغ پروژه‌های بعدی در راستای پروژه حاضر؛ کاهش هزینه‌های سلامت؛ افزایش تعداد افراد تحت پوشش؛ افزایش استفاده از سرویس‌های خدمات سلامت؛ درآمد حاصل از فروش اختراعات ثبت شده	کاهش سطح مرگ و میر مادران و نوزادان؛ ترویج و توسعه علمی؛ افزایش سرمایه‌گذاری در بخش علوم سلامت؛ افزایش پرستیژ ملی؛ افزایش یادگیری یا حافظه نهادی؛ افزایش میزان اشتغال‌زایی؛ افزایش کیفیت زندگی عموم افراد جامعه؛ افزایش امنیت ملی؛ کاهش نرخ بیماری‌های مربوط به پژوهش مورد نظر؛ تغییر در سیاست‌های حوزه سلامت؛ تغییر در رفتارهای مربوط

www.SID.ir

Archive of SID

جدول ۹) مجموعه شاخص‌های علوم سلامت [۳۹و۴۰]

شاخص‌های درون‌دادی	شاخص‌های برون‌دادی	شاخص‌های پیامدی	شاخص‌های اثرگذاری
			به سلامت افراد؛ افزایش دسترسی به خدمات سلامت؛ افزایش پیشگیری و اهمیت دادن به سلامتی؛ افزایش میزان امید به زندگی یا طول عمر؛ کاهش میزان شیوع بیماری‌ها؛ کاهش میزان هزینه‌کرد برای بیماری‌ها؛ کاهش میزان ابتلاء به بیماری‌ها؛ کاهش میزان مرگ و میر عمومی؛ Out of packet (چقدر مردم از هزینه شخصی برای بیماری‌ها می‌پردازند به عنوان مثال ۷۰٪ بیمه و ۳۰٪ مردم)

پس از گردآوری و غربالگری مجموعه شاخص‌ها در هر حوزه تخصصی، لازم است انتخاب شاخص‌های نهایی و امتیازدهی آنها در مطالعه موردی خاص انجام شود که می‌توان بر این اساس رتبه‌بندی افراد و مؤسسات را هم انجام داد. شایان ذکر است معمولاً تعداد محدودی شاخص باید انتخاب گردد و باید توجه داشت که انتخاب هر شاخص، جهت‌گیری ارزیابی در آن حوزه را تعیین می‌کند.

به منظور نمایش تصویری مدل نهایی، در هر حوزه تخصصی از نمودار راداری استفاده گردیده و مجموعاً برای هر ارزیابی، ۵ نمودار راداری در ۵ حوزه تخصصی ترسیم و تحت عنوان داشبورد ارزیابی نامیده می‌شود. در ادامه مطالعه موردی ارزیابی عملکرد STI در دانشگاه‌ها و پژوهشگاه‌های ایران به منظور تبیین بیشتر نتایج مدل‌سازی ارائه شده است.

۶- مطالعه موردی ارزیابی عملکرد STI در دانشگاه‌ها و پژوهشگاه‌های ایران [۲۵]

روال مرسوم و قدیمی ارزیابی STI در ایران، همان روال کنترل همه پروژه‌های دولتی است؛ به این معنا که صرفاً میزان تخصیص بودجه را به معنای پیشرفت پروژه تلقی کرده و پروژه تحقیقاتی موفق، پروژه‌ای است که ورودی بیشتری را به خود جذب می‌کند. طبیعی است چنین روندی به میزان نتیجه‌بخش بودن فعالیت‌های STI توجهی ندارد.

به منظور فراهم‌سازی امکان مقایسه عملکرد دستگاه‌ها در طی سال‌های مختلف و تحلیل روند رشد شاخص‌های عملکردی پژوهش و فناوری ملی و دستگاهی، در سال ۱۳۸۸ مجلس نخستین گام قانونی را برای ایجاد یک نظام ارزیابی ملی STI برداشت و کلیه سازمان‌ها، دانشگاه‌ها، مراکز پژوهشی و شرکت‌های دولتی را موظف ساخت که گزارش عملکرد بودجه‌های پژوهش و فناوری خود را هر سه ماه یکبار به شورای عالی عتف گزارش دهند تا این شورا بتواند یک گزارش ملی سالانه به مجلس ارائه نماید. بدین منظور این شورا سامانه‌ای با نام «سامانه مدیریت اطلاعات تحقیقاتی (سمات)[1]» طراحی و پیاده‌سازی کرد که کلیه داده‌های مربوطه در آن ثبت و نگهداری شوند. تدوین مجموعه‌ای از حدود ۵۰ شاخص سنجش درون‌داد، برون‌داد، پیامد و اثر پژوهش‌ها و ابلاغ آن به همه دستگاه‌های دولتی برای ثبت داده‌های این شاخص‌ها در این سامانه از دیگر اقدامات این شورا است.

