

A Correlation Study of Bibliometric-Based and Informed-Peer-Review University Rankings: The Case of UK Research Excellence Framework (REF) and the World's Prestigious University Ranking Systems

Somayeh Hesabi¹

Hajar Sotudeh^{2*}

Zahra Yousefi³

-  1. M. A. in Knowledge & Information Sciences, Faculty of Education & Psychology, Shiraz University. Shiraz, Iran.
Email: s.hesabbi@gmail.com
-  2. Professor, Department of Knowledge & Information Sciences, Faculty of Education & Psychology, Shiraz University. Shiraz, Iran.
(Corresponding Author)
-  3. Assistant Professor, Department of Knowledge & Information Sciences, Faculty of Literature & Humanities, Persian Gulf University, Bushehr, Iran. Email: z.yousefi@gmail.com

Email: sotudeh@shirazu.ac.ir

Abstract

Purpose: Informed peer-review university ranking systems like REF involve peer review to achieve high-quality university performance evaluations while considering bibliometric facts. Despite its advantages and methodological success, it is not possible to implement the review-based exercise in all countries due to unaffordable costs, vulnerability to bias, and differences in cultural, economic, managerial, and infrastructural conditions. As an alternative, some science systems prefer to consider using the results of international university rankings to gauge the academic performance of their higher education institutions. Using a different methodology, the international ranking systems tend to maintain their efficiency by heavily relying on the bibliometric information extracted from databases, the performance data gathered from official authorities, and academic and employer surveys. This gives rise to the question of to what extent these two types of systems are convergent in their results. If their different methodologies lead to similar results, they could be interpreted to have similar effectiveness in their evaluation of academic performance. Therefore, by relying on the results of the international ranking systems, one can avoid the shortcomings of the review-based method, and maintain both the efficiency and effectiveness of the evaluation systems.

To reveal the convergence of the results obtained from the methods, the present study explores a sample of British universities evaluated by REF (2014) to investigate the correlation between their scores in REF and the world's prestigious university rankings.

Methodology: Using a quantitative content analysis method, the present study concentrates on a collection of 150 British universities evaluated simultaneously by REF (2014) and at least one of the world's prestigious ranking systems including QS, THE, Leiden and ARWU. Due to the small size of the population, all REF members are examined without sampling. The

Date of Reception:
21/04/2021

Date of Acceptation:
13/09/2021



Somayeh Hesabi ¹

Hajar Sotudeh ^{2*}

Zahra Yousefi ³

Date of Reception:
21/04/2021

Date of Acceptation:
13/09/2021



evaluation results of these universities are extracted from these systems and entered into a checklist. Also, the subject fields and disciplines covered by the universities are collected to investigate their probable effects on the results. The universities' subject coverage similarity is calculated using the Cosine similarity measure and K Nearest Neighbor technique in the KNIME data mining platform. Finally, correlation and regression analyses are applied to analyze the data.

Findings: UK universities' scores in REF are found to be significantly correlated to theirs in international ranking systems. They are highly correlated to QS and THE's, while being moderately associated with ARWU's, and weakly-to-strongly correlated to Leiden's. The regression analyses show no significant effects of subject coverage on the overall scores, except for medical tendency's effect on QS. However, the subject coverage affects some dimension scores. While it does not significantly predict any of the dimensions of ARWU, it partially predicts the Citations and Industry Income dimensions in THE, with medical and technical, and engineering subjects respectively having the highest positive predicting power. In QS, the subject coverage partially predicts the dimensions of Academic reputation, Faculty/student ratio, International faculty ratio, International student ratio, and Citation Per faculty. Medical science has the highest positive effect on the dimensions of Academic reputation, Faculty/student ratio, and Citations per faculty. Moreover, basic sciences have the highest negative effect on the International student ratio and International faculty ratio.

Also, the subject coverage can predict Leiden scores for the dimensions PP (top 10%) (i.e. the proportion of a university's publications that, compared with other publications in the same field and the same year, belong to the top 10% most frequently cited), MCS (i.e. the average number of citations of the publications of a university), MNCS (i.e. the average number of citations of the publications of a university, normalized for field and publication year), PP(collab) (i.e. the proportion of a university's publications that have been co-authored with one or more other organizations), PP(int collab) (i.e. the proportion of a university's publications that have been co-authored by two or more countries), PP(industry) (i.e. the proportion of a university's publications that have been co-authored with one or more industrial organizations) and PP(>1000 km) (i.e. the proportion of a university's publications with a geographical collaboration distance of more than 1000 km). Medical science has the highest positive predictive power for the scores in all the mentioned dimensions except for PP(industry).

Conclusion: The results of REF, as a peer-review-based university ranking informed with bibliometric data, are highly correlated to those of the international evaluations based on performance statistics and enriched by surveys, while being moderately correlated to those performed by just performance statistics. The subject coverage impact on the rankings challenges the application of the results in comparing universities with different subject coverages.

Keywords: Research Evaluation, University Rankings, REF, THE, QS, ARWU, Leiden.

همبستگی بین رتبه‌بندی دانشگاه‌ها به دو روش کتاب‌سنجی محض و داوری غنی‌شده با کتاب‌سنجی: نمونه مورد مطالعه نظام تعالی پژوهش انگلیس و نظام‌های رتبه‌بندی معتبر جهانی

سمیه حسابی^۱

۱. کارشناس ارشد علم اطلاعات و دانش‌شناسی، دانشگاه شیراز. Email: s.hesabbi@gmail.com

هاجر ستوده^{۲*}

۲. استاد، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم تربیتی و روان‌شناسی، دانشگاه شیراز. (نویسنده مسئول)

زهرا یوسفی^۳

۳. استادیار، گروه علم اطلاعات و دانش‌شناسی، دانشکده ادبیات و علوم انسانی، دانشگاه خلیج فارس. Email: z.yusefi@gmail.com

Email: sotudeh@shirazu.ac.ir

چکیده

هدف: برخی رتبه‌بندی‌های دانشگاهی مانند چارچوب تعالی پژوهش (رف)، برای دستیابی به کیفیت بالاتر، از روش داوری استفاده می‌کنند. با این حال، کارایی این روش پایین است. پژوهش حاضر به منظور بررسی میزان هم‌گرایی نتایج رتبه‌بندی‌های دارای رویکرد متفاوت به داوری، همبستگی نمره‌های دانشگاه‌های انگلیس را در «رف» و سامانه‌های رتبه‌بندی معتبر لایدن، شانگهای، کیو-اس و تایمز بررسی می‌کند.

روش‌شناسی: پژوهش حاضر، به روش اسنادی با رویکرد تحلیل محتوای کمی انجام شده است. داده‌ها به روش سرشماری گردآوری و با تحلیل همبستگی و رگرسیون تحلیل شده‌اند.

یافته‌ها: قوت همبستگی نمره دانشگاه‌ها در «رف» با «تایمز» و «کیو-اس» قوی، با «شانگهای» متوسط و با ابعاد «لایدن» از ضعیف تا قوی است. نمره کل در «کیو-اس» متأثر از گرایش پزشکی است. تأثیر گرایش موضوعی بر نمره در برخی ابعاد سامانه لایدن مشاهده می‌شود.

نتیجه‌گیری: رتبه‌بندی مبتنی بر داوری غنی‌شده با کتاب‌سنجی، هم‌گرایی قوی و در نتیجه مشابهت بالایی را با روش مبتنی بر تلفیق کتاب‌سنجی با آمار عملکردی و نظرسنجی نشان می‌دهد. اما تلفیق کتاب‌سنجی با آمار عملکردی صرف مشابهت متوسطی با آن دارد و نمی‌تواند جایگزین آن شود. تأثیر گرایش موضوعی بر نتایج رتبه‌بندی، کاربرد این نتایج برای مقایسه دانشگاه‌های دارای پوشش موضوعی متفاوت را به چالش می‌کشد.

واژگان کلیدی: ارزیابی پژوهش، نظام رتبه‌بندی، رف، تایمز، کیو-اس، شانگهای، لایدن.

صفحه ۷۵-۹۸

دریافت: ۱۴۰۰/۰۲/۰۱

پذیرش: ۱۴۰۰/۰۶/۲۲

مقدمه و بیان مسئله

روند رو به رشد پژوهش و اهمیت فزاینده آن در توسعه جوامع دانش‌بنیان، نیازمند ارزیابی و اعتبارسنجی کارآمد پژوهش‌هاست. از این رو، سامانه‌هایی برای ارزشیابی عملکرد دانشگاه‌ها، به‌عنوان خاستگاه اصلی پژوهش و آفرینش‌های علمی، در گستره ملی و جهانی پدیدار شده‌اند. در میان نظام‌های ارزیابی ملی دانشگاه‌ها می‌توان از نظام ارزیابی انگلستان، نیوزلند، هلند، ایتالیا و استرالیا (Wilsdon, 2016; Goldfinch & Yamamoto, 2012; Geuna & Martin, 2003) و در سطح جهانی، از سامانه‌های رتبه‌بندی «شانگهای»^۱، «تایمز»^۲، «کیو-اس»^۳ و «لایدن»^۴ نام برد. دانشگاه‌ها در این سامانه‌ها بر پایه ابعاد عملکردی مختلف مانند آموزش، پژوهش، ارتباط با صنعت، یا دستاوردهای این فعالیت‌ها، مانند شهرت و چشم‌انداز جهانی، درآمد حاصل از همکاری با صنعت ارزیابی می‌شوند. تفاوت در ابعاد عملکردی در سامانه‌های رتبه‌بندی به تفاوت در روش‌های آنها می‌انجامد. برخی مانند لایدن تنها بر آمار کتاب‌سنجی، یعنی برون‌دادهای عینی پژوهشی برگرفته از پایگاه‌های استنادی تمرکز دارند.^۵ برخی دیگر، مانند «شانگهای» علاوه بر آمار کتاب‌سنجی، دیگر ابعاد عملکرد دانشگاه‌ها را بر پایه آمارهای عینی گردآمده از منابع رسمی تحلیل می‌کنند.^۶ در دسته‌ای دیگر، مانند «تایمز» و «کیو-اس»، داده‌های عملکردی و کتاب‌سنجی با نظرسنجی از محققان غنی می‌شود.^۷ برخی نیز می‌کوشند به کمک داوری، به‌تنهایی یا در تلفیق با داده‌های عینی، به ارزیابی کیفیت پژوهش دست یابند.^۸

چارچوب تعالی پژوهش انگلستان (رف)^۹ از معدود نظام‌های مبتنی بر تلفیق داوری و کتاب‌سنجی است. این نظام، عملکرد دانشگاه‌ها را در سه بعد برون‌داد پژوهشی، محیط پژوهشی و تأثیر ارزیابی می‌کند. این نظام از دستاوردهای موفق در ارزیابی پژوهش برخوردار بوده است (Stern, 2016; Geuna & Piolatto, 2016; Aksnes & Taxt, 2004). استواربودن بر اصل انصاف، برابری و شفافیت از جمله مهم‌ترین ویژگی‌های «رف» است. همچنین، تلفیق دو روش کتاب‌سنجی و داوری تخصصی نه تنها ربط، اصالت، کیفیت و تأثیر بالقوه اقتصادی-اجتماعی برون‌دادهای پژوهشی را کنترل می‌کند، بلکه نظرات متخصصان به کمک اطلاعات کتاب‌سنجی از دقت و عینیت بالاتری برخوردار می‌شود (Robbins, Wield & Wilson, 2017; De Boer et al, 2015; Abramo, D'Angelo, 2011; Box, 2010). از دیگر امتیازات «رف»، سنجش تأثیر اجتماعی پژوهش است.^{۱۰}

با وجود مزایای بسیار و موفقیت روش شناختی «رف» (Stern, 2016; Geuna & Piolatto, 2016)، آشکار است که اجرای این تجربه برای همه کشورها با سامانه‌های علمی و پژوهشی متفاوت به‌سادگی امکان‌پذیر نیست؛ زیرا از یک سو، روش داوری نه تنها با چالش‌های بسیاری به‌ویژه سوگیری و تأثیرپذیری از ذهنیت و سلیقه شخصی داوران روبه‌رو است (Taylor, 2011)، بلکه در شرایط فرهنگی، اقتصادی، مدیریتی و زیرساختی متفاوت لزوماً به پاسخی یکسان و رضایت‌بخش منجر نمی‌شود. از سوی دیگر، هزینه‌های بسیار سنگین داوری باعث می‌شود این گونه نظام‌ها به لحاظ اقتصادی مقرون‌به‌صرفه نباشند (Geuna & Piolatto, 2016; Franceschet, Costantini, 2011).

1. Academic Ranking of World Universities (ARWU)
2. Times Higher Education (THE)
3. Quacquarelli Symonds (QS)
4. Leiden
5. <http://www.leidenranking.com>
6. <http://www.shanghairanking.com>
7. <https://www.timeshighereducation.com> , <https://www.topuniversities.com/university-rankings>
8. <http://www.ref.ac.uk>
9. Research Excellence Framework (REF)
10. <http://www.ref.ac.uk/about> , <http://www.sussex.ac.uk/ref/>, <http://www.hefce.ac.uk/rsrch/ref2021>

(Chubin, 1994). بنابراین، به جای ایجاد نظام ارزیابی ملی پرهزینه، می‌توان بهره‌گیری از نتایج رتبه‌بندی‌های بین‌المللی را به عنوان راهکاری جایگزین بررسی کرد. چنانچه این دو نوع نظام به نتایجی مشابه منجر شوند، ضمن صرفه‌جویی در هزینه‌ها می‌توان از کاستی‌های روش مبتنی بر داوری نیز مصون ماند. نظر به اهمیت این امر، پژوهش حاضر می‌کوشد به این پرسش اساسی پاسخ گوید که آیا رویکردهای مبتنی بر تلفیق روش داوری و کتاب‌سنجی در مقایسه با رویکردهای مبتنی بر داده‌های کتاب‌سنجی و نظرسنجی به نتایج مشابهی دست می‌یابند. بدین منظور، با تمرکز بر آخرین نتایج ارزیابی «رف» ۲۰۱۴، همبستگی میان نمره‌های دانشگاه‌های انگلیسی در این نظام و نظام‌های معتبر رتبه‌بندی بررسی می‌شود تا از این رهگذر میزان واگرایی یا هم‌گرایی نتایج حاصل از روش‌های ارزیابی متفاوت آشکار شود.

