PAPER NAME

Achsan_2020_J._Phys. _Conf._Ser._1566_012009.pdf

AUTHOR

Harry T. Y. Achsan

WORD COUNT

4221 Words

CHARACTER COUNT

22514 Characters

PAGE COUNT

10 Pages

FILE SIZE

972.6KB

SUBMISSION DATE

Feb 9, 2023 2:42?PM GMT+7

REPORT DATE

Feb 9, 2023 2:43?PM GMT+7

18% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 18% Internet database
- 1% Publications database
- Crossref Posted Content database
- 0% Submitted Works database

Excluded from Similarity Report

- Crossref database
- Bibliographic material
- Quoted material
- Cited material

An Approach for Measuring Research Strength Map of an Institution

An Approach for Measuring Research Strength Map of an Institution

H T Y Achsan[1*], H Suhartanto[1], W C Wibowo[1] and W T H Putri[2]

[1] Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
[2] Faculty of Science & Engineering, Universitas Paramadina, Jakarta, Indonesia

*Email: harry.tursulistyono71@ui.ac.id

Abstract. The real research strength of an institution is important for establishing research priorities and developing an institutional research plan. It can also be used as a basis for SWOT analysis, developing vision-mission of an institution and also useful for research funder in distributing their research grants. Unfortunately, there was only a small number of researches found. In this article, we proposed a new method to quantify the quality of research outputs. The method is suitable for measuring research strength. We applied the method to an institution using research articles downloaded from a reputable document database/index. The results were research strength map in 27 subject areas and 337 subject categories.

1. Introduction

The research output of an institution is vital to its competitiveness and standing in the grading of universities; it is an increasingly substantial part of the resources that distinct institutions have at their disposal. Higher education research is central for knowledge generation, occupying a serious position in endorsing a nation's prosperity and its citizens' well-being in this era [1]. Research also provides for the economic growth of the nation and advantageously positions the national economy in the internationally modest knowledge economy. Whether it is clearly acknowledged, or only implied in policy, the international attractiveness of education institution research is a crucial condition for the competitiveness of the nationwide innovation system [2].

In this era when the Indonesian government is employing emphasis on research as a significant motor for driving the economy and the knowledge society, the effective management of research has become a main contemporary issue for institutions [3]. Each institution needs to formulate its vision/mission based on a thorough evaluation of research quality, thus the need for a research strength map. Research strength map also needed as the foundation for establishing research plans and research priorities as stated in Indonesia National Research Master Plan (Rencana Induk Riset Nasional) 2017-2045.

Unfortunately, in our inquiries, we have not found articles about the research strength map for Indonesian institutions. Therefore, we try to develop a technique for measuring the research strength map of an institution. To do it, we apply the research to an institution as a case study. The institution was Universitas Indonesia (University of Indonesia/UI). The reason is that Universitas Indonesia ranked the highest among any other Indonesian higher education institutions in 2019, based on Times Higher Education World University Rankings [4]. It is also ranked as the best research institution in Indonesia, based on Indonesian Science and Technology Index (Sinta) for the last 3 years [5]. Although we chose UI for our case study, our research is also applicable to any institution.

The methods that we employ in this research were web mining and scientometrics. The web mining part for this research was done through the means of a focused web crawler to gather the available data from the web. Scientometrics then employed to assess the research output quality and to reveal the research strength.

2. Related works

Uddin et.al [6] presented on their paper a sciento-text framework, a way to be able to fine-grained characterize and assess the research performance of institutions. They devise a way to strongly identify research themes of a subject. The framework consists of standard scientometric and text analytics components. They did this by systematically classified data into different areas by theme and followed by standard scientometric methodology. They chose the computer science domain from WoS. The data came from 530 institutions/research organizations worldwide. After data cleaning processes, they are left with 444 institutions and 498,488 publications. The two main parts of the framework are thematic area classification and selecting and measuring performance indicators. Their paper used the taxonomy for the computer science field comprised of Microsoft Academic Search (MAS) for thematic area classification. Furthermore, they used WoS data for keyword extraction so that it aligned well with their data collection. Next, they select ten criteria for measuring performance indicators of each institution. Some of these indicators are also being used by well-known ranking schemes. The list of indicators is publication based, citation-based, and collaboration. Their result was around 92 percent accuracy rate for thematic classification. For the determination of research strength, the paper used field normalization to see the comparative strength in the given thematic area. The approach is due to the variation in productivity levels for different thematic areas. Their result was artificial intelligence ranked as the highest theme, followed by algorithms & theory, and networks & communication.