مدل ارزیابی اولیه این سامانه، دارای یک سری شاخص استاندارد برای همه سازمان‌ها بود. این موضوع اشکالات متعددی را برای سازمان‌های دولتی ایجاد می‌کرد؛ به عنوان مثال وزارت بهداشت، درمان و آموزش پزشکی ادعا می‌کرد شاخص‌های مرتبط با حوزه پژوهش و فناوری سلامت، باید به طور خاص برای این دستگاه وزارتی و سازمان‌های تابعه

1- www.semat.ir

www.SID.ir

Archive of SID

در سامانه سمات لحاظ شود و یا وزارت کشور معتقد بود برخی پیامدهای فعالیت‌های علم و فناوری آن به ویژه در بحث اثرات اجتماعی در سامانه فعلی دیده نشده است. بنابراین، بر سر کفایت شاخص‌ها و تناسب آنها توافقی وجود نداشت.

به منظور ارتقاء مدل ارزیابی عملکرد STI طراحی شده، بر اساس آنچه در این مقاله ارائه شد طراحی جدیدی از شاخص‌های ارزیابی سازمان‌های مذکور انجام گرفت (مطالعات نظری و تجربی انجام شده برای این مطالعه موردی، طی سال‌های ۱۳۹۱ تا ۱۳۹۵ انجام شده و داده‌های کیفی آن از طریق اقدام‌پژوهی و انجام مصاحبه- پرسشنامه و داده‌های ثانویه مورد نیاز نیز از مستندات، تجربیات و آرشیو دیداری و شنیداری موجود استخراج گردیده‌اند [۲۵]) که نتایج آن بر اساس گام‌های طراحی مدل پیش‌گفته به شرح ذیل است:

- سطح ارزیابی: سازمانی (خُرد)

همچنان که در شکل ۳ مشخص شده است سطح ارزیابی عملکرد فعالیت‌های علم و فناوری مورد نظر، سازمانی است. شایان ذکر است که بر اساس تحلیل و تجمیع نتایج ارزیابی عملکرد سازمانی، می‌توان عملکرد سطح ملی مجموعه سازمان‌ها را تفسیر و ارزیابی نمود.

علاوه بر تعیین سطح ارزیابی باید توجه داشت دستگاه‌های مشمول از حیث رسالت سازمانی نیز با یکدیگر تفاوت اساسی دارند. طی مصاحبه با خبرگان، حداقل سه گروه سازمانی: ۱) دانشگاه‌ها و پژوهشگاه‌ها، ۲) دستگاه‌های اجرایی و ۳) شرکت‌های دولتی و بانک‌ها، شناسایی شدند. در این مقاله مدل ارزیابی گونه سازمانی دانشگاه‌ها و پژوهشگاه‌ها در ادامه تبیین و بررسی خواهد شد (نتایج مدل‌سازی برای سایر گونه‌های سازمانی در [۲۵] موجود است).

- گونه‌شناسی فعالیت‌های علم و فناوری بر اساس طبقه‌بندی بانک جهانی

محدوده فعالیت‌های علم و فناوری دانشگاه‌ها و پژوهشگاه‌ها، حوزه پژوهش و توسعه است. با توجه به ماهیت فعالیت‌های پژوهشی در دانشگاه‌ها و پژوهشگاه‌های ایران، از میان طبقه‌بندی‌های مختلف تفکیک پنج‌گانه بانک جهانی مشتمل بر: «علوم طبیعی»، «علوم انسانی»، «علوم اجتماعی»، «علوم سلامت» و «علوم مهندسی» انتخاب شده است. ممکن است فعالیت‌های برخی از دانشگاه‌ها و پژوهشگاه‌ها ترکیبی از این حوزه‌ها را پوشش دهد که در آن‌صورت شاخص‌های همه حوزه‌های تحت پوشش برای ارزیابی عملکرد پژوهش و توسعه آن سازمان مورد نیاز خواهد بود.