این پژوهش، بر چهار نظام «شانگهای»، «تایمز»، «کیو-اس» و «لایدن» که از شناخته‌ترین نظام‌های رتبه‌بندی دانشگاه‌ها هستند تمرکز دارد. هیچ‌یک از نظام‌های یادشده، در ارزیابی از روش داوری استفاده نمی‌کنند، بلکه از داده‌های کتاب‌سنجی و آمار عملکردی دانشگاه‌ها بهره می‌برند (به‌استثنای سامانه «لایدن» که تنها بر روش کتاب‌سنجی تکیه دارد). «تایمز» و «کیو-اس» از نظرسنجی نیز بهره می‌برند. انتخاب نظام‌هایی با روش‌شناسی‌های متفاوت کمک می‌کند که اثربخشی روش تلفیق داوری و کتاب‌سنجی با نظام‌هایی که از داوری بهره نمی‌گیرند، به محک آزمون گذاشته شود. نکته مهم دیگر، تأثیر پوشش موضوعی دانشگاه‌هاست. دانشگاه‌های صنعتی و تخصصی نسبت به دانشگاه‌های جامع، به لحاظ گرایش موضوعی همگن‌تر و در نتیجه دارای تفاوت‌های بین‌رشته‌ای کمتری هستند. دانشگاه‌های صنعتی و آنهایی که بیشتر بر علوم و مهندسی تمرکز دارند - به دلیل نزدیکی بیشتر به علوم سخت - در مقایسه با دانشگاه‌هایی با گرایش علوم انسانی و اجتماعی، غالباً هماهنگی بیشتری با روش‌های کتاب‌سنجی دارند (Geuna & Piolatto, 2016; Mryglod, Kenna, Holovatch & Berche, 2013; Abramo, D'Angelo, 2011). بنابراین، پرسش دیگری که در اینجا مطرح می‌شود میزان تأثیر گرایش‌های موضوعی بر نتایج این ارزیابی‌هاست.

سؤال‌های پژوهش

- ۱- آیا میان نمره دانشگاه‌ها در نظام «رف» و نمره آنها در هر یک از نظام‌های رتبه‌بندی معتبر جهانی («تایمز»، «کیواس»، و «شانگهای») همبستگی معناداری وجود دارد؟
- ۲- آیا میان نمره دانشگاه‌ها در نظام «رف» و نمره آنها در هر یک از ابعاد نظام‌های رتبه‌بندی معتبر جهانی همبستگی معناداری وجود دارد؟
- ۳- آیا گرایش موضوعی تأثیری معنادار بر نمره دانشگاه‌ها در نظام «رف» و نظام‌های رتبه‌بندی معتبر جهانی دارد؟
- ۴- گرایش‌های موضوعی دانشگاه‌ها بر نمره آنها در کدام یک از ابعاد نظام‌های رتبه‌بندی معتبر جهانی تأثیرگذار است؟

چارچوب نظری

نظام‌های ارزیابی پژوهش از رویکردهای متنوعی استفاده می‌کنند. در رویکرد بودجه‌بندی مبتنی بر عملکرد^۱، تخصیص و توزیع بودجه بر اساس عملکرد گذشته مؤسسه یا فرد در ابعاد گوناگون مانند آموزش، پژوهش، ارتباط با

1 . <http://www.shanghai ranking.com/ARWU2017.html>
2 . performance-based research funding systems (PBRFS)

همبستگی بین رتبه‌بندی دانشگاه‌ها به دو روش کتاب‌سنجی محض و داوری غنی‌شده با کتاب‌سنجی...

صنعت، تأثیرگذاری بر پیشبرد علم و اجتماع، و همچنین تأثیرات اقتصادی، اجتماعی، رفاهی و بهداشت و سلامت انجام می‌شود. Jonkers & Zacharewicz, 2016; De Boer et al, 2015; Hicks, 2012; Butler, 2010; Box, (2010). ارزیابی مبتنی بر عملکرد که به هدف ارتقای کیفیت در بخش آموزش عالی اجرا می‌شود، سیستم‌هایی پیچیده و پویا را بر اساس اصولی مهم مانند برقراری تعادل بین داوری و شاخص‌های کمی، به هنجارسازی تفاوت‌های میان‌رشته‌ای، و شفافیت در داده‌ها و نتایج بنا می‌نهد (De Boer et al, 2015; Hicks, 2012; Box, 2010). نتایج چنین نظام‌هایی، گذشته از کمک به توزیع منطقی بودجه میان دانشگاه‌ها، می‌تواند با ایجاد رقابت برای کسب اعتبار، انگیزه‌هایی قوی در سیستم‌های دانشگاهی ایجاد کند (Jonkers & Zacharewicz, 2016; Hicks, 2012; Butler, 2010).

به‌طور کلی، نظام‌های ارزیابی مبتنی بر عملکرد بر سه رویکرد عمده استوار است. ارزیابی برون‌دادهای تحقیق ممکن است از طریق کتاب‌سنجی (کمیت انتشارات و استناد) یا داوری (ارزیابی کیفی طیف وسیعی از برون‌دادها و فعالیت‌های پژوهشی) انجام شود. روش مبتنی بر کتاب‌سنجی ضعف‌هایی نظیر تمرکز بیشتر بر برخی انواع برون‌دادهای پژوهشی خاص، تأکید بر کمیت به جای کیفیت، نادیده‌گرفتن تفاوت‌های میان‌رشته‌ای و تفاوت‌های پایگاهی، احتمال افزایش خوداستنادی، خطاهای فنی و محتوایی، ناپایداری در طول زمان، و سوگیری انگلیسی-امریکایی دارد (Hood, Wilson, 2001; Kostoff, 1995; Oberski, 1988). روش داوری نیز از معایبی از جمله کندی و در نتیجه طولانی‌شدن فرایند داوری، کارایی پایین و هزینه بالا (Phillips, Maes, 2012; Taylor, 2011) به‌ویژه در ارزیابی حجم بالایی از آثار (Priem, Hemminger, 2010)، کمبود داوران متخصص و سوگیری (Phillips, Maes, 2012; Chubin, Hackett, 1990) متأثر می‌شود. در رویکرد سوم که رویکردی تلفیقی است، آرای داوران با شاخص‌های کتاب‌سنجی غنی و تلاش می‌شود که از مزایای هر دو روش بهره‌گرفته شود. باور بر آن است که سامانه‌های مبتنی بر داوری همتایان در مقایسه با سامانه‌های مبتنی بر کتاب‌سنجی که طیف وسیعی از برون‌دادها و فعالیت‌های پژوهشی را به روش کمی ارزیابی می‌کنند، انعطاف‌پذیری بیشتری دارند. در مقابل، سامانه‌های کتاب‌سنجی ارزان‌تر بوده و به انطباق کمتری نیاز دارند (De Boer, 2015; Box, 2010). از این‌رو، ترکیب این دو رویکرد، ارزیابی منطقی‌تر، اصولی‌تر و منصفانه‌تری در پی خواهد داشت (Stern, 2016; Geuna & Piolatto, 2016; Aksnes & Taxt, 2004).

در دهه گذشته، انگلستان یکی از پیشرفته‌ترین سیستم‌های ارزیابی پژوهش در اروپا را به نام چارچوب تعالی پژوهش^۱ «رف» توسعه داده است. «رف»، از نظام‌های ارزیابی مبتنی بر عملکرد است که موفق به تلفیق دو روش داوری و کتاب‌سنجی شده است. درحقیقت، نتایج ارزیابی داوران متخصص، با اطلاعات استنادی و شاخص‌های کمی مانند ضریب تأثیر مجلات و شاخص اچ^۲ پشتیبانی می‌شود (Geuna & Piolatto, 2016; Abramo, D'Angelo, 2003; Robbins, Wield & Wilson, 2011). «رف» به‌عنوان یک نظام ارزیابی تأثیر (Robbins, Wield & Wilson, 2011) درصدد است تا با رتبه‌بندی واحدهای تحقیقاتی دانشگاه‌ها بر اساس کیفیت، به توزیع منصفانه بودجه کمک نماید. ارزیابی در این نظام توسط هیئت‌های کارشناسی متشکل از دانشگاهیان، اعضای بین‌المللی و کاربران پژوهش انجام می‌شود. درنهایت، مؤسسات بر اساس میانگین نمره کیفیت^۲، رتبه‌بندی می‌شوند. این نمره، ترکیبی از نمرات

1 . REF: Research Excellence Framework

2 . GPA

برون داده‌های پژوهشی (۶۵ درصد)، تأثیر (۲۰ درصد) و محیط پژوهش (۱۵ درصد) است^۱. به دلیل به‌کارگیری روش ترکیبی، بسیاری از محدودیت‌های ذاتی روش مبتنی بر داوری یا کتاب‌سنجی صرف در این نظام به چشم نمی‌خورد. ترکیب این دو رویکرد موجب پوشاندن نقاط ضعف یکی با دیگری شده و به ارزیابی منطقی‌تر و اصولی‌تری می‌انجامد (Stern, 2016; Geuna & Piolatto, 2016). باین‌حال، اجرای تجربه «رف» پرهزینه بوده و در عمل به‌سادگی امکان‌پذیر نیست. در مقایسه با نظام «رف»، نظام‌های بین‌المللی مبتنی بر داده‌های کتاب‌سنجی و نظر‌سنجی، برای دانشگاه‌ها کم‌هزینه‌ترند. گرچه این نظام‌ها در سطح ملی اجرا نمی‌شوند، اما پیوستن شمار هر چه بیشتری از دانشگاه‌های یک کشور می‌تواند به ایجاد یک زیرسامانه ملی در این سامانه‌ها کمک کند و درعین‌حال، امکان مقایسه بین دانشگاهی را نه تنها در سطح ملی، بلکه در سطح منطقه‌ای و بین‌المللی فراهم آورد.

چهار نظام «شانگهای»، «لایدن»، «کیو-اس» و «تایمز» از قدیمی‌ترین و معروف‌ترین نظام‌های ارزیابی دانشگاهی جهانی هستند که از روش‌شناسی‌های متفاوت بهره می‌برند. «شانگهای» از سال ۱۹۹۸ آغاز به کار کرده و هر ساله، به معرفی برترین دانشگاه‌ها از نظر دستاوردهای علمی و پژوهشی می‌پردازد. این سامانه، عمدتاً به پژوهش و کسب جوایز بین‌المللی معطوف است و برای رتبه‌بندی، از معیارهایی نظیر کیفیت آموزش، کیفیت اعضای هیئت علمی، برون‌داد پژوهشی و سرانه عملکرد بهره می‌برد و از نظام‌های مبتنی بر داده‌های عینی عملکرد (شامل داده‌های کتاب‌سنجی و آمار رسمی) است. سامانه «لایدن» سال‌ها بعد در ۲۰۰۷، ارزیابی سالانه عملکرد دانشگاه‌های جهان را بر اساس مجموعه پیچیده‌ای از شاخص‌های کتاب‌سنجی آغاز کرد. «لایدن» تنها بر عملکرد پژوهشی تأکید دارد و آن را به کمک دسته‌ای از شاخص‌های مبتنی بر انتشارات، تأثیر استنادی و همکاری‌های علمی می‌سنجد. رتبه‌بندی «کیو-اس» که از سال ۲۰۰۴ توسط مؤسسه «کاکارلی سیموندز» انجام می‌گیرد، تا پیش از سال ۲۰۱۰، با همکاری تایمز و پس‌از آن، هر ساله به‌صورت مستقل منتشر می‌شود. شهرت دانشگاه، شهرت کارفرمایان، نسبت اعضای هیئت علمی به دانشجو، سرانه استناد و سهم اعضای هیئت علمی بین‌المللی پنج شاخص اصلی آن به شمار می‌آیند. رویکرد «کیو-اس» به ارزیابی و رتبه‌بندی دانشگاه‌ها مبتنی بر داده‌های عینی غنی شده با داده‌های نظر‌سنجی است. نظام رتبه‌بندی «تایمز» نیز از سال ۲۰۱۰ سالانه منتشر شده است و همانند «کیو-اس»، تلفیق دو روش کتاب‌سنجی و نظر‌سنجی را به کار می‌برد^۲. این رتبه‌بندی شامل پنج معیار آموزش، پژوهش، استناد، وجهه بین‌المللی و درآمد صنعتی است.