Another work by González-Albo et.al [7] in the context of CSIC (Spanish National Research Council), uses bibliometric indicators for the analysis of the research performance of a multidisciplinary institution. Their work analyses CSIC scientific activity, a national center that conducts research in all fields of knowledge, both basic and advanced research. The CSIS consists of 7 centers and 128 institutes, which manages an annual budget of EUR 737.1 Million [8]. In the article, they use absolute indicators of activity and impact and showed relative indicators to compare CSIC's research activity against overall national research in different areas or disciplines. Their sources of data are the CSIC annual report, and the WoS database. The bibliometric indicators used in this article are activity indicators, impact indicators, level of research, and scientific collaboration. Their result was the greatest number of articles were in Physics, followed by Agriculture/Biology/Environment, Biomedicine, and Chemistry.

Another article by Johnes & Yu [9], used data envelopment analysis (DEA) to measure the research performance of Chinese higher education institutions. They stated that DEA has become popular tools for measuring the efficiency of non-profit institutions such as hospitals, schools, and universities. DEA is a non-parametric linear programming (LP) technique, they combined DEA with stochastic frontier analysis (SFA) that allowed them to draw statistical inferences from the result. They used the data from Chinese university rankings, that have been available for nine consecutive years. Their result was divided into three main areas, geographical location is significantly related to efficiency, Higher Education Institutions (HEI) administration method of central or local in nature, and lastly, comprehensive universities consistently have higher average efficiency than specialist institutions.

3. Methods

This work aims to determine the research strength of an institution by its research output. Research outputs usually published in the form of scientific articles; therefore, this work requires scientific articles published by the institution. Since many scientific articles are not free, we then use bibliometric data. Bibliometrics is the science of bibliography, and a bibliography is a list of references or a list of articles or lists of documents. References or bibliography contains the author's name, title, year of publication, publisher, etc. All of those are parts of documents metadata. This research utilized documents metadata, especially documents of scientific articles to measure the research strength of an institution.

3.1. Data gathering

We choose Scopus as the main source of the document's metadata. Scopus is one of the bibliometrics, citations and abstracts databases that have become typical in the field of bibliometrics and scientometrics [10, 11, 12]. The justification of choosing Scopus among others are: a) Scopus provides a tool to download many (as many of 2000) documents metadata simultaneously, b) Scopus provides documents metadata that has more than 40 features (metadata fields) like abstract, author keywords, index keywords, author's affiliation, etc., and c) it has advance search engine that satisfies this work.

By using Scopus advanced search engine, we have downloaded all documents metadata published by Universitas Indonesia during the period of 2009-2018 in July 2019. We developed a small focused web crawler software to automate the downloading process. The software was plugged-in into the Chrome web browser to mimic a person. Mimicking a person or a web browser user is important in at least two reasons: a) to avoid flooded Scopus web servers that prevent the crawler from denial of service attack, and b) to prevent our Scopus subscription from being banned. The downloaded files comprised 11,031 rows of documents metadata. Most parts of the document metadata were journal articles (50%), articles in proceedings (44%) as seen in Table 1.

Table 1. Documents metadata downloaded from Scopus.