- ارزیابی کارایی، اثربخشی و سودمندی در دانشگاه‌ها و پژوهشگاه‌ها

بر اساس آنچه در جداول ۵ تا ۹ ارائه شد ابتدا مجموعه شاخص‌های هر حوزه تخصصی به تفکیک نوع شاخص، تدوین شده و پس از غربال‌گری اولیه در پرسشنامه-مصاحبه با خبرگان هر حوزه، از میان آنها شاخص‌های مناسب برای دانشگاه‌ها و پژوهشگاه‌های کشور انتخاب شدند. برخی از شاخص‌های برون‌دادی، پیامدی و اثرگذاری میان همه یا تعدادی از حوزه‌های تخصصی مشترک بوده، اما در مرحله وزن‌دهی و اولویت‌بندی شاخص‌ها تفاوت معناداری مشاهده می‌شود. به عنوان مثال، برون‌دادهای انتشاراتی همچون مقاله در هر پنج گونه وجود دارد، اما با توجه به ماهیت اغلب بنیادی پژوهش‌های حوزه علوم طبیعی نسبت به کاربردی بودن پژوهش‌های حوزه علوم مهندسی، اولویت و وزن این شاخص در ارزیابی این دو نوع حوزه کاملاً متفاوت است. بدین ترتیب هرگونه سازمانی دسته‌ای شاخص‌های مشترک و دسته‌ای شاخص‌های تخصصی دارد.

پس از محاسبه درصد وزنی امتیاز هر یک از شاخص‌ها از مجموع امتیازات اخذ شده توسط خبرگان، نتایج نهایی شاخص‌های انتخابی در هر ۵ حوزه تخصصی برای دانشگاه‌ها و پژوهشگاه‌ها به شرح جدول ۱۰ جمع‌بندی شده است.

در شکل ۷ نمونه‌ای از نمودار راداری برای دانشگاه‌ها و پژوهشگاه‌های فعال در حوزه علوم طبیعی مشاهده می‌شود. در این شکل، شاخص‌های هر ربع نمودار به همراه درصد وزنی (ضریب تأثیر) آنها مشخص شده است. به منظور ارزیابی سازمان می‌باید حداکثر میزان هر شاخص مشخص شده و میزان شاخص در سازمان مورد نظر نسبت به آن سنجیده خواهد شد.

از تلفیق هر پنج نمودار راداری مربوط به پنج حوزه تخصصی در یک داشبورد واحد، ارزیابی عملکرد سازمان در همه حوزه‌های پژوهشی مقدور خواهد شد (شکل ۸).

www.SID.ir

Archive of SID

جدول ۱۰) شاخص‌های ارزیابی عملکرد دانشگاه‌ها و پژوهشگاه‌ها به تفکیک حوزه‌های پژوهشی

حوزه	شاخص‌های برون‌دادی	وزن (٪)	شاخص‌های پیامدی	وزن (٪)	شاخص‌های اثرگذاری	وزن (٪)
علوم طبیعی	ساخت نمونه اولیه (پایلوت)	۲۵	توسعه روش‌شناسی جدید تست	۳۱	افزایش پرستیژ ملی	۲۵
	مدل‌ها و شبیه‌سازها	۲۳	ایجاد شرکت‌های نوپا	۲۵	ترویج و توسعه علمی	۲۲
	جوایز معتبر	۱۹	درآمدهای پژوهشی	۲۰	افزایش سرمایه‌گذاری در علوم طبیعی	۲۰
	تعداد ارجاعات	۱۷	همکاری با صنعت	۱٤	افزایش امنیت ملی	۲۰
	مقاله (ISC/ISI/علمی-ترویجی/کنفرانس)	۱٦	جذب اعتبار از شخص ثالث	۱۰	افزایش کیفیت زندگی عموم افراد جامعه	۱۳
علوم مهندسی	محصولات، مواد و فرآیندهای جدید یا بهبودیافته	۲۷	ایجاد شرکت‌های نوپا	۲۵	کاهش مصرف انرژی	۲۳
	ثبت اختراع/ نشان تجاری/ برند ملی	۲۵	همکاری با صنعت	۲۰	افزایش کیفیت زندگی عموم افراد جامعه	۲۲
	ساخت نمونه اولیه (پایلوت)	۱۷	درآمدهای پژوهشی	۲۰	محافظت از محیط زیست	۲۲
	گزارش‌های فنی	۱۷	جذب اعتبار از شخص ثالث	۱۸	ترویج و توسعه علمی	۲۰
	کاهش هزینه‌های تولید	۱٤	انتقال دانش و فناوری	۱۷	افزایش سرمایه‌گذاری در بخش علوم مهندسی	۱۳
علوم انسانی	کتاب (درسی)/ فصل کتاب	۳۰	ایجاد شرکت‌های نوپا	۲۸	ترویج و توسعه علمی	۲٤
	بررسی‌های هم‌تراز/ عضویت در شوراهای تخصصی/مشارکت در جوامع یادگیری	۲۵	راه‌اندازی کارهای خلاقانه	۲٤	افزایش سرمایه‌گذاری در بخش علوم انسانی	۲٤
	کارهای کتاب‌شناسی عمده/ کارهای مروری سالیانه در یک حوزه	۱۵	درآمدهای پژوهشی	۱۸	حفظ میراث فرهنگی	۲۲
	حضور در رسانه‌ها/ سخنرانی‌ها	۱۲	اجرای موسیقی/ تئاتر	۱۷	تغییر فهم عمومی مردم	۱٦
	ویژه‌نگاشت یا رساله	۹	آئین‌نامه‌ها و سیاست‌های جدید یا تغییر یافته	۱۳	تغییر سیاست‌های عمومی	۱٤
علوم اجتماعی	مقاله	۲۳	ایجاد مکتب فکری جدید	۳۲	تغییر فهم عمومی مردم	۲۳
	ارجاعات	۱۹	درآمدهای پژوهشی	۲۸	افزایش کیفیت زندگی عموم افراد جامعه	۲۲
	کتاب	۱۷	همکاری با صنعت و بازار	۱۵	ترویج و توسعه علمی	۱۸
	همکاری‌های بین‌المللی	۱٦	ایجاد شرکت‌های نوپا	۱٤	افزایش سرمایه‌گذاری در بخش علوم اجتماعی	۱٤
	پایگاه داده	۱۳	همکاری در سیاست‌گذاری	۱۱	افزایش پرستیژ ملی	۱۲