پیشینه پژوهش

بررسی پژوهش‌ها درباره سامانه‌های رتبه‌بندی، به شناسایی دو گروه آثار مرتبط با مسئله در دست بررسی انجامید. دسته نخست پژوهش‌ها، به مقایسه یا تحلیل همبستگی نتایج سامانه‌های رتبه‌بندی گوناگون پرداخته‌اند. برای مثال، بررسی دانشگاه‌های برتر در سامانه‌های «شانگهای»، «کیو-اس»، «فراایتر نشنال»، «وبومتریکس»، «هییکت»، «تایمز»، و «لایدن» و «سایماگو» حاکی از وجود همبستگی قوی میان نتایج این نظام‌های رتبه‌بندی جهانی است (خسروچردی و زراعت‌کار، ۱۳۹۱؛ خانی‌زاد و منتظر، ۱۳۹۶). همچنین، بررسی همبستگی نمرات دانشگاه‌ها در سامانه‌های رتبه‌بندی نمایه نیچر و لایدن (به‌عنوان سامانه‌های رتبه‌بندی ساده مبتنی بر کتاب‌سنجی) با تایمز و کیو-اس (به‌عنوان سامانه‌های پیچیده مبتنی بر داده‌های کتاب‌سنجی، آمار عملکردی و نظر‌سنجی) نشان از معناداری رابطه میان نمره‌ها در نمایه نیچر و لایدن با تایمز و کیو-اس دارد. میزانی از همسویی میان سامانه‌های رتبه‌بندی ساده و پیچیده به لحاظ تشابه در ابعاد

1 . <http://www.ref.ac.uk>

2 . www.timeshighereducation.co.uk

همبستگی بین رتبه‌بندی دانشگاه‌ها به دو روش کتاب‌سنجی محض و داوری غنی‌شده با کتاب‌سنجی...

و شاخص‌ها به‌ویژه ابعاد مرتبط با بهره‌وری پژوهشی وجود دارد، که البته چندان چشمگیر نیست (مشتاق، ستوده، یقطین و جوکار، ۱۴۰۰). ارزیابی سه رتبه‌بندی اصلی جهانی «تایمز»، «کیو-اس» و «شانگهای» نشان می‌دهند که این سامانه‌ها، با وجود پایداری نسبی در طول زمان تفاوت‌هایی داشته‌اند. نتایج تحلیل مؤلفه‌های اصلی و تحلیل عاملی اکتشافی، حاکی از این است که متغیرهای مورد استفاده در این سامانه‌ها عمدتاً دو عامل اساسی شهرت و بهره‌وری پژوهشی را می‌سنجند. همچنین، این متغیرها ممکن است عیناً همان مفاهیم مورد ادعا را اندازه‌گیری نکنند (Selten, Neylon, Huang & Groth, 2020). با بررسی روش‌شناسی‌های سامانه‌های رتبه‌بندی «کیواس»، «ویومتریکس»، «شانگهای»، «تایمز» و «لایدن» نشان داده شده که دانشگاه‌های نامدار و با سابقه دارای رتبه‌های بالاتری هستند. بخشی از آن احتمالاً به دلیل سوگیری در نظرسنجی‌های جهانی است. همچنین، تفاوت در اهداف اصلی هر دانشگاه در پژوهش، آموزش و خدمات اجتماعی و تنوع برنامه‌ها و دوره‌های ارائه‌شده، به ناسازگاری در سامانه‌های رتبه‌بندی دامن زده است. دانشگاه‌های دارای رتبه پایین، گاهی در برخی ابعاد عملکرد بهتری نسبت به دانشگاه‌های برتر دارند. از طرفی، هیچ سامانه رتبه‌بندی کامل نیست و نیاز به بازبینی و بهبود مستمر آنها در انعکاس مأموریت، چشم‌انداز و اهداف واقعی دانشگاه‌ها وجود دارد. استفاده از نتایج این رتبه‌بندی‌ها می‌بایستی آگاهانه و با پرهیز از اعتماد بیش از اندازه باشد (Fauzi, Tan, Daud, & Awalludin, 2020). با تحلیل تطبیقی ۵ نظام رتبه‌بندی «شانگهای»، «لایدن»، «تایمز»، «کیو-اس» و «یو-مالتی‌رنک»، میزان هم‌پوشانی، روش‌های رتبه‌بندی، نحوه گزینش و محاسبه شاخص‌ها، چولگی توزیع داده‌ها، همبستگی‌های آماری میان شاخص‌ها و تأثیر نرمال‌سازی بر رتبه مؤسسات مورد بررسی قرار گرفته است. نتایج نشان می‌دهد که فقط ۳۵ دانشگاه از ۱۰۰ دانشگاه برتر در این سامانه‌ها مشترک بوده‌اند و قرارگرفتن دانشگاه در رتبه‌های برتر آشکارا به سامانه رتبه‌بندی وابسته است. به این ترتیب، می‌توان چنین نتیجه‌گیری کرد که هر سامانه به جای رتبه‌های موازی جداگانه، می‌بایست دادگان و ابزارهایی را برای مشاهده الگوهای موجود در داده‌های چندوجهی ارائه کند (Moed, 2017). به‌رغم تفاوت در برخی شاخص‌های نظام‌های مختلف رتبه‌بندی جهانی، خروجی نهایی این نظام‌ها تا اندازه‌ای مشابه است (پوریزدیان و کرمی، ۱۳۹۸). تحلیل شباهت‌ها و تفاوت‌های نظام‌های رتبه‌بندی جهانی نشان می‌دهد که با تنگ‌شدن دامنه مقایسه از سطح جهانی به منطقه‌ای، شباهت میان داده‌ها افزایش می‌یابد (نورمحمدی و صفری، ۱۳۹۲).

دسته دوم پژوهش‌های انجام‌گرفته بر نتایج ارزیابی به کمک داوری و کتاب‌سنجی تمرکز کرده‌اند. در این راستا، بررسی نقش شاخص‌های کمی در فرایند ارزیابی پژوهش ۲۰۰۸ انگلستان، وجود همبستگی میان نتایج ارزیابی متخصصان و نتایج کتاب‌سنجی را تأیید نموده است (Taylor, 2011). اما در تحقیقات مختلف، میزان این همبستگی بسته به رشته‌های مختلف، متفاوت گزارش شده است. به‌عنوان نمونه، همبستگی ضعیفی میان سنج‌های کیفیت و تأثیر پژوهشی برای رشته‌های شیمی، فیزیک، مهندسی مکانیک، هوافضا، جغرافیا و مطالعات محیطی، جامعه‌شناسی و تاریخ نشان داده شده است. همچنین، همبستگی میان نمرات داوری و استنادها در علوم نرم ضعیف گزارش شده است. شواهد نشان می‌دهد گرچه استناد می‌تواند نمایانگر توانمندی گروه پژوهشی، به‌ویژه برای علوم سخت باشد، اما تفاوت این شاخص در علوم سخت و نرم باعث می‌شود که رتبه‌بندی کاملاً متفاوت و غیرواقعی‌بینانه‌ای را نسبت به داوری ارائه دهد (Mryglod et al, 2013). مقایسه دو رویکرد داوری و کتاب‌سنجی در ارزیابی پژوهش بر اساس

پارامترهای اساسی دقت، قدرت^۱، اعتبار، کارآمدی (عملکرد)، زمان و هزینه‌ها، نیز نشان از تأثیر تفاوت‌های بین‌رشته‌ای بر همسویی روش داوری و کتاب‌سنجی داشته، چنان‌که در علوم طبیعی و صوری^۲، عملکرد روش کتاب‌سنجی بهتر از داوری نشان داده شده است (Abramo & D'Angelo, 2011). همچنین، بررسی‌ها نشان داده که توافق میان این دو روش در رشته تاریخ در مقایسه با اقتصاد، مدیریت و آمار کمتر است (Bertocchi et al, 2015). البته همه تحقیقات، مؤید وجود رابطه مستقیم میان نتایج حاصل از روش داوری و روش کتاب‌سنجی نیستند، تا بدانجا که Baccini & De Nicolao با بررسی نتایج ارزیابی کتاب‌سنجی در مقابل داوری غنی‌شده در برنامه ارزیابی تحقیقات ایتالیا، به وجود تفاوت میان نتایج حاصل از این دو روش اذعان نموده‌اند (Baccini & De Nicolao, 2016)

در مجموع، می‌توان چنین برداشت کرد که تحقیقات انجام‌شده پیرامون همبستگی نتایج سامانه‌های رتبه‌بندی دانشگاهی نشانگر شباهت‌هایی در برخی ابعاد و شاخص‌هاست که تا اندازه‌ای به هم‌گرایی در نتایج رتبه‌بندی‌ها می‌انجامد. مقایسه نتایج کتاب‌سنجی و داوری نیز نشان می‌دهد که این دو روش می‌توانند تا اندازه‌ای هم‌گرا باشند. این هم‌گرایی بستگی به رشته‌ها دارد. کتاب‌سنجی برای رشته‌های مختلف به یک اندازه قابل کاربرد نیست. به‌ویژه در علوم نرم همبستگی قوی میان نتایج داوری و کتاب‌سنجی دیده نشده است. بر این اساس، آشکار نیست نتایج ارزیابی‌های دانشگاه‌هایی که ترکیب ناهمگنی از رشته‌ها و گرایش‌های موضوعی از علوم سخت تا نرم را دربرمی‌گیرند، در مقایسه با دانشگاه‌های همگن‌تر یا با تخصص‌های خاص‌تر تا چه اندازه متفاوت است.

روش‌شناسی پژوهش

پژوهش حاضر، به روش اسنادی با رویکرد تحلیل محتوای کمی انجام شده است. داده‌ها به روش سرشماری گردآوری شده‌اند. جامعه پژوهش را دانشگاه‌های بریتانیا که هم‌زمان در نظام «رف» ۲۰۱۴ و دست کم در یکی از سامانه‌های رتبه‌بندی معتبر جهانی ارزیابی شده‌اند تشکیل می‌دهد (جدول ۱). به دلیل کوچکی جامعه، همه اعضا بدون نمونه‌گیری بررسی شده‌اند.

جدول ۱. وضعیت دانشگاه‌های انگلیس در «رف» و رتبه‌بندی‌های جهانی

فراوانی دانشگاه‌ها	نظام رتبه‌بندی
۷۰	کیواس
۴۹	تایمز
۴۴	لایدن
۳۷	شانگهای
۱۵۰	نظام «رف»

بررسی اولیه نشان داد ۱۵۰ دانشگاه ارزیابی شده در «رف» ۲۰۱۴، دست کم در یکی از نظام‌های رتبه‌بندی جهانی مورد بررسی وارد شده‌اند. نتایج ارزیابی این دانشگاه‌ها از این نظام‌ها، گردآوری و در یک سیاهه واریسی وارد شد. شایان ذکر است نمره کل برخی دانشگاه‌ها در نظام‌های جهانی گزارش نمی‌شود و رتبه آنها نیز به صورت دامنه‌ای گزارش می‌شود. بنابراین، برای محاسبه نمره کل هر دانشگاه، مجموع نمرات آن در ابعاد مختلف پس از اعمال ضرایب

1 . robustness
2 . natural and formal sciences

همبستگی بین رتبه‌بندی دانشگاه‌ها به دو روش کتاب‌سنجی محض و داوری غنی‌شده با کتاب‌سنجی...

وزنی تعیین شده در روش شناسی هر نظام به دست آمد. برای کنترل اعتبار محاسبات، به نمرات دانشگاه‌هایی که نمره کل آنها به طور دقیق اعلام شده بود مراجعه و تفاوت اندکی میان نمرات به دست آمده با نمرات اعلام شده در نظام رتبه‌بندی به دلیل برخی به‌هنجارسازی‌ها در الگوریتم‌های مربوطه مشاهده شد. برای اطمینان از اعتبار نمرات به دست آمده، همبستگی آنها با نمره‌های گزارش شده در نظام‌ها تحلیل شد. ضرایب همبستگی برای «تایمز»، «شانگهای» و «کیو-اس» (به ترتیب ۱.۰۰۰، ۱.۰۰۰ و ۰.۹۹۹ در سطح اطمینان ۹۹ درصد) حاکی از همبستگی بسیار قوی و مثبت بود. در گام بعد، نمره گرایش موضوعی هر دانشگاه محاسبه شد. بدین منظور، نخست رشته‌های موجود در تایمز (مشمول بر ۴۰ رشته) در چهار گروه موضوعی «هنر، علوم انسانی و اجتماعی»، «علوم پزشکی»، «مهندسی و فناوری»، «علوم پایه و کاربردی» شناسایی شد. سپس، گرایش‌های موضوعی هر دانشگاه از «تایمز» استخراج و به کمک نرم‌افزار داده‌کاوی نایم^۱، میزان مشابهت آنها با هر یک از این چهار گروه موضوعی به دست آمد. مشابهت با استفاده از معیار شباهت کسینوسی^۲ و تکنیک نزدیک‌ترین همسایه^۳ محاسبه شد. از آنجاکه شباهت پوشش موضوعی دانشگاه به هر یک از چهار گروه موضوعی مورد نیاز بود، تعداد نزدیک‌ترین همسایه‌ها برابر با چهار در نظر گرفته شد. به هر دانشگاه چهار نمره بر اساس میزان شباهت آن به هر یک از گروه‌های موضوعی داده شد. داده‌ها در اسپاس ۱۶ و به کمک تحلیل همبستگی و رگرسیون خطی تحلیل شد. البته در روش سرشماری، آمار استنباطی در برآورد پارامترهای جامعه و تعمیم‌پذیری کاربردی ندارد؛ بلکه هدف، اطمینان از دقت و ارزش همبستگی یا تأثیر مشاهده شده است.

یافته‌های پژوهش

پاسخ به سؤال اول پژوهش. آیا میان نمره دانشگاه‌ها در نظام «رف» و نمره آنها در هر یک از نظام‌های رتبه‌بندی معتبر جهانی («تایمز»، «کیواس»، «شانگهای») همبستگی معناداری وجود دارد؟

نتایج آزمون کولموگروف-اسمیرنوف نشان از نرمال بودن توزیع نمرات در «رف» ($P\text{-value} > 0.05$) و نرمال نبودن نمره‌ها در نظام‌های معتبر جهانی ($P\text{-value} < 0.05$) داشت. بنابراین، از تحلیل همبستگی ناپارامتری اسپیرمن استفاده شد.