Year	Article	Conf. Paper	Review	Book Chapter	Article in Press	Book	Editorial	Erratum	Letter	Note	Short Survey	Total
2009	190	44	7	21			4			1		267
2010	203	62	20	8		1	6	2	2	1	1	306
2011	262	129	19	14		3	3	2	2	3		437
2012	298	184	23	20			6	2				533
2013	345	226	22	23		2	10	4		1		633
2014	387	202	17	14		1	11		1	3		636
2015	582	145	28	18		1	11	1	7	4		797
2016	736	383	45	31	3	2	9	2	10	3		1,224
2017	1,251	1,218	55	29	1	8	22	4	1	1	1	2,591
2018	1,220	2,239	51	46	21	3	15	1	7	4		3,607
Total	5,474	4,832	287	224	25	21	97	18	30	21	2	11,031

To measure the quality of research outputs, we also retrieved journals metadata from Scimago Journal Ranks (SJR). SJR ranks reputable journals all over the world. SJR also categorizes journals in the form of Q1, Q2, Q3, and Q4. Journals categorized as Q1 are the most reputable journals, and Q4 means the least reputable journals.

3.2. Scientometrics

Scientometrics was initiated in the 1920s when information about citation rates can be used by librarians to make procurement decisions on academic journals for libraries with a restricted budget. Over time, scientometrics has advanced into an established interdisciplinary area of research that can be applicable to entirely natural and social sciences disciplinary research [13]. Garfield [14] recommended that a citation count of articles be more efficient than counting the number of articles for scholars' productivity. Based on Garfield's work, Price [15] made it possible for scientometrics to develop discipline over-analyzing huge citation data. Through numerical modeling, this study revealed how scientific networks were linked through published scientific articles in natural sciences. It also exposed that citation-based analysis was able to recognize the "nature of the scientific research fronts" for any discipline.

In addition, citation amounts are often considered to have predictive ability. Garfield found that the predictable factor of Nobel Prize winners on the scientific community is reflected expressively in their citation records long before they obtain the prizes [16]. In current years, citation rates are becoming progressively important in judging the research quality of individual faculty members, journals, departments and institutions [17].

3.3. Measuring publication quality

SJR has ranks international reputable journals since years ago. SJR also categorizes journals in the form of Q1, Q2, Q3, and Q4 as mentioned above. The journal ranks and categories were used by many researchers to quantify the productivity, quality, and/or performance of research by its output [18, 19, 20, 21, 22]. We used journal quartile to quantify the quality of the research article as used by Uddin et.al [6] in their research:

$$Q_d = \sum_{i=1}^{N_d} k\,(5 - q_i) \quad (1)$$

where Q is the quality, k (we used k=5) is a weighting constant and q_i is the articles' quartile. The value of q_i is in the range of 1 to 4, $q_i = 1$ is the first quartile or the best quality journal. While d can be applied to a document, an individual, a department, an institution, or a country. Here we applied d as an institution, so N_d is the number of documents authored by a researcher affiliated to the institution that indexed by Scopus. One weakness of the above formula is the formula cannot measure the quality of research articles published in proceedings because proceedings have no quartile. The formula cannot be used to measure the research strength of Universitas Indonesia because almost half (44%) of its research outputs published in proceedings.

To overcome the problem, we proposed a new method. Since citation count (denote as ci) can be used in the assessment of research strength, then we add citation count to formula (1). After adding the citation count, the formula become:

$$Q_d = \sum_{i=1}^{N_d} [k\,(5 - q_i) + c_i] \quad (2)$$

The above formula can reckon articles in proceedings, it will return non zero for articles in proceedings that have been cited.

The formula (2) can also be used to quantify the quality of research papers by its Subject Area or by its Subject Category. The source of research papers like journals and proceedings are categorized by its Subject Area. Some researchers used the Subject Area to assess research excellence [23, 24], research productivity [25], and research quality [25, 26, 27, 28]. We assess institution research output quality in more detail using Subject Category. Subject Category is a sub-category of the Subject Area. In this research, we used Subject Areas and Subject Categories developed by SJR. SJR categorized journals and proceedings into 27 Subject Areas. Those subject areas sub-categorized into 337 Subject Categories.

3.4. Strength map

To visualize the strength map, we use a Radar chart. The data used in the Radar chart was from formula (2) that applied to each Subject Area and Subject Category. A visual strength map has the advantage to deliver information quickly. But it has a disadvantage cannot inform accurately. To overcome the problem, we also display the research strength of an institution in tables.