www.SID.ir

Archive of SID

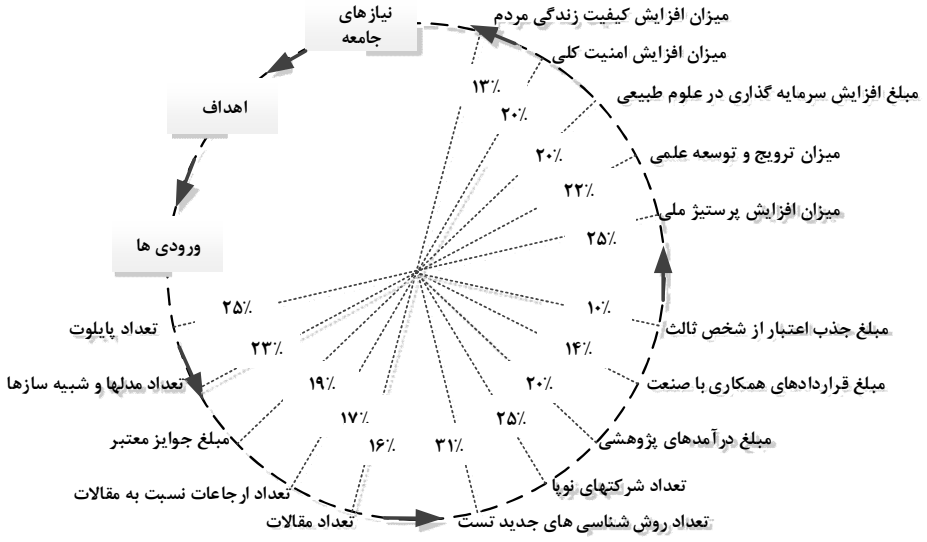
جدول ۱۰) شاخص‌های ارزیابی عملکرد دانشگاه‌ها و پژوهشگاه‌ها به تفکیک حوزه‌های پژوهشی

حوزه	شاخص‌های برون‌دادی	وزن (٪)	شاخص‌های پیامدی	وزن (٪)	شاخص‌های اثرگذاری	وزن (٪)
علوم سلامت	ساخت نمونه اولیه (پایلوت)	۳۲	توسعه استانداردهای جدید سلامت	۳۳	میزان افزایش امید به زندگی باکیفیت یا طول عمر	۲۳
	راهنماها، رویه‌ها و دستورالعمل‌ها	۲۳	ایجاد شرکت‌های نوپا	۲٤	میزان کاهش شیوع بیماری‌ها	۲۲
	انتشارات حرفه‌ای	۱٦	درآمدهای پژوهشی	۱۹	افزایش سرمایه‌گذاری در بخش علوم سلامت	۲۱
	بانک اطلاعاتی، نرم‌افزار، الگوریتم، ابزار سنجش	۱۵	کاهش هزینه‌های سلامت	۱۳	کاهش میزان ابتلاء به بیماری‌ها	۱۹
	عضویت در شوراهای تخصصی/ مشارکت در جوامع یادگیری	۱٤	استفاده حرفه‌ای از رویه‌ها، دستورالعمل‌ها و راهنماها	۱۱	کاهش میزان هزینه‌کرد بیماری‌ها	۱۵