جدول ۲. همبستگی نمره دانشگاه‌ها در «رف» با نمرات کل آنها در نظام‌های رتبه‌بندی جهانی

نظام‌ها	ضریب همبستگی	تعداد
تایمز	۰.۸۱۸**	۴۹
کیواس	۰.۸۸۲**	۷۰
شانگهای	۰.۵۵۷**	۳۷

**همبستگی در سطح ۰.۰۱ معنادار است (دوسویه)

با توجه به نتایج تحلیل همبستگی، بین نمره دانشگاه‌ها در نظام «رف» و نمره آنها در هر یک از نظام‌های رتبه‌بندی معتبر جهانی («تایمز»، «کیواس»، «شانگهای») همبستگی مثبت و معناداری وجود دارد ($P\text{-value} < 0.01$). شدت

1. KNIME
2. Cosine similarity
3. k-nearest neighbors (KNN)

همبستگی نمره در «رف» با نمره کل در «تایمز» و «کیو-اس»، بسیار قوی ($r=0.818$ و $r=0.882$) و با نمره کل در «شانگهای»، در سطح متوسط ($r=0.557$) ارزیابی می‌شود (جدول ۲).

پاسخ به سؤال دوم پژوهش. آیا میان نمره دانشگاه‌ها در نظام «رف» و نمره آنها در هر یک از ابعاد نظام‌های رتبه‌بندی معتبر جهانی همبستگی معناداری وجود دارد؟

بر اساس آزمون کولموگروف-اسمیرنوف، توزیع نمرات ابعاد «استناد»، «درآمد صنعتی»، «چشم‌انداز بین‌المللی»، «اعضای هیئت علمی پراستناد در ۲۱ گروه موضوعی»، «مقالات در نمایه‌نامه علوم و علوم اجتماعی»، «انتشارات پراستناد دانشگاه در ۱۰ درصد انتشارات پراستناد در همان سال و در همان زمینه»، «میانگین نمره استناد»، «انتشارات مشترک بین‌سازمانی»، «انتشارات مشترک بین‌المللی»، «انتشارات مشترک دانشگاه با بخش صنعت»، «انتشارات مشترک دانشگاه با فاصله جغرافیایی کمتر از ۱۰۰ کیلومتر»، «انتشارات (همکاری با فاصله جغرافیایی طولانی) بیشتر از ۱۰۰۰ کیلومتر» نرمال بود ($P\text{-value} > 0.05$)، اما برای دیگر ابعاد نرمال نبود ($P\text{-value} < 0.05$). از این رو، در صورت نرمال بودن هر دو متغیر از همبستگی پیرسون و در صورت نرمال نبودن یکی از متغیرها از همبستگی اسپیرمن استفاده شد.

جدول ۳. همبستگی میان نمره دانشگاه‌ها در «رف» و نمرات آنها در ابعاد نظام‌های رتبه‌بندی

ابعاد	نظام	ابعاد مشابه در نظام‌ها	ضریب همبستگی	سطح معناداری
آموزش	تایمز آموزش		۰.۷۸۳**	۰.۰۰۰
	شانگهای دانش‌آموختگان برنده جایزه نوبل یا مدال فیلدز		۰.۴۵۸**	۰.۰۰۴
	شانگهای سرانه عملکرد دانشگاه		۰.۶۰۴**	۰.۰۰۰
	کیواس نسبت اعضای هیئت علمی به دانشجو		۰.۸۳۸**	۰.۰۰۰
پژوهش	تایمز پژوهش در نظام تایمز		۰.۸۱۰**	۰.۰۰۰
	شانگهای مقالات در نمایه‌نامه علوم و علوم اجتماعی		۰.۵۸۶**	۰.۰۰۰
	شانگهای مقالات منتشر شده در مجلات نیچر و ساینس		۰.۵۳۷**	۰.۰۰۱
	لایدن انتشارات (تعداد مقالات)		۰.۵۸۲**	۰.۰۰۰
	لایدن انتشارات پراستناد دانشگاه در ۱۰ درصد انتشارات پراستناد		۰.۵۰۸**	۰.۰۰۱
استناد	تایمز استناد		۰.۴۳۷**	۰.۰۰۲
	کیو-اس سرانه استناد (نسبت استناد به ازای هر عضو هیئت علمی)		۰.۸۰۳**	۰.۰۰۰
	لایدن میانگین نمره استناد		۰.۵۰۲**	۰.۰۰۱
چشم‌انداز بین‌المللی	لایدن میانگین نمره استناد نرمال شده		۰.۵۴۲**	۰.۰۰۱
	تایمز چشم‌انداز بین‌المللی		۰.۴۰۰**	۰.۰۰۴
	کیو-اس شهرت علمی دانشگاه		۰.۸۶۷**	۰.۰۰۰
	کیو-اس شهرت از دید کارفرمایان		۰.۷۸۳**	۰.۰۰۰
	کیو-اس سهم اعضای هیئت علمی بین‌المللی		۰.۵۸۷**	۰.۰۰۰

ادامه جدول ۳. همبستگی میان نمره دانشگاه‌ها در «رف» و نمرات آنها در ابعاد نظام‌های رتبه‌بندی

ابعاد	نظام	ابعاد مشابه در نظام‌ها	ضریب همبستگی	سطح معناداری
چشم‌انداز بین‌المللی	کیو-اس	سهام دانشجویان بین‌المللی	۰.۲۷۸*	۰.۰۲۰
	اعضای هیئت علمی برنده جایزه نوبل یا مدال فیلدز		۰.۴۹۳**	۰.۰۰۰۲
	اعضای هیئت علمی پرستاد در ۲۱ گروه موضوعی		۰.۵۶۶**	۰.۰۰۰
	لایدن	انتشارات مشترک بین‌سازمانی	۰.۶۰۵**	۰.۰۰۱
	لایدن	نسبت انتشارات مشترک بین‌سازمانی	۰.۰۱۳	۰.۹۳۵
	لایدن	انتشارات مشترک بین‌المللی	۰.۲۹۳	۰.۰۵۴
	لایدن	انتشارات مشترک یک دانشگاه با بخش صنعت	۰.۰۰۹	۰.۹۵۶
	لایدن	انتشارات مشترک یک دانشگاه با فاصله جغرافیایی کمتر از ۱۰۰ کیلومتر	۰.۰۱۷	۰.۹۱۴
	لایدن	انتشارات (همکاری با فاصله جغرافیایی طولانی) بیشتر از ۱۰۰۰ کیلومتر	۰.۲۲۸	۰.۱۳۶
	درآمد صنعتی	تایمز	درآمد صنعتی	۰.۵۴۳**

** همبستگی در سطح ۰.۰۱ معنادار است (دوسویه).

* همبستگی در سطح ۰.۰۵ معنادار است (دوسویه).

همان‌گونه که در جدول ۳ مشاهده می‌شود، برای دستیابی به درکی بهتر از روابط، ابعاد مختلف نظام‌های بین‌المللی بر اساس مشابهت آنها گروه‌بندی شدند. ابعاد «تایمز» به‌عنوان مبنای دسته‌بندی به کار رفته است. گرچه برخی ابعاد را می‌توان در بیش از یک گروه دسته‌بندی کرد، در اینجا گروه‌بندی تنها به منظور نظم‌بخشی به مطالب و درک بهتر نتایج انجام شده است. نمره دانشگاه‌ها در «رف» با نمره آنها در اغلب ابعاد نظام‌های رتبه‌بندی همبستگی معناداری نشان می‌دهد ($P\text{-value} < 0.01$). اما، میان این نمره با ابعاد «نسبت انتشارات مشترک بین‌سازمانی»، «انتشارات مشترک بین‌المللی»، «انتشارات مشترک یک دانشگاه با بخش صنعت»، «انتشارات مشترک یک دانشگاه با فاصله جغرافیایی کمتر از ۱۰۰ کیلومتر»، «انتشارات مشترک یک دانشگاه با فاصله جغرافیایی بیشتر از ۱۰۰۰ کیلومتر» در «لایدن» همبستگی معناداری دیده نمی‌شود ($P\text{-value} > 0.05$).

نمره در «رف» با بعد «آموزش» در «تایمز»، همبستگی قوی ($r=0.783$) نشان می‌دهد؛ درحالی‌که با «سرانه عملکرد دانشگاه» ($r=0.604$) و «دانش‌آموختگان برنده جایزه نوبل یا مدال فیلدز» ($r=0.458$) همبستگی متوسط نشان می‌دهد. اما، با بعد «نسبت استاد به دانشجو» ($r=0.838$) در «کیو-اس» و بعد «پژوهش» در «تایمز» ($r=0.810$) همبستگی بسیار قوی دارد. در «شانگهای» «مقالات در نمایه‌نامه علوم و علوم اجتماعی» ($r=0.586$)، و «مقالات در مجلات نیچر و ساینس» ($P\text{-value} < 0.01$ و $r=0.537$) همبستگی متوسط؛ و در «لایدن» ابعاد «تعداد مقالات» ($r=0.582$)، و «انتشارات پرستاد دانشگاه در ۱۰ درصد انتشارات پرستاد در همان سال و در همان زمینه» ($r=0.508$) همبستگی متوسط با نمره در «رف» نشان می‌دهند.

بعد «استناد» نیز در نظام «تایمز» ($r=0.437$)، همبستگی متوسط؛ و در نظام «کیواس» بعد «سرانه استناد» ($r=0.803$)، همبستگی بسیار قوی با نمره در «رف» نشان می‌دهد. در لایدن نیز بعد «میانگین نمره استناد» ($r=0.502$) و «میانگین نمره استناد نرمال‌شده» ($r=0.542$)، همبستگی متوسطی با نمره در «رف» نشان می‌دهد.

همچنین، میان نمره در «رف» و بعد «چشم‌انداز» در نظام «تایمز»، همبستگی متوسط ($r=0.400$) دیده می‌شود. در نظام «کیو-اس» بعد شهرت علمی دانشگاه ($r=0.867$) و بعد «شهرت از دید کارفرمایان» ($r=0.783$) همبستگی قوی و بعد «سهم اعضای هیئت علمی بین‌المللی» همبستگی متوسط ($r=0.587$) و بعد «نسبت دانشجویان بین‌المللی»، همبستگی ضعیفی ($r=0.278$) نشان می‌دهند. در «شانگهای» نیز در ابعاد «اعضای هیئت علمی برنده جایزه نوبل یا مدال فیلدز» ($r=0.493$) و «اعضای هیئت علمی پراستناد در ۲۱ گروه موضوعی» ($r=0.566$) همبستگی متوسط مشاهده می‌شود؛ اما در «لایدن» تنها در بعد «انتشارات مشترک بین‌سازمانی» ($r=0.605$) همبستگی متوسط مشاهده می‌شود و در بقیه ابعاد همبستگی وجود ندارد. بعد «درآمد صنعتی» در «تایمز» ($r=0.543$)، همبستگی متوسطی با نمره در «رف» نشان می‌دهد.

پاسخ به سؤال سوم پژوهش. آیا گرایش موضوعی تأثیری معنادار بر نمره دانشگاه‌ها در نظام «رف» و نظام‌های رتبه‌بندی معتبر جهانی دارد؟

برای بررسی تأثیر گرایش موضوعی بر نمره دانشگاه‌ها در نظام‌های رتبه‌بندی، از تحلیل رگرسیون خطی چندگانه استفاده شد. پیش از انجام این تحلیل نیز، ابتدا به کمک آزمون کولموگروف-اسمیرنوف از نرمال بودن توزیع داده‌های متغیر وابسته اطمینان حاصل شد.

جدول ۴. تأثیر گرایش موضوعی بر نمره دانشگاه‌ها در «رف» و نظام‌های رتبه‌بندی جهانی

متغیر پیش‌بین	متغیر وابسته	ضریب تعیین	تی	بتا	سطح معناداری
علوم پایه	نمره در «رف»	۰.۰۴۷	هنر، علوم انسانی و علوم اجتماعی	-۰.۱۲۵	۰.۳۹۶
			پزشکی	۰.۰۹۲	۰.۹۲۷
			فنی و مهندسی	۰.۷۹۲	۰.۴۳۱
علوم پایه	نمره در تایمز	۰.۰۵۴	هنر، علوم انسانی و علوم اجتماعی	-۰.۱۵۴	۰.۸۷۵
			پزشکی	۰.۱۵۲	۰.۸۸۰
			فنی و مهندسی	۰.۶۹۷	۰.۴۸۹
علوم پایه	نمره در کیو-اس	۰.۱۴۷	هنر، علوم انسانی و علوم اجتماعی	-۱.۳۸۹	۰.۱۷۰
			پزشکی	۰.۸۷۹	۰.۳۸۳
			فنی و مهندسی	۲.۴۶۸	۰.۰۱۶
علوم پایه	نمره در شانگهای	۰.۰۶۳	هنر، علوم انسانی و علوم اجتماعی	-۰.۷۲۸	۰.۴۷۲
			پزشکی	۰.۰۹۱	۰.۹۲۸
			فنی و مهندسی	۰.۲۱۶	۰.۸۳۱
					۰.۴۰۲

با توجه به معنادار نبودن مدل برای «رف»، «تایمز» و «شانگهای» ($P\text{-value} > 0.05$)، گرایش‌های موضوعی در دانشگاه‌های انگلیس پیش‌بین نمره در این سامانه‌ها نیستند. اما در «کیو-اس» گرایش پزشکی تأثیر معناداری را نشان می‌دهد ($P\text{-value} < 0.05$). به این ترتیب، نمره دانشگاه در «کیو-اس»، به‌طور معنادار به میزان گرایش آن به رشته پزشکی بستگی دارد، ولی دیگر گرایش‌ها تعیین‌کننده نیستند ($P\text{-value} > 0.05$). گرایش موضوعی ۱۴.۷ درصد از واریانس نمره در «کیو-اس» را پیش‌بینی می‌کند. پزشکی با ضریب رگرسیون ۰.۳۲۸ بالاترین قابلیت پیش‌بینی مثبت نمره در «کیو-اس» را دارد (جدول ۴).