4. Results

We reveal Universitas Indonesia's research strength after implementing equation (2) to the downloaded metadata, see Table 2. The table shows the quantified quality of all subject areas of research output articles authored by researchers affiliated to Universitas Indonesia in the last 5 years. The visualization of the table can be seen in Figure 1. Figure 1 is a research strength map of Universitas Indonesia. Since the top 5 of subject areas ("Medicine", "Engineering", "Computer Science", "Physics and Astronomy", and "Business, Management & Accounting") make other subject areas look small, then we removed it and redraw the research strength map, the result displayed in Figure 2.

To see more detail of the research strength of Universitas Indonesia, we choose one subject area "Computer Science". Its research strength is distributed into 12 subject categories, see Table 3. Figure 3 shows the research strength map for the subject area "Computer Science".

Table 2. The research strength of Universitas Indonesia based on its Subject Areas for the last 5 years.

No	Subject Area	Year					Total
		2014	2015	2016	2017	2018	
1	Medicine	6,782	9,476	11,054	20,036	21,258	68,606
2	Engineering	2,333	4,768	6,063	11,528	15,307	39,999
3	Computer Science	2,145	1,770	3,723	9,102	10,578	27,319
4	Physics and Astronomy	905	794	2,199	6,953	10,190	21,041
5	Business, Management and Accounting	990	3,757	4,136	5,933	5,833	20,648
6	Materials Science	316	1,021	1,167	4,525	6,630	13,659
7	Social Sciences	1,271	1,409	2,104	4,197	4,160	13,140
8	Environmental Science	710	820	1,140	3,475	6,943	13,088
9	Pharmacology, Toxicology and Pharmaceutics	708	762	1,339	4,747	4,353	11,909
10	Energy	587	364	1,462	3,117	5,375	10,905
11	Biochemistry, Genetics and Molecular Biology	873	882	2,230	2,492	4,190	10,667
12	Agricultural and Biological Sciences	588	769	1,141	2,238	2,560	7,296
13	Earth and Planetary Sciences	264	247	354	732	4,875	6,473
14	Chemistry	487	639	617	1,244	3,193	6,181
15	Immunology and Microbiology	975	760	657	1,160	1,738	5,290
16	Chemical Engineering	237	285	354	1,227	2,282	4,384
17	Mathematics	355	617	813	1,493	950	4,228
18	Nursing	353	749	671	658	1,683	4,113
19	Economics, Econometrics and Finance	277	369	547	1,059	1,177	3,429
20	Dentistry	76	41	725	1,391	1,110	3,343
21	Arts and Humanities	224	140	333	469	1,083	2,249
22	Multidisciplinary	63	911	191	178	445	1,788
23	Neuroscience	115	42	384	358	820	1,720
24	Psychology	107	84	209	334	613	1,348
25	Decision Sciences	122	135	299	306	437	1,298
26	Health Professions	56	118	28	197	308	706
27	Veterinary	0	7	54	112	175	348