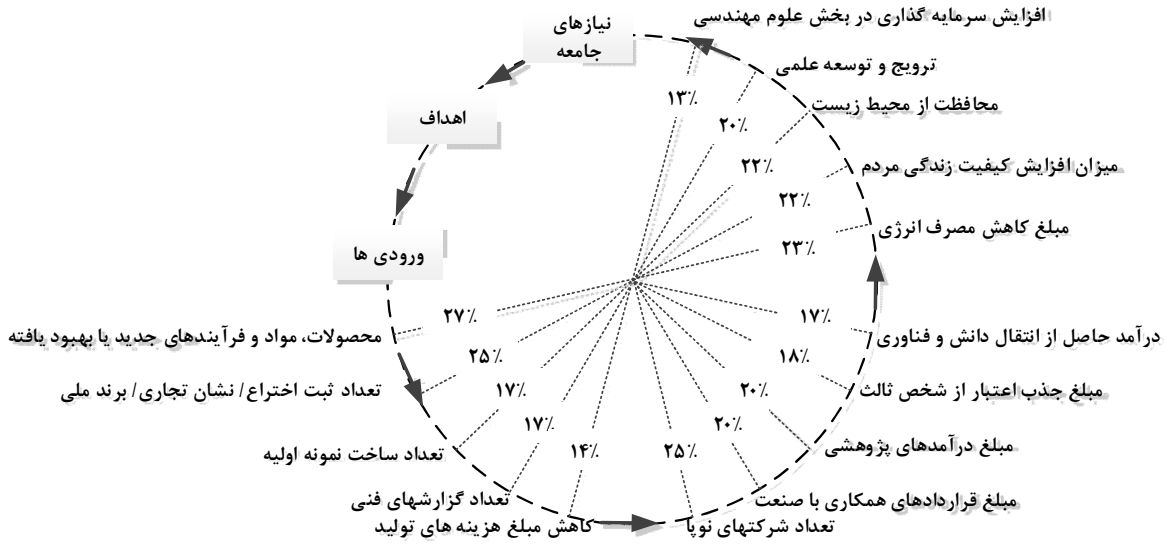
شکل ۷) نمودار راداری ارزیابی عملکرد دانشگاه‌ها و پژوهشگاه‌ها در حوزه علوم طبیعی