پاسخ به سؤال چهارم پژوهش. گرایش‌های موضوعی دانشگاه‌ها بر نمره آنها در کدام‌یک از ابعاد نظام‌های رتبه‌بندی معتبر جهانی تأثیر گذار است؟

تحلیل رگرسیون خطی چندگانه نشان داد که گرایش موضوعی هیچ‌یک از ابعاد نظام «شانگهای» را به‌طور معناداری پیش‌بینی نمی‌کند ($P\text{-value} > 0.05$).

جدول ۵. تأثیر گرایش موضوعی بر نمره ابعاد در «تایمز»

متغیر وابسته	متغیر پیش‌بین	ضریب تعیین	تی	بتا	سطح معناداری
آموزش	هنر، علوم انسانی و اجتماعی	۰.۰۵۷	-۰.۶۲۸	-۰.۱۱۵	۰.۵۳۳
	علوم پایه		-۰.۹۱۵	-۰.۱۴۶	۰.۳۶۵
	پزشکی		-۰.۲۷۹	-۰.۰۴۹	۰.۷۸۲
پژوهش	فنی و مهندسی	۰.۰۵۷	۰.۵۵۸	۰.۰۸۹	۰.۵۸۰
	هنر، علوم انسانی و اجتماعی		-۰.۶۵۰	-۰.۱۱۹	۰.۵۱۹
	علوم پایه		-۱.۰۲۰	-۰.۱۶۳	۰.۳۱۳
استاد	پزشکی	۰.۲۰۵	-۰.۱۹۹	-۰.۰۳۵	۰.۸۴۳
	فنی و مهندسی		۰.۳۲۵	۰.۰۵۲	۰.۷۴۷
	هنر، علوم انسانی و اجتماعی		۱.۱۱۷	۰.۱۸۸	۰.۲۷۰
درآمد صنعتی	علوم پایه	۰.۲۴۸	۰.۴۰۰	۰.۰۵۹	۰.۶۹۱
	پزشکی		۳.۲۴۴	۰.۵۲۶	۰.۰۰۲
	فنی و مهندسی		-۰.۰۸۸	-۰.۰۱۳	۰.۹۳۰
چشم‌انداز بین‌المللی	هنر، علوم انسانی و اجتماعی	۰.۱۴۳	-۲.۱۱۸	-۰.۳۴۸	۰.۰۴۰
	علوم پایه		-۰.۰۶۱	-۰.۰۰۹	۰.۹۵۲
	پزشکی		-۰.۷۳۷	-۰.۱۱۶	۰.۴۶۵
	فنی و مهندسی		۲.۰۳۰	۰.۲۹۰	۰.۰۴۸
	هنر، علوم انسانی و اجتماعی		۰.۳۶۳	۰.۰۶۴	۰.۷۱۸
	علوم پایه		-۱.۷۸۱	-۰.۲۷۲	۰.۰۸۲
	پزشکی		-۱.۳۷۷	-۰.۲۳۲	۰.۱۷۵
	فنی و مهندسی		۱.۳۳۹	۰.۲۰۴	۰.۱۸۷

در «تایمز»، گرایش موضوعی پیش‌بین معنادار نمره در ابعاد «آموزش»، «پژوهش» و «چشم‌انداز بین‌المللی» نیست ($P\text{-value} > 0.05$)؛ اما، ۲۰.۵ درصد از واریانس نمره بعد «استناد» را پیش‌بینی می‌کند. گرایش پزشکی با ضریب رگرسیون ۰.۵۲۶ دارای بالاترین قدرت پیش‌بینی مثبت نمره در این بعد است. همچنین، گرایش موضوعی ۲۴.۸ درصد از واریانس نمره بعد «درآمد صنعتی» را پیش‌بینی می‌کند. گرایش فنی و مهندسی با ضریب ۰.۲۹۰ بیشترین تأثیر مثبت را بر این متغیر دارد. گرایش هنر، علوم انسانی و اجتماعی با ضریب -0.348 دارای بالاترین قابلیت پیش‌بینی منفی در این بعد می‌باشد (جدول ۵).

جدول ۶. تأثیر گرایش موضوعی بر نمره ابعاد در «کیو-اس»

متغیر وابسته	متغیر پیش‌بین	ضریب تعیین	تی	بتا	سطح معناداری
شهرت علمی دانشگاه	هنر، علوم انسانی و اجتماعی	۰.۱۲۹	-۱.۱۹۷	-۰.۲۱۷	۰.۲۳۶
	علوم پایه		-۰.۷۱۳	-۰.۰۹۶	۰.۴۷۹
	پزشکی		۲.۴۹۳	۰.۳۳۵	۰.۰۱۵
شهرت از دید کارفرما	فنی و مهندسی	۰.۰۸۳	-۰.۶۲۸	-۰.۱۱۶	۰.۵۳۲
	هنر، علوم انسانی و اجتماعی		-۰.۹۲۰	-۰.۱۷۱	۰.۳۶۱
	علوم پایه		-۰.۵۴۵	-۰.۰۷۶	۰.۵۸۸
نسبت استاد به دانشجو	پزشکی	۰.۱۷۹	۰.۵۹۱	۰.۰۸۱	۰.۵۵۶
	فنی و مهندسی		۰.۵۷۱	۰.۱۰۸	۰.۵۷۰
	هنر، علوم انسانی و اجتماعی		-۱.۷۳۳	-۰.۳۰۵	۰.۰۸۸
سهم اعضای هیئت علمی بین‌المللی	علوم پایه	۰.۰۸۵	-۱.۲۳۱	-۰.۱۶۲	۰.۲۲۳
	پزشکی		۲.۷۵۶	۰.۳۵۹	۰.۰۰۸
	فنی و مهندسی		-۰.۷۰۸	-۰.۱۲۷	۰.۴۸۱
سهم دانشجویان بین‌المللی	هنر، علوم انسانی و اجتماعی	۰.۱۳۳	-۱.۷۰۸	-۰.۳۱۷	۰.۰۹۲
	علوم پایه		-۲.۲۰۲	-۰.۳۰۶	۰.۰۳۱
	پزشکی		۱.۱۹۶	۰.۱۶۵	۰.۲۳۹
سرانه استناد	فنی و مهندسی	۰.۲۱۹	-۱.۸۳۲	-۰.۳۴۶	۰.۰۷۲
	هنر، علوم انسانی و اجتماعی		-۰.۸۲۷	-۰.۱۴۹	۰.۴۱۱
	علوم پایه		-۳.۰۰۶	-۰.۴۰۶	۰.۰۰۴
	پزشکی		۰.۱۹۱	۰.۰۲۶	۰.۸۴۹
	فنی و مهندسی		-۰.۴۷۲	-۰.۰۸۷	۰.۶۳۸
	هنر، علوم انسانی و اجتماعی		-۱.۴۱۸	-۰.۲۴۳	۰.۱۶۱
	علوم پایه		-۰.۲۷۴	-۰.۰۳۵	۰.۷۸۵
	پزشکی		۳.۱۴۵	۰.۴۰۰	۰.۰۰۳
	فنی و مهندسی		-۰.۴۸۰	-۰.۴۸۰	۰.۶۳۳

در نظام «کیو-اس» گرایش موضوعی پیش‌بینی‌کننده معنادار نمره در بعد «شهرت از دید کارفرما» نیست ($P > 0.05$)، اما ۱۲.۹ درصد از واریانس نمره «شهرت علمی دانشگاه» را پیش‌بینی می‌کند؛ گرایش پزشکی با ضریب رگرسیون ۰.۳۳۵ دارای بالاترین قابلیت پیش‌بینی مثبت و معنادار نمره در این بعد می‌باشد. همچنین، گرایش موضوعی ۱۷.۹ درصد از واریانس نمره بعد «نسبت استاد به دانشجو» را پیش‌بینی می‌کند. گرایش پزشکی با ضریب رگرسیون ۰.۳۵۹ دارای بالاترین قابلیت پیش‌بینی مثبت و معنادار نمره در این بعد می‌باشد. علاوه بر این، گرایش موضوعی ۸.۵ درصد از واریانس نمره در بعد «سهم اعضای هیئت علمی بین‌المللی» را پیش‌بینی می‌کند. گرایش علوم پایه با ضریب رگرسیون ۰.۳۰۶- دارای بالاترین قابلیت پیش‌بینی منفی و معنادار نمره بعد «سهم اعضای هیئت علمی بین‌المللی» می‌باشد. همچنین، گرایش موضوعی ۱۳.۳ درصد از واریانس نمره بعد «سهم دانشجویان بین‌المللی» را پیش‌بینی می‌کند؛ علوم پایه با ضریب رگرسیون ۰.۴۰۶- دارای بالاترین قابلیت پیش‌بینی منفی و معنادار نمره بعد «سهم دانشجویان بین‌المللی» می‌باشد. علاوه بر این، گرایش موضوعی ۲۱.۹ درصد از واریانس نمره بعد «سرانه استناد» را پیش‌بینی می‌کند. گرایش پزشکی با ضریب ۰.۴۰۰ دارای بالاترین قابلیت پیش‌بینی مثبت و معنادار نمره بعد «سرانه استناد» است (جدول ۶).

جدول ۷. تأثیر گرایش موضوعی بر نمره بعد «تأثیر» در لایدن

متغیر وابسته	متغیر پیش‌بین	ضریب تعیین	تی	بتا	سطح معناداری
	هنر، علوم انسانی و اجتماعی		-۰.۱۷۵	-۰.۰۳۴	۰.۸۶۲
انتشارات (تعداد مقالات)	علوم پایه	۰.۱۱۲	۰.۳۸۱	۰.۰۸۱	۰.۷۰۵
	پزشکی		۱.۱۶۶	۰.۲۰۶	۰.۲۵۱
	فنی و مهندسی		۱.۸۰۱	۰.۲۹۶	۰.۰۷۹
انتشارات پراستناد دانشگاه در ۱۰ درصد انتشارات پراستناد	هنر، علوم انسانی و اجتماعی		-۰.۶۸۹	-۰.۱۱۹	۰.۴۹۵
	علوم پایه	۰.۳۰۱	۱.۴۶۰	۰.۲۷۵	۰.۱۵۲
پزشکی	۳.۶۷۰		۰.۵۷۶	۰.۰۰۱	
	فنی و مهندسی		-۰.۸۷۹	-۰.۱۲۸	۰.۳۸۵
	هنر، علوم انسانی و اجتماعی		-۰.۵۹۹	-۰.۱۰۳	۰.۵۵۳
میانگین نمره استناد	علوم پایه	۰.۳۰۷	۱.۶۳۲	۰.۳۰۶	۰.۱۱۱
	پزشکی		۳.۹۷۱	۰.۶۲۰	۰.۰۰۰
	فنی و مهندسی		۱.۸۰۱	۰.۲۹۶	۰.۰۷۹
میانگین نمره استناد نرمال‌شده	هنر، علوم انسانی و اجتماعی		-۰.۶۸۳	-۰.۱۱۸	۰.۴۹۹
	علوم پایه	۰.۳۰۱	۱.۰۷۱	۰.۲۰۲	۰.۲۹۱
پزشکی	۳.۴۹۴		۰.۵۴۸	۰.۰۰۱	
	فنی و مهندسی		-۱.۰۱۰	-۰.۱۴۷	۰.۳۱۹

در «لایدن»، گرایش موضوعی بر نمره در بعد «انتشارات» تأثیری ندارد ($P\text{-value} > 0.05$)، اما می‌تواند ۳۰.۱ درصد از واریانس نمره در بعد «نسبت انتشارات پراستناد دانشگاه در ۱۰ درصد انتشارات پراستناد در همان سال و در همان زمینه» را پیش‌بینی کند. گرایش پزشکی با ضریب ۰.۵۷۶ بالاترین قدرت پیش‌بینی مثبت برای نمره در این بعد را دارد.

همچنین، گرایش موضوعی ۳۰.۷ درصد از واریانس نمره در بعد «میانگین نمره استاد» را پیش‌بینی می‌کند، گرایش پزشکی با ضریب ۰.۶۲۰ بیشترین قابلیت پیش‌بینی نمره این بعد را داراست. علاوه بر این، گرایش موضوعی ۳۰.۱ درصد از واریانس نمره در بعد «میانگین نمره استاد نرمال‌شده» را پیش‌بینی می‌کند؛ در این بعد نیز، گرایش پزشکی با ضریب ۰.۵۴۸ قوی‌ترین پیش‌بین است (جدول ۷).