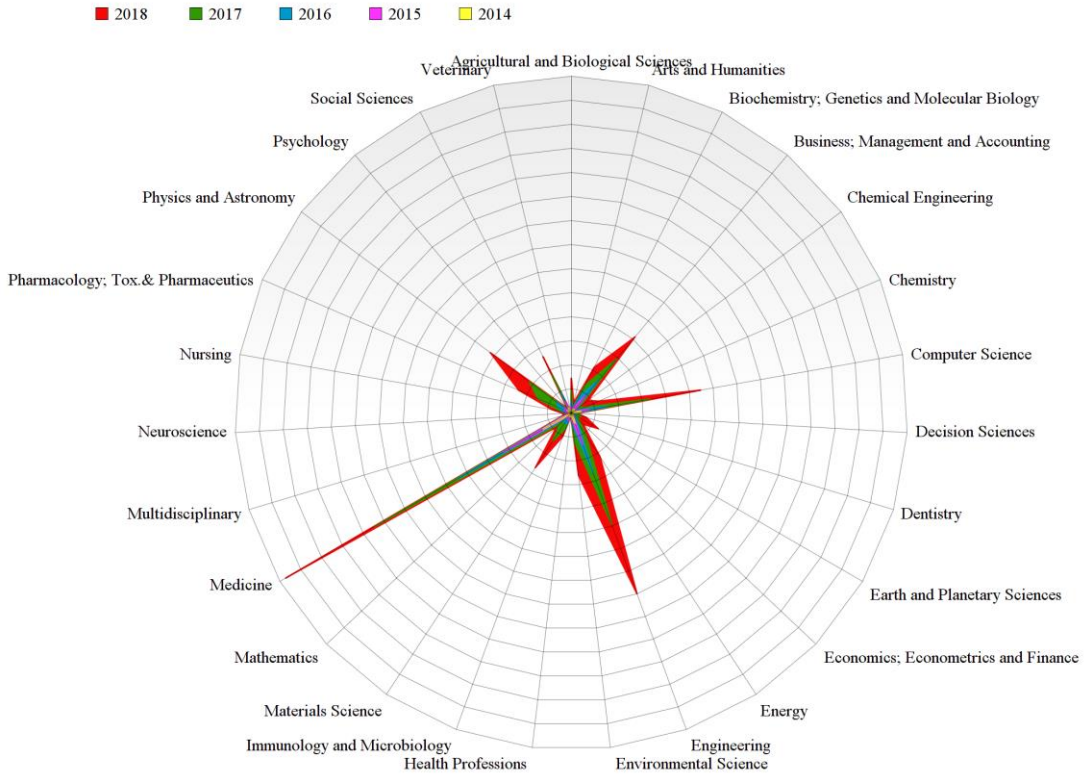

Figure 1. Visualization of research strength map of Universitas Indonesia for the last 5 years.

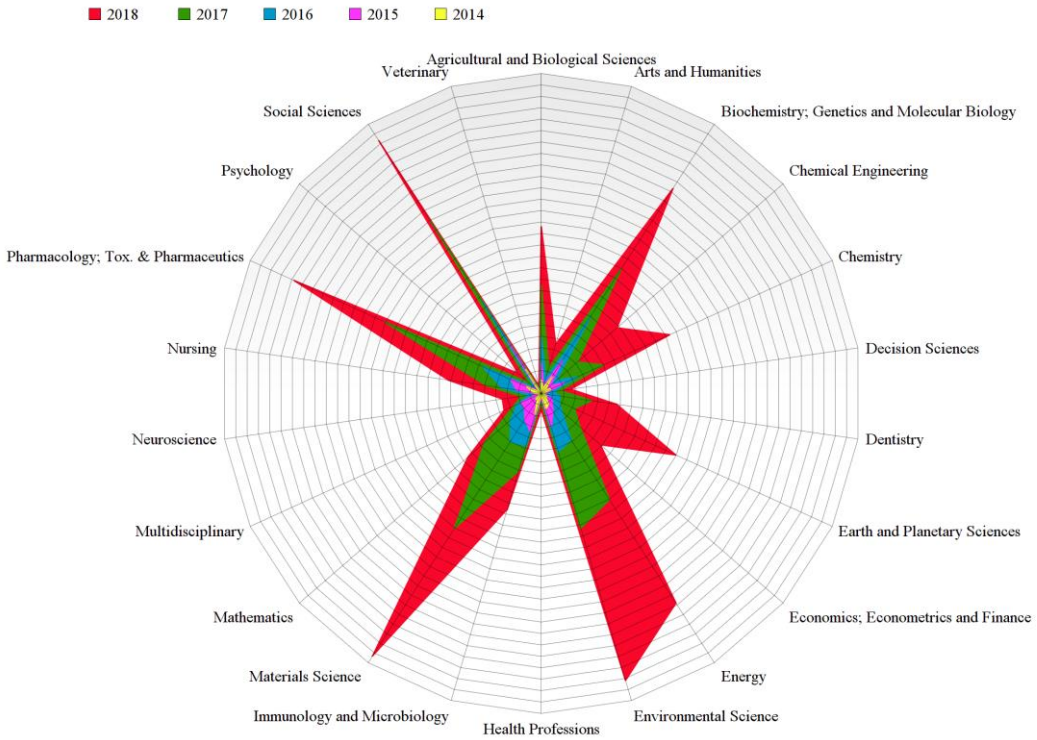

Figure 2. Research strength map of Universitas Indonesia for the last 5 years without the top 5.