www.SID.ir

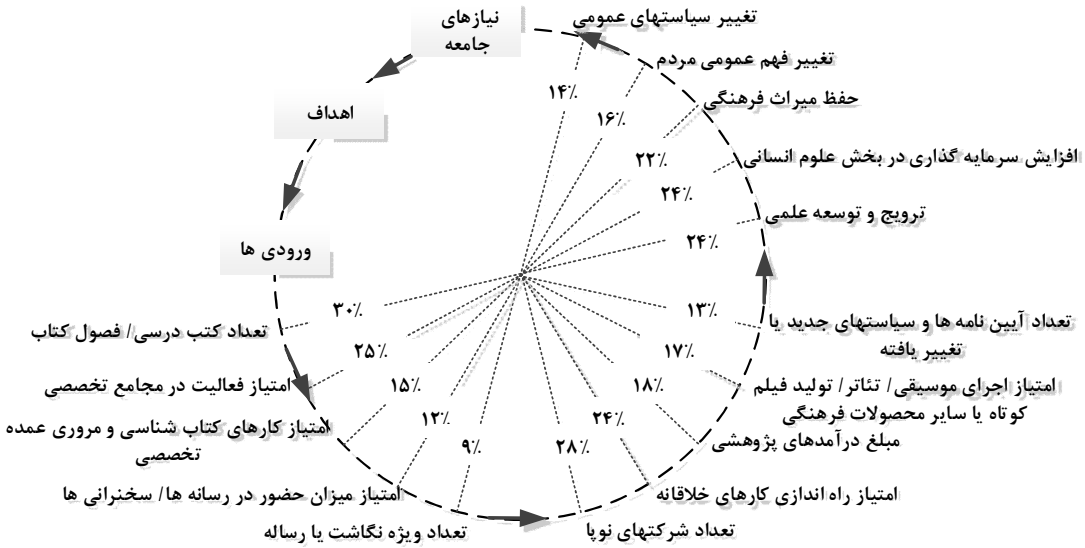
Archive of SID



حوزه تخصصی علوم طبیعی



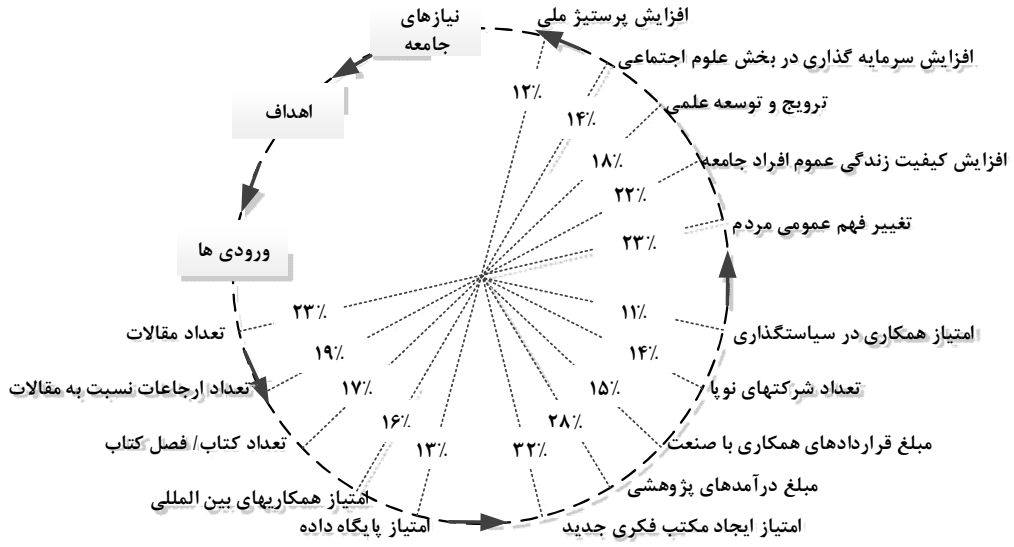
حوزه تخصصی علوم مهندسی



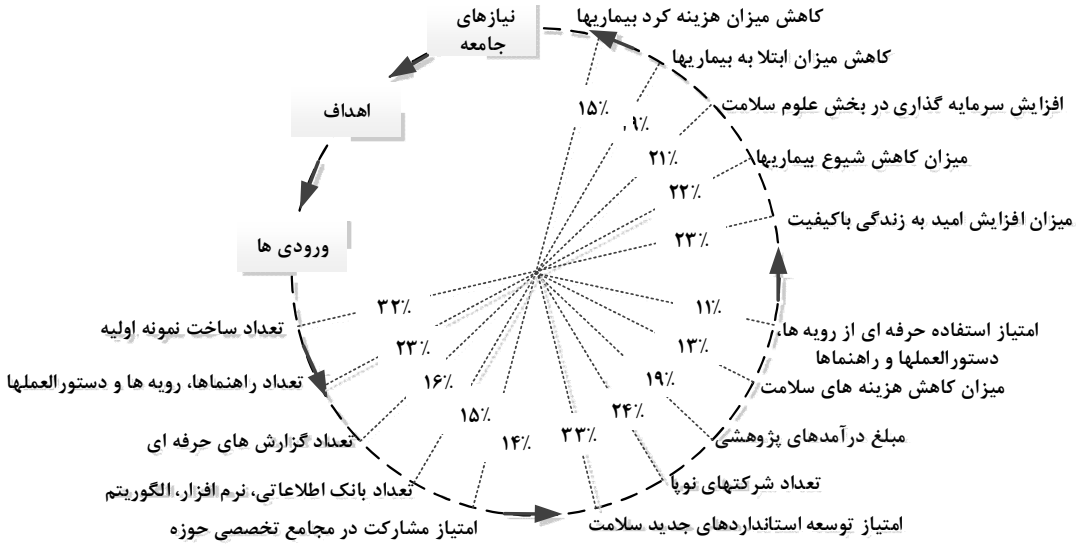
حوزه تخصصی علوم انسانی

www.SID.ir

Archive of SID



حوزه تخصصی علوم اجتماعی



حوزه تخصصی علوم سلامت

شکل ۸) داشبورد ارزیابی عملکرد پژوهشی دانشگاه‌ها و پژوهشگاه‌ها

به منظور ارزیابی کارایی، اثربخشی و سودمندی عملکرد پژوهشی دانشگاه‌ها و پژوهشگاه‌ها ضرایب کارایی، اثربخشی و سودمندی به شرح ذیل استفاده می‌شود:

$$\text{ضریب کارایی} = \left(\sum_{j=1}^{5}\sum_{k=1}^{5}(wp_k \cdot Ip_k)_j\right)/B_i \quad (1)\ \text{معادله}$$

$$\text{ضریب اثربخشی} = \left(\sum_{j=1}^{5}\sum_{k=1}^{5}(wc_k \cdot Ic_k)_j\right)/B_i \quad (2)\ \text{معادله}$$

$$\text{ضریب سودمندی} = \left(\sum_{j=1}^{5}\sum_{k=1}^{5}(wi_k \cdot Ii_k)_j\right)/B_i \quad (3)\ \text{معادله}$$