جدول ۸. تأثیر گرایش موضوعی بر نمره بعد «همکاری» در «لایدن»

متغیر وابسته	متغیر پیش‌بین	ضریب تعیین	تی	بتا	سطح معناداری
	هنر، علوم انسانی و اجتماعی		-۰.۲۶۱	-۰.۰۵۱	۰.۷۹۵
انتشارات مشترک	علوم پایه	۰.۱۲۰	۰.۳۷۷	۰.۰۸۰	۰.۷۰۸
بین‌سازمانی	پزشکی		۱.۳۶۶	۰.۲۴۰	۰.۱۸۰
	فنی و مهندسی		۱.۷۶۷	۰.۲۸۹	۰.۰۸۵
	هنر، علوم انسانی و اجتماعی		۰.۶۷۸	۰.۰۹۰	۰.۵۰۲
نسبت انتشارات	علوم پایه	۰.۵۸۴	۰.۰۹۷	۰.۰۱۴	۰.۹۲۳
مشترک بین‌سازمانی	پزشکی		۶.۷۶۹	۰.۸۱۹	۰.۰۰۰
	فنی و مهندسی		۱.۶۲۱	۰.۱۸۳	۰.۱۱۳
	هنر، علوم انسانی و اجتماعی		-۰.۱۶۳	-۰.۰۲۶	۰.۸۷۱
انتشارات مشترک	علوم پایه	۰.۳۹۰	-۰.۰۴۸	-۰.۰۰۸	۰.۹۶۲
بین‌المللی	پزشکی		۴.۰۹۱	۰.۵۹۹	۰.۰۰۰
	فنی و مهندسی		-۰.۳۳۶	-۰.۰۴۶	۰.۷۳۸
نسبت انتشارات	هنر، علوم انسانی و اجتماعی		-۲.۱۳۸	-۰.۳۶۴	۰.۰۳۹
مشترک یک دانشگاه	علوم پایه	۰.۳۲۳	-۱.۲۸۵	-۰.۲۳۸	۰.۲۰۷
با بخش صنعت	پزشکی		-۱.۷۸۳	-۰.۲۷۵	۰.۰۸۲
	فنی و مهندسی		۲.۱۰۸	۰.۳۰۳	۰.۰۴۲
نسبت انتشارات	هنر، علوم انسانی و اجتماعی		-۰.۳۹۰	-۰.۰۷۶	۰.۶۹۹
مشترک با فاصله	علوم پایه	۰.۱۰۵	-۱.۴۹۶	-۰.۳۱۹	۰.۱۴۳
جغرافیایی کمتر از	پزشکی		-۰.۶۸۹	-۰.۱۲۲	۰.۴۹۵
۱۰۰ کیلومتر	فنی و مهندسی		۰.۷۴۷	۰.۱۲۳	۰.۴۶۰
نسبت انتشارات	هنر، علوم انسانی و اجتماعی		-۰.۲۰۶	-۰.۰۳۳	۰.۸۳۸
مشترک با فاصله	علوم پایه	۰.۴۰۵	۰.۰۴۲	۰.۰۰۷	۰.۹۶۶
جغرافیایی بیشتر از	پزشکی		۴.۳۰۳	۰.۶۲۲	۰.۰۰۰
۱۰۰۰ کیلومتر	فنی و مهندسی		-۰.۱۸۱	-۰.۰۲۴	۰.۸۵۷

در «لایدن»، گرایش موضوعی بر نمره در ابعاد «انتشارات مشترک بین‌سازمانی» و «انتشارات مشترک یک دانشگاه با فاصله جغرافیایی کمتر از ۱۰۰ کیلومتر» تأثیری ندارد ($P\text{-value} > 0.05$). با این حال، ۵۸.۴ درصد از واریانس نمره در بعد «نسبت انتشارات مشترک بین‌سازمانی» توسط این متغیر قابل پیش‌بینی است. گرایش پزشکی با ضریب ۰.۸۱۹

همبستگی بین رتبه‌بندی دانشگاه‌ها به دو روش کتاب‌سنجی محض و داوری غنی‌شده با کتاب‌سنجی...

دارای بالاترین قابلیت پیش‌بینی مثبت نمره در این بعد است. همچنین، گرایش موضوعی، ۳۹ درصد از واریانس نمره در بعد «نسبت انتشارات مشترک بین‌المللی» را پیش‌بینی می‌کند. در این بعد، نیز گرایش پزشکی با ضریب ۰.۵۹۹ بیشترین تأثیر را دارد. همچنین، این متغیر، ۳۲.۲ درصد از واریانس نمره در بعد «نسبت انتشارات مشترک یک دانشگاه با بخش صنعت» را پیش‌بینی می‌کند. گرایش علوم انسانی با ضریب ۰.۳۶۴- تأثیر منفی بر این بعد دارد، درحالی‌که گرایش فنی و مهندسی با ضریب ۰.۳۰۳ این بعد را به‌طور مثبت پیش‌بینی می‌کند. علاوه‌براین، گرایش موضوعی، ۴۰.۵ درصد از نمره بعد «نسبت انتشارات (همکاری با فاصله جغرافیایی طولانی) بیشتر از ۱۰۰۰ کیلومتر» را پیش‌بینی می‌کند؛ باز هم این گرایش پزشکی (با ضریب ۰.۶۲۲) است که بالاترین قابلیت پیش‌بینی نمره این بعد را دارد (جدول ۸).

بحث و نتیجه‌گیری

پژوهش حاضر به‌منظور آگاهی از تشابه نتایج حاصل از به‌کارگیری روش داوری غنی‌شده با کتاب‌سنجی با روش‌های مبتنی بر کتاب‌سنجی، تلفیق کتاب‌سنجی و نظرسنجی و دیگر داده‌های عملکردی دانشگاه‌ها انجام گرفت. بدین منظور، نظام ملی «رف» به‌عنوان نظام مبتنی بر داوری غنی‌شده با کتاب‌سنجی و «لایدن» به‌عنوان نظام مبتنی بر کتاب‌سنجی صرف و «تایمز» و «کیو-اس» به‌عنوان نظام‌های مبتنی بر کتاب‌سنجی، نظرسنجی و آمار عملکردی و نظام «شانگهای» به‌عنوان نظام مبتنی بر داده‌های کتاب‌سنجی و آمار عملکردی انتخاب شدند.

در پاسخ به سؤال اول پژوهش، بر اساس نتایج حاصل از تحلیل همبستگی مشخص شد که بین نمره کیفیت پژوهش دانشگاه‌ها در سیستم «رف» و نمره آنها در سیستم‌های مبتنی بر کتاب‌سنجی غنی‌شده با نظرسنجی («تایمز» و «کیو-اس») و سیستم ارزیابی مبتنی بر کتاب‌سنجی («شانگهای») همبستگی معناداری وجود دارد. این یافته نشان از هم‌سویی نتایج این نظام‌های رتبه‌بندی با یکدیگر دارد. مهم‌ترین وجه تمایز نظام «رف» با دو نظام «تایمز» و «کیو-اس»، استفاده از داوری در برابر نظرسنجی است. از آنجایی‌که این نظام‌ها در ابعاد مرتبط با کتاب‌سنجی و عملکرد دانشگاه شباهت‌های بسیاری با هم دارند، ضریب همبستگی بسیار قوی را می‌توان دال بر نتایج نزدیک به هم در دو روش مبتنی بر داوری و نظرسنجی دانست. به‌عبارت‌دیگر، احتمال می‌رود غنی‌سازی داده‌های کتاب‌سنجی با نظرسنجی نتایجی بسیار مشابه با روش داوری غنی‌شده با کتاب‌سنجی به بار آورد. اهمیت این بخش از یافته‌ها را می‌توان در آن دانست که نظام ملی ارزیابی بریتانیا به‌ازای صرف بودجه‌های هنگفت، به نتایج دقیق‌تری دست می‌یابد (Stern, 2016). از طرف دیگر، بین سامانه «رف» که نظام مبتنی بر داوری غنی‌شده با کتاب‌سنجی است با نظام «شانگهای» که تنها مبتنی بر کتاب‌سنجی است همبستگی ضعیف‌تر و در سطح متوسط ($r=0.557$) مشاهده می‌شود. این نظام نه‌تنها در ابعاد عملکرد دانشگاه تفاوت‌هایی با نظام «رف» دارد، بلکه از هیچ یک از دو روش داوری و نظرسنجی استفاده نمی‌کند. این یافته را می‌توان هم‌راستا با نتایج پژوهش (Aksnes & Taxt, 2004) در دانشگاه برگن نروژ دانست که همبستگی مثبت، اما نسبتاً ضعیفی را بین نتایج ارزیابی داوران و شاخص‌های کتاب‌سنجی مشاهده کردند. بنابراین، اگر زمانی دانشگاه‌هایی با عملکرد مشابه با دانشگاه‌های انگلیس بخواهند بین این دو نوع نظام انتخاب کنند، نظام «تایمز» و «کیو-اس» نتایج مشابهی را با نتایج نظام «رف» نشان می‌دهند. دلیل ارجحیت این دو نظام بر نظام «رف» را می‌توان لزوم صرف هزینه قابل توجه برای سیستم‌های مبتنی بر داوری دانست (Geuna & Piolatto, 2016; Bertocchi et al, 2015; Aksnes & Taxt, 2004).

پرسش دوم در پژوهش حاضر در امتداد پرسش اول و در ابعاد جزئی‌تر طرح شده است. یافته‌ها نشان می‌دهد که

بین نمره دانشگاه‌ها در «رف» و نمرات آنها در تمام ابعاد نظام‌های رتبه‌بندی -منهای ابعادی از نظام لایدن- همبستگی مثبت و معناداری وجود دارد. یکی از دلایل محتمل برای این یافته، همبستگی قوی این نظام‌ها با یکدیگر در بخش‌هایی از روش‌شناسی -به جز بخش داوری- است. از طرفی، احتمال می‌رود بخش مربوط به نتایج داوری در «رف»، ناهمسو با شاخص‌های کتاب‌سنجی نباشد تا تصویری متفاوت ایجاد کند. این یافته با پژوهش (2004) Aksnes & Taxt مبنی بر همبستگی مثبت ارزیابی داوران با شاخص‌های کتاب‌سنجی هم‌راستا است. علاوه بر این، Rinia et al. (1998) نیز همبستگی‌هایی را با درجات متفاوت بین شاخص‌های مختلف کتاب‌سنجی و نتایج داوری مشاهده کرده‌اند.

در پاسخ به پرسش سوم پژوهش نیز، یافته‌ها نشان داد نمره کل دانشگاه‌ها در نظام‌های مورد بررسی متأثر از گرایش موضوعی آنها نیست، تنها استثناء نظام «کیو-اس» بود که در آن، گرایش پزشکی تأثیری معنادار را نشان می‌داد. اگرچه، با توجه به پایین بودن ضریب به‌دست‌آمده می‌توان چنین استنباط کرد که این متغیر تأثیر بسیار ناچیزی بر نمره کل دانشگاه در نظام «کیو-اس» دارد. این یافته در حالی به دست آمد که پیش از این، نتایج پژوهش (2011) Rauhvargers حاکی از گرایش شاخص ضریب تأثیر به سمت علوم طبیعی و پزشکی بود. (۲۰۱۲) Radicchi & Castellano نیز نشان دادند که مقادیر ضریب تأثیر در رشته زیست‌شناسی سلولی مولکولی، بیش از رشته ریاضی بوده و یا مقدار شاخص اچ در شیمی بیش از کامپیوتر است. علاوه بر این، تفاوت استنادی در گرایش‌های یک رشته نیز در پژوهش (2012) Smolinsky & Lercher تأیید شده است. همچنین، Mryglod et al (2013) با مقایسه شاخص‌های استنادی و شاخص‌های مبتنی بر داوری در چندین رشته دریافتند که به جهت ماهیت علوم سخت (علوم طبیعی، مهندسی و علوم پزشکی)، نمره استناد در این علوم، بیش از علوم نرم (مانند علوم اجتماعی و انسانی) است. (2011) Abramo & D'Angelo نیز با مقایسه دو رویکرد داوری و کتاب‌سنجی در ارزیابی پژوهش دریافتند که کارکرد روش کتاب‌سنجی در علوم طبیعی و صوری، به مراتب بهتر از داوری است. با این حال، نتایج پژوهش حاضر، تأثیر گرایش‌های موضوعی دانشگاه‌های انگلیس را بر نمره آنها در نظام‌های رتبه‌بندی چندان مؤثر نیافته است. این امر می‌تواند به این دلیل باشد که احتمالاً این دانشگاه‌ها در گرایش‌های موضوعی خود مشابهت داشته‌اند یا اینکه روش‌شناسی‌های به کار گرفته‌شده در این نظام‌ها موفق به خنثی کردن تأثیر تفاوت‌های موضوعی شده‌اند. تحقیقات بیشتر برای روشن شدن دلایل این ناهمسویی با تحقیقات پیشین ضروری است.

پرسش چهارم هم در امتداد پرسش سوم و به‌منظور بررسی تأثیر گرایش‌های موضوعی بر نمره ابعاد مختلف در نظام‌های رتبه‌بندی طرح شد. در این بخش، برای هر یک از نظام‌های مورد بررسی یافته‌های متفاوتی به دست آمد. برای مثال، در نظام «شانگهای» گرایش به موضوعی خاص نمی‌تواند بر نمره در ابعاد این نظام مؤثر باشد؛ اما در «تایمز»، گرایش‌های موضوعی دانشگاه‌ها بر نمره در ابعاد استناد و درآمد صنعتی دانشگاه‌ها تأثیر دارد؛ به‌گونه‌ای که گرایش پزشکی تأثیری مثبت بر نمره بعد استناد دارد. در تبیین چرایی این یافته می‌توان به نابرابری میزان انتشارات علمی در رشته‌های مختلف اشاره داشت (Mryglod et al, 2013). در بعد درآمد صنعتی، گرایش مهندسی به صورت مثبت و گرایش علوم انسانی به صورت منفی و قوی‌تر از گرایش مهندسی، نمره را پیش‌بینی می‌کند. به این ترتیب، اگر دانشگاهی در گرایش مهندسی قوی‌تر و در گرایش علوم انسانی ضعیف‌تر باشد نمره آن در بعد درآمد صنعتی بیشتر خواهد بود. با توجه به سرشت رشته‌های فنی و مهندسی، احتمال بیشتری برای همکاری با بخش صنعت و در نتیجه، درآمدهای بیشتر می‌رود. به گواه تحقیقات پیشین، کیفیت برون‌دادهای پژوهشی حاصل از همکاری دانشگاه با شرکای

همبستگی بین رتبه‌بندی دانشگاه‌ها به دو روش کتاب‌سنجی محض و داوری غنی‌شده با کتاب‌سنجی...