Table 3. The details of research strength in the area of Computer Science.

No	Subject Category	Year					Total
		2014	2015	2016	2017	2018	
1	Computer Networks and Communications	330	253	928	2,062	2,333	5,905
2	Computer Science Applications	468	237	430	1,403	1,852	4,390
3	Computer Science (miscellaneous)	326	636	721	1,403	962	4,048
4	Signal Processing	218	81	192	910	1,075	2,477
5	Human-Computer Interaction	245	17	27	874	925	2,088
6	Software	208	168	292	455	897	2,020
7	Hardware and Architecture	22	21	356	480	695	1,574
8	Artificial Intelligence	129	101	242	392	623	1,487
9	Information Systems	60	92	290	309	650	1,401
10	Computer Graphics and Computer-Aided Design	73	64	122	368	205	832
11	Computer Vision and Pattern Recognition	60	47	102	314	282	804
12	Computational Theory and Mathematics	6	51	22	133	80	291

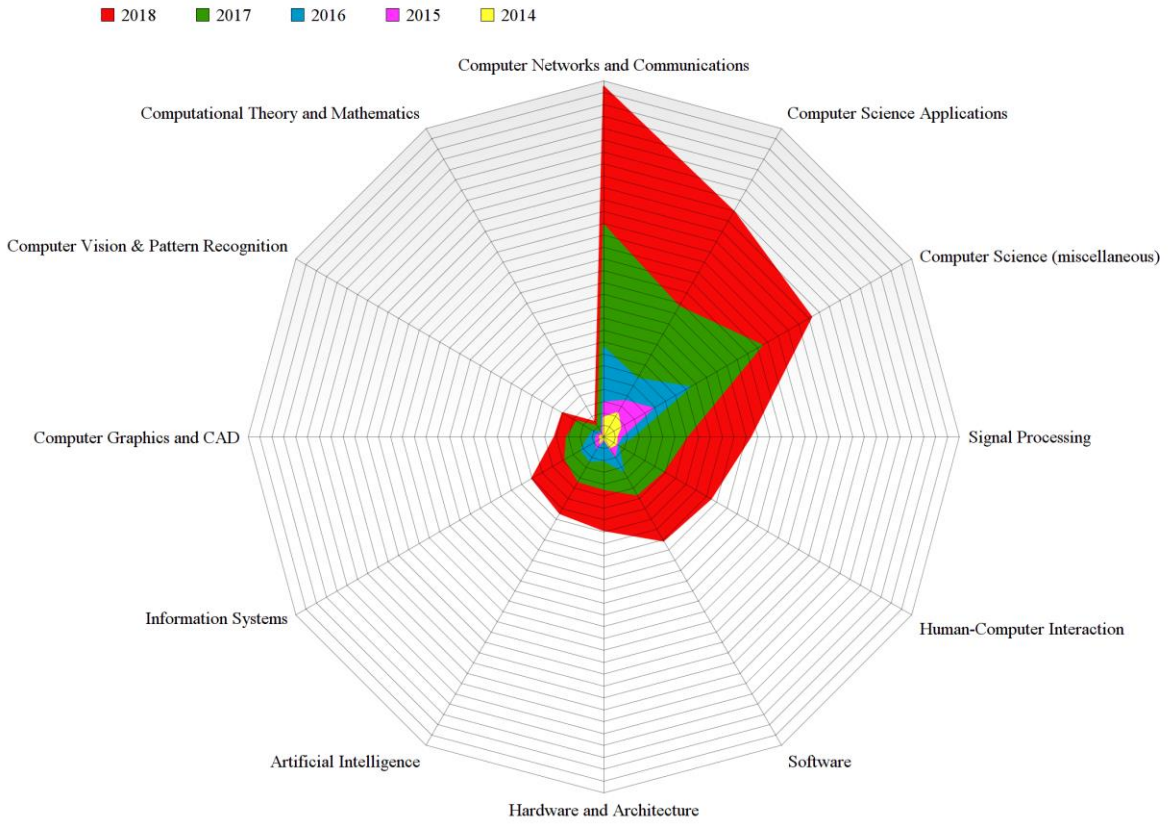

Figure 3. Research strength map of Universitas Indonesia for the last 5 years in the subject area of Computer Science.