در معادله (۱)، wp[1] ضریب وزنی و Ip[2] مقدار شاخص‌های برون‌دادی در حوزه پژوهشی است. در معادله (۲)، wc[3] ضریب وزنی و Ip[4] مقدار شاخص پیامدی در حوزه پژوهشی است و در معادله (۳)، wi[5] ضریب وزنی و Ii[6] مقدار شاخص اثرگذاری در حوزه پژوهشی است. در هر سه معادله بالا، B[7] مقدار بودجه سازمان، اندیس i سازمان مربوطه، اندیس j (از

1- Weight of Output Indicator
2- Output Indicator
3- Weight of Outcome Indicator
4- Outcome Indicator
5- Weight of Impact Indicator
6- Impact Indicator
7- Organization Budget

www.SID.ir

Archive of SID

یک تا ۵) شماره حوزه پژوهشی مربوطه در صورت وجود بیش از یک حوزه فعالیت برای سازمان، و اندیس k (از یک تا ۵) شماره شاخص مربوطه در حوزه تخصصی jام، را نشان می‌دهد. بر مبنای ضرایب به دست آمده از معادله‌های فوق، امکان رتبه‌بندی سازمان‌های مختلف از حیث کارایی، اثربخشی و سودمندی فراهم شده است و البته می‌توان آن را در مورد هر زیرمجموعه سازمانی (مثل دانشکده‌ها و گروه‌های آموزشی و حتی اشخاص محقق) به کار گرفت.

پس از محاسبه این ضرایب، در دو سطح می‌توان عملکرد یک دانشگاه یا پژوهشگاه را ارزیابی نمود: ۱) تغییر عملکرد سازمان نسبت به عملکرد پیشین آن (مقایسه با خود)، و ۲) عملکرد نسبی سازمان در طبقه مربوط به آن (مقایسه با گروه رقیب نزدیک). با استفاده از این نتایج، رتبه‌بندی سازمان از ابعاد مختلف نیز قابل انجام است که می‌تواند مبنای سیاست‌گذاری‌های آتی یا تخصیص منابع پژوهش و فناوری به هر سازمان قرار گیرد.

References منابع

[1] Hansson, F. (2006). Organizational use of evaluations governance and control in research evaluation. Evaluation, 12(2), 159-178.

[2] Frederiksen, L. F., Hansson, F., & Wenneberg, S. B. (2003). The Agora and the role of research evaluation. Evaluation, 9(2), 149-172.

[3] 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.

[4] Godin, B. (2004). Measurement and Statistics on Science and Technology: 1920 to the Present. Routledge.

[5] UNESCO. (2010). UNESCO Science Report: The Current Status of Science around the World. Paris.

[6] Aitsi-Selmi, A., Murray, V., Wannous, C., Dickinson, C., Johnston, D., Kawasaki, A., ... & Yeung, T. (2016). Reflections on a science and technology agenda for 21st century disaster risk reduction. International Journal of Disaster Risk Science, 7(1), 1-29.

[7] Molas-Gallart, J. (2012). Research governance and the role of evaluation: A comparative study. American Journal of Evaluation, 33(4), 583-598.

[8] Barker, K. (2007). The UK Research Assessment Exercise: The evolution of a national research evaluation system. Research Evaluation, 16(1), 3-12.

[9] Chelimsky, E. (2006). The purposes of evaluation in a democratic society. The SAGE Handbook of Evaluation, 33-55.

[10] Cruz-Castro, L., & Sanz-Menéndez, L. (2007). Research evaluation in transition. In The Changing Governance of the Sciences (pp. 205-223). Springer Netherlands.

[11] Coryn, C. L., Hattie, J. A., Scriven, M., & Hartmann, D. J. (2007). Models and mechanisms for evaluating government-funded research: An international comparison. American Journal of Evaluation, 28(4), 437-457.

[12] Bazargan Harandi, A. (2004). Validation in Higher Education. Encyclopedia of Higher Education, 1(1), 163-166. {In Persian}.