صنعتی معتبر ارتقا می‌یابد (Banal-Estañol, Macho-Stadler & Pérez-Castrillo). اگرچه، نتایج پاره‌ای تحقیقات نشان داده است که در ارتباط بین دانشگاه و صنعت، ویژگی‌های فردی محققان تأثیر قوی‌تری نسبت به ویژگی‌های بخش‌ها و دانشگاه‌ها دارد (D'Este & Patel, 2007). این در حالی است که در علوم انسانی، به دلیل ماهیت این حوزه و سنت‌های حاکم بر آن، تمایل بیشتر پژوهشگران به آثار انفرادی یا مشارکت در گروه‌های کوچک است (عرفان‌منش و مروتی اردکانی، ۱۳۹۵؛ رحیمی و فتاحی، ۱۳۸۷). بنابراین، در این حوزه همکاری با بخش غیردانشگاهی از جمله بخش صنعتی بسیار کم انجام می‌شود.

در نظام «کیو-اس»، گرایش‌های موضوعی پیش‌بینی‌کننده معنادار نمره در ابعاد شهرت دانشگاه، نسبت استاد به دانشجو، نسبت اعضای هیئت علمی بین‌المللی به داخلی، نسبت دانشجویان بین‌المللی به داخلی و سرانه استناد هستند. سه بعد شهرت دانشگاه، نسبت استاد به دانشجو و سرانه استناد متأثر از گرایش پزشکی هستند. تأثیر این گرایش بر نمره در بعد سرانه استناد بالاتر از دو بعد دیگر است. گرایش علوم پایه نیز بر نمره در دو بعد سهم اعضای هیئت علمی بین‌المللی و سهم دانشجوی بین‌المللی تأثیر می‌گذارد. بنابراین، می‌توان گفت که هر چه گرایش یک دانشگاه به طرف رشته‌های پزشکی بیشتر باشد، شهرت دانشگاه، نسبت تعداد استاد به دانشجو و میزان استناد آن بالاتر می‌رود و در نتیجه نمره کل دانشگاه در این نظام بالاتر می‌رود. بالعکس، هر چه محور فعالیت یک دانشگاه در زمینه علوم پایه بیشتر باشد، سهم اعضای هیئت علمی و دانشجوی بین‌المللی کمتری جذب می‌کند و در نتیجه، نمره‌ای که در این دو بعد دریافت می‌کند پایین‌تر خواهد بود. شاید بتوان این یافته مرتبط با بعد شهرت در این نظام را در پرتو انتقاداتی که به نظام‌های رتبه‌بندی مشهور چون «تایمز» و «کیواس» وارد شده است، تبیین کرد. این دو نظام از شاخص‌های عینی و ذهنی چون کتاب‌سنجی، شاخص‌های عملکردی و نظرسنجی استفاده می‌کنند. نتیجه چنین روشی این می‌شود که دانشگاه‌هایی که به لحاظ تاریخی پیشرو بوده‌اند، همچنان موقعیت خود را در صدر جدول حفظ می‌کنند (Moed, 2017). همچنین، بنا به نظر Marginson (2009) نظام‌های رتبه‌بندی بازتاب شهرت و قدرت دانشگاه‌ها هستند. از سویی دیگر، دانشگاه‌هایی که بیشتر گرایش‌های موضوعی مهندسی، پزشکی و علوم طبیعی را پوشش می‌دهند، و سبک غالب انتشار در آنها مقاله می‌باشد، در این نظام‌ها نمود بیشتری پیدا می‌کنند (Stratilatis, 2014; Boulton, 2011). درخصوص یافته بعد سرانه استناد در حوزه پزشکی، می‌توان به دلایل پیش‌تر گفته‌شده پیرامون رفتار انتشار و استناد در حوزه علوم پزشکی اشاره کرد. درخصوص یافته‌های بعد شهرت دانشگاه، نسبت استاد به دانشجو، سهم اعضای هیئت علمی بین‌المللی و دانشجوی بین‌المللی تبیین دلایل زیربنایی به پژوهش‌های بیشتر نیاز دارد.

در «لایدن»، نمره‌ها در ابعاد نسبت انتشارات ۱۰ درصد برتر، میانگین نمره استناد، میانگین نمره استناد نرمال‌شده، نسبت انتشارات مشترک بین‌سازمانی، نسبت انتشارات مشترک بین‌المللی، نسبت انتشارات (همکاری با فاصله جغرافیایی طولانی) بیشتر از ۱۰۰۰ کیلومتر متأثر از گرایش پزشکی است. با توجه به مقدار ضریب بتا، قدرت پیش‌بینی این گرایش در بعد نسبت انتشارات مشترک بین‌سازمانی بالاتر از ابعاد دیگر می‌باشد. یک دلیل می‌تواند این باشد که در حوزه علوم پزشکی، مسائل فرامرزی بیشتری وجود دارد و بنابراین مشارکت آنها در سطح جهانی بیشتر است. این بدین معناست که در قطب‌های مختلف علمی جهان، مسائل پزشکی یا مشترک است یا می‌توان از همکاری پژوهشگران دوردست برای حل این مسائل استفاده کرد. در نتیجه پزشکی گرایش بیشتری به همکاری در مرزهای دورتر را دارد. در حوزه‌هایی که برای انجام پژوهش، نیاز به ابزار و امکانات آزمایشگاهی، بودجه‌های کلان و تعداد زیاد پژوهشگر است، همکاری علمی ضرورت بیشتری می‌یابد (Henriksen, 2016). از جمله در رشته‌های پزشکی،

نیاز به امکانات آزمایشگاهی و تحقیقاتی، ضرورت همکاری‌های علمی بین سازمانی و بین‌المللی (ستارزاده و همکاران، ۱۳۹۵؛ عرفان‌منش و همکاران، ۱۳۹۲) را افزایش می‌دهد.

بعد نسبت انتشارات مشترک دانشگاه با بخش صنعت در نظام لایدن نیز، متأثر از گرایش علوم انسانی و فنی و مهندسی است. قدرت پیش‌بینی‌کنندگی علوم انسانی نسبت به فنی و مهندسی بیشتر و در جهت عکس است. بدین ترتیب، هر چقدر گرایش یک دانشگاه به حوزه علوم انسانی بیشتر باشد نسبت انتشارات مشترک آن با بخش صنعت کمتر خواهد بود. این یافته، با یافته‌های تحقیق در بعد درآمد صنعتی در نظام «تایمز» همسو است. در حوزه فنی و مهندسی، بین آفرینش علمی و صنعتی کردن و تجاری‌سازی علم رابطه مستقیم و مؤثری وجود دارد (ریاحی و نوروزی، ۱۳۹۰). این یافته تا حدودی مؤید ادعای Boulton (2011) و Stratilatis (2014) است که در دانشگاه‌هایی با گرایش موضوعی مهندسی، پزشکی و علوم طبیعی، سبک غالب انتشار مقاله در این نظام‌ها نمود بیشتری پیدا می‌کند. یکی از پیامدهای چنین رویکردی نادیده‌گرفتن حوزه‌هایی مانند هنر، علوم انسانی و اجتماعی است (Rauhvargers, 2013).

به‌طور کلی، نظام‌های رتبه‌بندی با روش‌شناسی‌های متفاوت، نتایج نسبتاً مشابهی به دست می‌آورند. بخشی از این شباهت به کاربرد ابعاد عملکردی یکسان در ارزیابی‌ها بازمی‌گردد و بخشی دیگر را می‌توان به نتایج مشابه حاصل از روش‌های متفاوت نسبت داد. برای نمونه، همبستگی «رف» با «تایمز» و «کیو-اس» بیش از «شانگهای» است. شانگهای از سنت متفاوتی پیروی می‌کند: عامل انسانی را حتی در سطح نظرسنجی حذف کرده است؛ درعین حال، به‌کارگیری شاخص‌هایی مانند جایزه نوبل و مدال فیلدز، انتشار مقاله در مجلات ساینس و نیچر و نشر آثار در حوزه علوم انسانی و اجتماعی آن را از «تایمز» و «کیو-اس» متمایز می‌کند. بنابراین، «تایمز» و «کیو-اس» از لحاظ شاخص‌های کتاب‌سنجی به نظام «رف» نزدیک‌تر هستند. همچنین، قضاوت انسانی را به شکل نظرسنجی به خدمت می‌گیرند. گرایش موضوعی تأثیری چشمگیر بر نمره کلی دانشگاه‌ها در نظام‌های رتبه‌بندی یا نمره آنها در ابعاد ندارد. تنها، در «کیو-اس» نمره کل دانشگاه‌ها متأثر از گرایش پزشکی است. همچنین، گرایش‌های پزشکی و مهندسی گاه در جهت عکس هنر، علوم انسانی و اجتماعی بر نمره‌های برخی ابعاد تأثیر می‌گذارند. بنابراین، ارزیابی دانشگاه‌های با گرایش‌های موضوعی متفاوت در کنار یکدیگر می‌تواند تا حدودی نتایج را متأثر سازد.

پیشنهادهای اجرایی پژوهش

۱. با توجه به نتایج تقریباً مشابه (همبستگی بسیار قوی) بین سیستم‌های مبتنی بر کتاب‌سنجی غنی‌شده با نظرسنجی با سیستم مبتنی بر داوری غنی‌شده با کتاب‌سنجی، تا آنجایی که دانشگاه‌های انگلیس یا دانشگاه‌هایی با عملکرد شبیه به دانشگاه انگلیس مدنظر باشد، می‌توان جهت کاهش هزینه‌های داوری و همچنین کاهش زمان ارزیابی از سامانه‌های مبتنی بر کتاب‌سنجی غنی‌شده با نظرسنجی استفاده کرد؛
۲. نمره دانشگاه‌ها تا حدودی متأثر از گرایش موضوعی آنهاست. بنابراین، اصلاح ساختار نظام‌ها برای رفع این اختلاف هر چند اندک توصیه می‌شود. برای اصلاح می‌توان ضرایب وزنی را برای گرایش‌های موضوعی در ابعادی که سوگیری موضوعی را نشان داده‌اند اعمال کرد؛
۳. با توجه به تأثیر -هر چند اندک- گرایش موضوعی بر نمره دانشگاه‌ها دقت لازم در تفسیر نتایج یا برنامه‌ریزی بر اساس آنها ضروری است. به‌ویژه، می‌بایست نتایج این ارزیابی‌ها برای مقایسه دانشگاه‌های جامع، تخصصی و

صنعتی با احتیاط به کار گرفته شود.

پیشنهاد برای پژوهش‌های آتی

۱. گرایش موضوعی بر نمره دانشگاه‌ها در برخی ابعاد (مانند شهرت دانشگاه، نسبت استاد به دانشجو و سهم اعضای هیئت علمی و دانشجوی بین‌المللی) مؤثر دیده شد. روشن شدن دلایل زیربنایی این مسئله نیاز به پژوهش‌های بیشتر دارد؛
۲. بین نظام‌هایی با روش‌شناسی‌های متفاوت همبستگی متوسط تا قوی وجود دارد. از آنجاکه روش‌شناسی این نظام‌ها از ابعاد مختلف مانند نوع شاخص، ضریب وزنی، نظرسنجی یا داوری، بازه زمانی و پایگاه‌های داده‌ای با هم متفاوت‌اند، تحقیقات بیشتر برای آگاهی از اینکه شدت همبستگی مشاهده‌شده تا چه اندازه متأثر از هر یک از این عوامل است ضروری است؛
۳. تکرار این پژوهش بر دیگر نظام‌های مبتنی بر داوری مانند نظام ایتالیا ضمن آزمون قابلیت اطمینان نتایج این پژوهش، می‌تواند وضعیت دیگر نظام‌های ارزیابی ملی را روشن کند.