5. Discussion

From the results above we can see that articles in subject areas “Medicine”, “Engineering”, “Computer Science”, “Physics and Astronomy”, and “Business, Management & Accounting” has dominated the research outputs of Universitas Indonesia. These subject areas were the strength research areas of Universitas Indonesia. This information is useful for an internal institution for its distinction or uniqueness, for SWOT (Strength, Weakness, Opportunity, and Threat) analysis, establishing research priorities and developing an institutional research plan. It is also valuable for research funder and government to allocate their grant to the right institution.

The research strength related to computer science was surprising because research output linked to Artificial Intelligence was weak beyond our expectations. Faculty of Computer Science of Universitas Indonesia (FCS-UI) could use Figure 3 as a foundation to manage their research plan and to change the priorities. The information can also be used by FCS-UI for SWOT analysis and for the foundation in changing its mission/objectives statements.

References

[1] M. Abbott and H. Doucoliagos, "Research output of Australian universities," *Education Economics,* vol. 12, no. 3, 2004.

[2] A. Martin-Sardesai and J. Guthrie, "Human capital loss in an academic performance measurement system," Journal of Intellectual Capital Forthcoming, 2017.

[3] Organization for Economic Co-operation and Development (OECD), "University Research Management," 2004.

[4] Universitas Indonesia, "Times Higher Education World University Ranking: UI of the Best Higher Education in Indonesia," Universitas Indonesia, 30 September 2019. [Online]. Available: https://www.ui.ac.id/en/times-higher-education-world-university-ranking-ui-of-the-best-higher-education-in-indonesia/. [Accessed 6 November 2019].

[5] Sinta Indonesia, "SINTA - Science and Technology Index," 2019. [Online]. Available: http://sinta2.ristekdikti.go.id/affiliations. [Accessed 5 August 2019].

[6] A. Uddin, J. Bhoosreddy, M. Tiwari and V. K. Singh, "A Sciento-text framework to characterize research strength of institutions at fine-grained thematic area level," Scientometrics, vol. 106, p. 1135–1150, 2016.

[7] B. González-Albo, L. Moreno, F. Morillo, and M. Bordons, "Bibliometric indicators for the analysis of the research performance of a multidisciplinary institution: The CSIC," Rev. Esp. Doc. Cient, vol. 35, no. 1, pp. 9-37, 2012.

[8] Agencia Estatal Consejo Superior de Investigaciones Científicas, "Consejo Superior de Investigaciones Científicas (CSIC) Annual Report 2017," 2018.

[9] J. J. a. L. YU, "Measuring the research performance of Chinese higher education institutions using data envelopment analysis," China Econ. Rev, vol. 19, no. 4, pp. 679-696, 2008.

[10] A.-W. Harzing and S. Alakangas, "Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison," Scientometrics, vol. 106, no. 2, pp. 787-804, 2016.

[11] P. Mongeon and A. Paul-Hus, "The journal coverage of Web of Science and Scopus: a comparative analysis," Scientometrics, vol. 106, no. 1, pp. 213-228, 2016.

[12] M. Rose and J. Kitchin, "pybliometrics: Scriptable bibliometrics using a Python interface to Scopus," SoftwareX, vol. 10, pp. -, 2019.

[13] G. G. Wang, J. W. Gilley and J. Y. Sun, "The “Science of HRD Research”: Reshaping HRD Research Through Scientometrics," Human Resource Development Review, vol. 11, no. 4, pp. 500-520, 2012.

[14] E. Garfield, "Citation indexes for science. A new dimension in documentation through association of ideas," Science, vol. 133, no. 3159, pp. 108-111, 1955.