[13] Hansson, F. (2003). How to evaluate and select new scientific knowledge?. Vest. a Journal for Science and Technology Studies.

[14] Hansen, H. F. (2009). Research Evaluation: Methods, Practice, and Experience. Danish Agency for Science, Technology and Innovation.

[15] Hansen, H. F. (2005). Choosing evaluation models A discussion on evaluation design. Evaluation, 11(4), 447-462.

[16] Rossi, P. H., Lipsey M. W., & Freeman, H. E. (2004). Evaluation: A Systematic Approach. Thousand Oaks, CA: SAGE.

[17] Mintzberg, H. (1983). Structure in Fives: Designing Effective Organizations. Englewood Cliffs, NJ: Prentice-Hall.

[18] Jørgensen, T. B. (2003). 'Værdier i harmoni, konflikt og forandring'. In Jørgensen, T. B. (ed.), På sporet af en offentlig identitet, pp. 63-80. Aarhus: Aarhus Universitetsforlag.

[19] Stame, N. (2004). Theory-Based Evaluation and Varieties of Complexity. Evaluation, 10(1), 58-76.

[20] Karimi, T. (2005). New Models of Organizational Performance Evaluation. Tadbir, 17(171), 22-27. {In Persian}.

[21] Amiran, H. (2003). Performance Evaluation: Dos and Don'ts. Tadbir, 14(136), 77-80. {In Persian}.

[22] Brown, M. G., & Svenson, R. A. (1998). RTM Classic: Measuring R&D Productivity. Research-Technology Management, 41(6), 30-35.

[23] European Commission. (2014). Indicators for monitoring and evaluation: a practical guide.

[24] 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. {In Persian}.

[25] Farazkish, M. (2017). Designing an evaluation model of science, technology and innovation process for Iranian government organizations. Ph.D. thesis in science and technology policy, Faculty of Management and Economics, Tarbiat Modarres University. {In Persian}.

[26] Nowruzi Chakly, A., Ghazavi, R., & Taheri, B. (2015). Valuation of research indicators for different fields of science in Iran. Journal of Science and Technology Policy, 7(4), 31-40. {In Persian}.

[27] The World Bank. (2004). World Development Indicators. Retrieved May 2012, from http://data.worldbank.org/indicator.

[28] Ghazinoory, S., & Farazkish, M. (2017). Innovation process evaluation of national organizations with an integrated approach. Collections of papers and negotiations. Tehran: Academy of sciences. pp. 128-

www.SID.ir

Archive of SID

C. (2006). Requirements engineering paper classification and evaluation criteria: a proposal and a discussion. Requirements Engineering, 11(1), 102-107.
[36] CCA (Council of Canadian Academies). (2012). Informing Research Choices: Indicators and Judgment (The Expert Panel on Science Performance and Research Funding).
[37] CCA (Council of Canadian Academies). (2012). The State of Science and Technology in Canada 2012 (The Expert Panel on the State of Science and Technology in Canada).
[38] Ochester, M. (2012). Indicators for Research Quality for Evaluation of Humanities Research: Opportunities and Limitations. BAND (1).
[39] Marjanovic, S., Hanney, S., & Wooding, S. (2009). A Historical Reflection on Research Evaluation Studies, Their Recurrent Themes and Challenges. Technical Report. RAND Corporation.
[40] NHMRC. (2013). National Health and Medical Research Council Submission to the Productivity Commission's Research Study on Public Support for Science and Innovation in Australia.
137. {In Persian}.
[29] Ghanei Rad, M. A., Mahmoudi, M., & Ebrahim Abadi, H. (2018). A multi-dimensional model of humanities development evaluation. Collections of papers and negotiations. Journal of Science and Technology Policy, 9(1), 85-103. {In Persian}.
[30] Chien, C. F., Chen, C. P., & Chen, C. H. (2009). Designing performance indices and a novel mechanism for evaluating government R&D projects. Journal of Quality, 16(2), 119.
[31] Fahrenkrog, G., Polt, W., Rojo, J., Tübke, A., & Zinöcker, K. (2002). RTD Evaluation Tool Box: Assessing the Socio-economic Impact of RTD Policies. IPTS Technical Report Series. Seville.
[32] Street, D. (2009). Developing key performance indicators for humanities. Royal Irish Academy.
[33] Royal Netherlands Academy of Arts and Sciences. (2013). Quality Indicators in the Social Sciences.
[34] European Commission. (2003). Third report on scientific and technological indicators 2003. Office for official publications of the European communities.
[35] Wieringa, R., Maiden, N., Mead, N., & Rolland,

www.SID.ir