فهرست منابع

- پوریزدیان، مهدی؛ کرمی، مرتضی (۱۳۹۸). ارزیابی نظام‌های رتبه‌بندی دانشگاه‌ها جهت ارزیابی عملکرد پژوهشی دانشگاه با رویکرد تعالی سازمانی. نخستین کنفرانس ملی علوم انسانی و توسعه. شیراز، ۱۶ ص.
- خانی‌زاد، رحیم؛ منتظر، غلامعلی (۱۳۹۶). ارزیابی تطبیقی نظام‌های رتبه‌بندی دانشگاه‌های جهان. سیاست علم و فناوری، ۳۱-۴۶، (۳) ۹. DOI: 10.22034/jstp.2017.9.3.537781
- خسروجردی، محمود و ندا زراعت‌کار (۱۳۹۱). مروری بر نتایج هفت نظام رتبه‌بندی دانشگاه‌های جهان. پژوهش‌نامه پردازش و مدیریت اطلاعات، ۲۸ (۱): ۷۱-۸۴.
- رحیمی، ماریه؛ فتاحی، رحمت‌الله (۱۳۸۷). بررسی وضعیت همکاری علمی اعضای هیئت علمی در چهار حوزه موضوعی در دانشگاه فردوسی مشهد. فصلنامه کتابداری و اطلاع‌رسانی، ۱۱ (۲) ص. ۹۵-۱۲۰.
- ریاحی، عارف؛ نوروزی، علیرضا (۱۳۹۰). بررسی میزان همکاری‌های علمی کشورهای حوزه خلیج فارس با کشورهای اسکاندیناوی در پایگاه اطلاعاتی اسکاپوس طی سال‌های ۱۹۸۹-۲۰۰۹. رهیافت (شماره ۴۸).
- ستارزاده، اصغر؛ گلین‌مقدم، گلنسا؛ مؤمنی، عصمت (۱۳۹۵). تحلیل ساختار شبکه همکاری‌های علمی پژوهشگران حوزه علوم پایه پزشکی ایران در نمایه استنادی علوم در بازه زمانی ۱۹۹۶-۲۰۱۳. فصلنامه مطالعات دانش‌شناسی، ۲ (۶) ص. ۱-۲۰. DOI: 10.22054/jks.2016.2707
- عرفان‌منش، محمدامین؛ روحانی، والاعلی؛ بصیریان جهرمی، رضا؛ غلامحسین‌زاده، زهره (۱۳۹۲). بررسی مشارکت پژوهشگران روان‌شناسی و روان‌پزشکی کشور در تولید علم. پژوهش‌نامه پردازش و مدیریت اطلاعات، ۲۹ (۱) ص. ۱۶۳-۱۳۷.
- عرفان‌منش، محمدامین؛ مروتی اردکانی، مرضیه (۱۳۹۵). مطالعه علم‌سنجی و تحلیل شبکه‌های همکاری علمی در

فصلنامه مطالعات میان‌رشته‌ای در علوم انسانی. فصلنامه مطالعات میان‌رشته‌ای در علوم انسانی، ۸ (۴). ص. ۷۷-۵۵.
DOI: 10.22035/isih.2016.230

مشتاق، مریم، ستوده، هاجر، یقطین، مریم و جوکار، طاهره (۱۴۰۰). همبستگی نتایج سامانه‌های رتبه‌بندی نمایه نیچر، لایدن با تایمز و کیو-اس. پژوهش‌نامه علم‌سنجی، ۷ (۱۴). ۱۷۲-۱۵۷.
DOI: 10.22070/rsci.2020.5488.1384

نورمحمدی، حمزه و صفری، فاطمه (۱۳۹۲). معرفی نظام‌های رتبه‌بندی جهانی دانشگاه‌ها و بررسی شاخص‌های این نظام‌ها. سیاست‌نامه علم و فناوری، ۳(۲): ۷۱-۸۶. DOR: 20.1001.1.24767220.1392.03.2.7.1

Abramo, G., D'Angelo, C. A., & Caprasecca, A. (2009). Allocative efficiency in public research funding: Can bibliometrics help? *Research policy*, 38(1), 206-215. doi: 10.1016/j.respol.2008.11.001

Abramo, G., & D'Angelo, C. A. (2011). Evaluating research: from informed peer review to bibliometrics. *Scientometrics*, 87(3), 499-514. doi: 10.1007/s11192-011-0352-7

Aksnes, D. W., & Taxt, R. E. (2004). Peer reviews and bibliometric indicators: a comparative study at a Norwegian university. *Research evaluation*, 13(1), 33-41. doi: 10.3152/147154404781776563

Baccini, A., & De Nicolao, G. (2016). Do they agree? Bibliometric evaluation versus informed peer review in the Italian research assessment exercise. *Scientometrics*, 108(3), 1651-1671. doi:10.1007/s11192-016-1929-y

Banal-Estañol, A., Macho-Stadler, I., & Pérez-Castrillo, D. (2013). Research output from university-industry collaborative projects. *Economic Development Quarterly*, 27(1), 71-81. doi: 10.1177/0891242412472535

Bertocchi, G., Gambardella, A., Jappelli, T., Nappi, C. A., & Peracchi, F. (2015). Bibliometric evaluation vs. informed peer review: Evidence from Italy. *Research Policy*, 44(2), 451-466. doi: 10.1016/j.respol.2014.08.004

Boulton, G. (2011). University rankings: Diversity, excellence and the European initiative. *Procedia-Social and Behavioral Sciences*, 13, 74-82. doi: 10.1016/j.sbspro.2011.03.006

Box, S. (2010). Performance-based funding for public research in tertiary education institutions: Country experiences. *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*, OECD Publishing. Dec. 01, 2010. Paris, 188 p. doi: 10.1787/9789264094611-en

Butler, L. (2010). Impacts of performance-based research funding systems: A review of the concerns and the evidence. *Performance-based Funding for Public Research in Tertiary Education Institutions: Workshop Proceedings*, OECD Publishing. Dec. 01, 2010. Paris, 188 p. doi: 10.1787/9789264094611-en

Chubin, D. E., Hackett, E. J. (1990). *Peerless science: Peer review and US science policy*. Suny Press. New York, USA. 257 P.

- Chubin, D. E. (1994). Grants peer review in theory and practice. *Evaluation Review*, 18(1), 20-30. doi: 10.1177/0193841X9401800103
- De Boer, H., Jongbloed, B., Benneworth, P., Cremonini, L., Kolster, R., Kottmann, A., & Vossensteyn, H. (2015). Performance-based funding and performance agreements in fourteen higher education systems. *Center for Higher Education Policy Studies*. <https://research.utwente.nl/en/publications/>
- D'Este, P., & Patel, P. (2007). University–industry linkages in the UK: What are the factors underlying the variety of interactions with industry?. *Research policy*, 36(9), 1295-1313. doi: 10.1016/j.respol.2007.05.002
- Erfanmanesh, M., Rohani, V. A., Jahromi, R. B., & Gholamhosseinzadeh, Z. (2014). Investigating scientific collaboration of Iranian psychology and psychiatry researchers. *Iranian Journal of Information Processing and Management*, 29(1), 137-163. [In Persian].
- Erfanmanesh, M. A., & Morovati Ardakani, M. (2016). A scientometrics and collaboration network analysis of the quarterly journal of interdisciplinary studies in the humanities. *Interdisciplinary Studies in the Humanities*, 8(4), 55-77. doi: 10.22035/isih.2016.230 [In Persian].
- Fauzi, M. A., Tan, C. N. L., Daud, M., & Awalludin, M. M. N. (2020). University rankings: A review of methodological flaws. *Issues in Educational Research*, 30(1), 79-96.
- Franceschet, M., & Costantini, A. (2011). The first Italian research assessment exercise: A bibliometric perspective. *Journal of informetrics*, 5(2), 275-291. doi: 10.1016/j.joi.2010.12.002
- Geuna, A., & Martin, B. R. (2003). University research evaluation and funding: An international comparison. *Minerva*, 41(4), 277-304. doi: 10.1023/B:MINE.0000005155.70870.bd
- Geuna, A., & Piolatto, M. (2016). Research assessment in the UK and Italy: Costly and difficult, but probably worth it (at least for a while). *Research Policy*, 45(1), 260-271. doi: 10.1016/j.respol.2015.09.004
- Goldfinch, S., & Yamamoto, K. (2012). Prometheus Assessed? Research Measurement, Peer Review, and Citation Analysis. *Chandos Publishing*. Minneapolis, USA. 384 p.
- Henriksen, D. (2016). The rise in co-authorship in the social sciences (1980–2013). *Scientometrics*, 107(2), 455-476. doi: 10.1007/s11192-016-1849-x
- Hicks, D. (2012). Performance-based university research funding systems. *Research policy*, 41(2), 251-261. doi: 10.1016/j.respol.2011.09.007
- Hood, W., & Wilson, C. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*, 52(2), 291-314. doi: 10.1023/a:1017919924342
- Jonkers, K., & Zacharewicz, T. (2016). Research Performance Based Funding Systems: a Comparative Assessment. *Publications Office of the European Union*. Luxembourg, Luxembourg. 108 p.
- Jonkers K and Zacharewicz T. Research Performance Based Funding Systems: a Comparative Assessment . EUR 27837. Luxembourg (Luxembourg);: 2016

- Khanizad, R., & Montazer, G. (2017). A Comparative Evaluation of the World University Rankings Systems. *Journal of Science and Technology Policy*, 9(3), 31-43. doi: 10.22034/jstp.2017.9.3.537781 [In Persian].
- Khosrowjerdi, M., & Zeraatkar, N. (2012). A review of outcomes of seven world university ranking systems. *Iranian Journal of Information processing and Management*, 28(1), 71-84. [In Persian].
- Kostoff, R. N. (1995). Federal research impact assessment: Axioms, approaches, applications. *Scientometrics*, 34(2), 163-206. doi: 10.1007/BF02020420
- Marginson, S. W. M. (2009). University rankings, government and social order. *In: Re-Reading Education Policies*. Brill. 584-604 pp. doi: 10.1007/BF02020420
- Moed, H. F. (2017). A critical comparative analysis of five world university rankings. *Scientometrics*, 110(2), 967-990. doi: 10.1007/978-3-319-60522-7_18
- Moshtagh, M., Sotudeh, H., Yaghtin, M., & Jowkar, T. (2021). The Correlation of Nature and Leiden Index Ranking Systems with Times and QS. *Scientometrics Research Journal*, 7(14), 157-172. doi: 10.22070/rsci.2020.5488.1384 [In Persian].
- Mryglod, O., Kenna, R., Holovatch, Y., & Berche, B. (2013). Comparison of a citation-based indicator and peer review for absolute and specific measures of research-group excellence. *Scientometrics*, 97(3), 767-777. doi: 10.1007/s11192-013-1058-9
- Nourmohammadi, H. A., & Safari, F. (2013). Introduction the global rankings of universities and review criteria of this system. *Science and Technology Policy Letters*, 3(2), 71-86. DOR: 20.1001.1.24767220.1392.03.2.7.1 [In Persian].
- Oberski, J. E. J. (1988). Some Statistical Aspects Of Co-Citation Cluster Analysis And A Judgment By Physicists. *In: Handbook of quantitative studies of science and technology*. Elsevier. 431-462 pp.
- Phillips, M., & Maes, K. (2012). Research Universities and Research Assessment, Position Paper for the League of European Research Universities (LERU). <https://www.leru.org/publications/research-universities-and-research-assessment>
- Pooryazdian, m., Karami, m (2019) Evaluation of university ranking systems to evaluate the university's research performance with the organizational excellence approach. *The first national conference of humanities and development*. Shiraz, 16p. [In Persian].
- Priem, J., & Hemminger, B. H. (2010). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*, 15(7). doi: 10.5210/fm.v15i7.2874
- Radicchi, F., & Castellano, C. (2012). A reverse engineering approach to the suppression of citation biases reveals universal properties of citation distributions. *PLoS One*, 7(3), e33833. doi: 10.1371/journal.pone.0033833
- Rahimi, M., & Fatahi, R. (2008). A survey of scholarly collaborations among academic staff of Ferdowsi University of Mashhad. *Library and Information Science*, 11(2), 95-120 [In Persian].
- Rauhvargers, A. (2013). Global university rankigs and their impact: Report II. European University Association. Brussels, Belgium. 82 p.

- Rinia, E. J., Van Leeuwen, T. N., Van Vuren, H. G., & Van Raan, A. F. (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria: Evaluation of condensed matter physics in the Netherlands. *Research policy*, 27(1), 95-107. doi: 10.1016/S0048-7333(98)00026-2
- Riyahi, A., & Norouzi, A. (2011). The survey of the scientific cooperation of the Gulf states with the Scandinavian countries in the Scopus database during 1989-2009. *Rahyaft*, 21(48), 91-110. [In Persian].
- Robbins, P. T., Wield, D., & Wilson, G. (2017). Mapping engineering and development research excellence in the UK: An analysis of REF2014 impact case studies. *Journal of International Development*, 29(1), 89-105. doi: 10.1002/jid.3255
- Sattarzadeh, A., Galyani Moghaddam, G., & Momeni, E. (2016). The Analysis of the structure of scientific collaboration networks in basic medical sciences in the Science Citation Index from 1996 to 2013. *Knowledge Retrieval and Semantic Systems*, 3(6), 1-20. doi: 10.22054/jks.2016.2707 [In Persian].
- Selten, F., Neylon, C., Huang, C. K., & Groth, P. (2020). A longitudinal analysis of university rankings. *Quantitative Science Studies*, 1(3), 1109-1135. doi: 10.1162/qss_a_00052
- Smolinsky, L., & Lercher, A. (2012). Citation rates in mathematics: A study of variation by subdiscipline. *Scientometrics*, 91(3), 911-924. doi: 10.1007/s11192-012-0647-3
- Stern, N. (2016). Building on success and learning from experience: an independent review of the Research Excellence Framework. <http://eprints.lse.ac.uk/87988/>
- Stratilatis, C. (2014). University rankings and the scientification of social sciences and humanities. *Ethics in Science and Environmental Politics*, 13(2), 177-192. doi: 10.3354/esep00144
- Taylor, J. (2011). "The assessment of research quality in UK universities: peer review or metrics?" *British Journal of Management*, 22(2): 202-217. doi: 10.1111/j.1467-8551.2010.00722.x
- van Raan, A. F. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), 133-143. doi: 10.1007/s11192-005-0008-6
- Wilsdon, J. (2016). *The metric tide: independent review of the role of metrics in research assessment and management*. Sage. London, UK. 167 p.