[15] D. J. Price, "Networks of scientific papers," Science, vol. 149, pp. 510-515, 1965.

[16] E. Garfield, "Do Nobel Prize winner write citation classics?," Current Content, vol. 23, pp. 3-8, 1986.

[17] D. B. Klein and E. Chiang, "Citation counts and SSCI in personnel decisions: A survey of economics departments," Economics Journal Watch, vol. 1, pp. 166-174, 2004.

[18] A.-I. Petrişor and A. Mitrea, "Romanian spatial planning research facing the challenges of globalizing sciences," Urbanism. Architecture. Constructions, vol. 9, no. 1, pp. 81-88, 2018.

[19] R. Eskrootchi and N. Sanee, "Comparison of medical research performance by thermodynamic and citation analysis methods," Scientometrics, vol. 117, no. 3, pp. 2159-2168, 2018.

[20] L. Aldieri, G. Guida, M. Kotsemir and C. Vinci, "An investigation of impact of research collaboration on academic performance in Italy," Quality and Quantity, pp. -, 2019.

[21] P. Osborn, S. Ames, N. Turner, M. Caird, M. Karam, M. Mormino and C. Krueger, "An Analysis of Research Quality and Productivity at Six Academic Orthopaedic Residencies," Journal of Surgical Education, vol. 75, no. 6, pp. 1635-1642, 2018.

[22] C. Krueger, J. Hoffman, G. Balazs, A. Johnson, B. Potter and P. Belmont, "Protected Resident Research Time Does Not Increase the Quantity or Quality of Residency Program Research Publications: A Comparison of 3 Orthopedic Residencies," Journal of Surgical Education, vol. 74, no. 2, pp. 264-270, 2017.

[23] P. Haddawy, S.-U. Hassan, C. Abbey and I. Lee, "Uncovering fine-grained research excellence: The global research benchmarking system," Journal of Informetrics, vol. 11, no. 2, pp. 389-406, 2017.

[24] A. Noorhidawati, M. Aspura, M. Zahila and A. Abrizah, "Characteristics of Malaysian highly cited papers," Malaysian Journal of Library and Information Science, vol. 22, no. 2, pp. 85-99, 2017.

[25] M. Christopher and A. Marusic, "Geographic trends in research output and citations in veterinary medicine: Insight into global research capacity, species specialization, and interdisciplinary relationships," BMC Veterinary Research, vol. 9, pp. -, 2013.

[26] D. Ocholla, J. Mostert and D. Rotich, "Visibility of University of Zululand and Moi University researchers in web of science and scopus from 2003 to 2013," African Journal of Library Archives and Information Science, vol. 26, no. 1, pp. 3-15, 2016.

[27] M. Cacace, S. Ettelt, N. Mays and E. Nolte, "Assessing quality in cross-country comparisons of health systems and policies: Towards a set of generic quality criteria," Health Policy, vol. 112, no. 1-2, pp. 156-162, 2013.

[28] P. Lansley, "Research assessment and the activity hypothesis," Engineering, Construction and Architectural Management, vol. 20, no. 1, pp. 7-28, 2013.

18% Overall Similarity

Top sources found in the following databases:

- 18% Internet database
- Crossref Posted Content database
- 1% Publications database
- 0% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	journals.sagepub.com Internet	4%
2	researchgate.net Internet	3%
3	eprints.qut.edu.au Internet	2%
4	backend.orbit.dtu.dk Internet	1%
5	iopscience.iop.org Internet	1%
6	redc.revistas.csic.es Internet	1%
7	core.ac.uk Internet	<1%
8	munin.uit.no Internet	<1%
9	pure.york.ac.uk Internet	<1%

10	china.iopscience.iop.org Internet	<1%
11	aaltodoc2.org.aalto.fi Internet	<1%
12	icoen.org Internet	<1%
13	koreascience.or.kr Internet	<1%
14	alm.plos.org Internet	<1%
15	medlibrary.org Internet	<1%
16	repo.uni-hannover.de Internet	<1%
17	sdbindex.com Internet	<1%
18	senate.be Internet	<1%