Webology (ISSN: 1735-188X)
Volume 19, Number 3, 2022

Metric Comparison Between Google Scholar And Research Gate For Engineering Academics

Arthur James Swart

Central University of Technology.

Abstract

A researcher's reputation is influenced by the quantity and quality of his or her publications. A record of these publications is usually kept by an academic in a resume or on an online platform, such as LinkedIn, Google Scholar or Research Gate. The purpose of this article is to contrast the major metrics of Google Scholar with that of Research Gate of a number of engineering academics employed at a university of technology in South Africa, in order to determine any notable differences. A quantitative study is undertaken to gather the total number of citations, h-index values and Research Gate scores for engineering researchers present on both databases. Results indicate that Research Gate has the highest number of authors present from the Faculty of Engineering, Built Environment and Information Technology at the Central University of Technology. However, Google Scholar records the highest number of citations for these authors. Only 31 out of the possible 86 academic staff members maintains a profile on both databases, where only 4 researchers have more than 500 citations. It is recommended that management mandate the presence of their academic staff on one of many available databases in this regard, thereby enhancing the visibility of the research done at the university and enabling an easier review of the achievements of staff for performance management purposes or for promotion.

Keywords: citations, h-index, score, quantitative, data mining, data analytics

1. Introduction

According to Harvey Mackay, “You can't buy a good reputation; you must earn it” [1]. Mackay is one of America’s most popular and entertaining speakers and author of a number of best-selling inspirational business books. Mackay rightly states that a good reputation must be earned, which obviously requires much hard work and effort. This is well-known to many researchers around the globe who often struggle to get their own research work published. However, a researcher’s reputation for ongoing scholarly work in a specific field is significantly enhanced through such publications.

A researcher's reputation is influenced by the quantity and quality of his or her publications [2]. These publications are often considered for selection and promotion in academic institutions. In fact, as is usually the case in academia, a researcher's publications are considered more important by an institution than is the actual teaching of the researcher [3]. A record of these publications is usually kept by an academic in a resume or on an online platform, such as LinkedIn, Google Scholar or Research Gate. Both Google Scholar and Research Gate provide specific metrics that are related to a researcher's publications, such as the total number of citations, the h-Index, and the Research Gate Score (RGS).

However, one recent comparison of such platforms cautions that Web of Science and Scopus impact metrics should be humorously taken with a cautionary cup of salt, whereas metrics from Google Scholar and Research Gate should be taken with a bountiful classroom of salt [4]. These metrics should therefore not be considered as the sole indicator, or pinnacle, of one's reputation, but should rather be considered as an initial stepping stone to establish the reputation. It can indeed be useful for self-monitoring and to indicate a measure of success [5]. Although both databases may provide similar publication ranks [6], great differences have been found between the number of citations present on Google Scholar and those present on Research Gate [7]. The following research questions arise:

- What difference exists in the average number of citations between Google Scholar and Research Gate for a specific list of researchers?
- What relationship exists between the main metrics of Google Scholar and Research Gate for a specific list of researchers?
- What percentage of researchers listed on Google Scholar do not engage in sole authorship?

The purpose of this article is therefore to contrast the major metrics of Google Scholar with that of Research Gate of a number of engineering academics employed at a university of technology in South Africa (SA), in order to determine any notable differences. This is a form of data mining or data analytics, although not linked to big data. Data mining refers to all aspects of an automated or semi-automated process to extract unknown and useful knowledge and patterns from large databases [8] while data analytics refers to a method or technique that uses data, information, and knowledge to learn, describe and predict something [9]. A quantitative study is used to gather the total number of citations, h-index values and RGS for academic researchers present on both databases. The article starts with a brief discussion on the purpose and value of the two databases, followed by the context of this study. Results, discussions and the conclusions follow.

2. Google Scholar versus Research Gate

Google Scholar is a searchable database which focuses on scholarly literature, such as academic journals. It allows users to access information, cross reference that information with other sources and keep up with new research as it is released. It also allows users to access journals, conference papers, academic books, pre-prints, theses and dissertations, abstracts and other scholarly literature from other areas of research. Google Scholar started to operate on the 20th of November 2004 [10].

A study investigated the use of Google Scholar in searches for academic and grey literature. Grey literature is information produced outside of traditional publishing and distribution channels. Common types of grey literature include reports, working papers, newsletters, government documents, speeches, white papers and urban plans. Grey literature can be seen as good, as it is often more current than traditionally published sources, due to it not having to go through a lengthy peer-review process. Grey literature can also provide access to raw data and data sets which would normally be inaccessible. Grey literature is furthermore not tied to a conventional structure, and as such can be longer and provide more detail. The results showed that the majority of grey literature occurred on page 80 of the Google Scholar results [11].

A study in 2017 tested whether or not the identification of highly-cited documents through Google Scholar is feasible and reliable. This study was done to see whether or not Google Scholar could be used as a viable tool for bibliometric analysis. The results showed that there was a strong correlation between a document's citations and where it appeared in the search results. This study thus concluded that Google Scholar was able to identify highly-cited documents effectively. During this study, 64,000 documents were obtained from Google Scholar [12].

A recent study investigated the differences between Google Scholar, Web of Science and Scopus, as there was no recent or systematic evidence of the differences between them. This study found that Google Scholar consistently found the largest percentage of citations across all areas. Google Scholar found most of the Web of Science and Scopus citations. This study investigated 2,448,055 citations to 2,299 highly-cited documents from 252 subject categories published in 2006 [13].

Research Gate is a social network for researchers. It allows users to share their research publications, find collaborators, access job boards and ask and answer questions in real-time. The major disciplines on Research Gate include: biology, medicine, computer science, physics and chemistry. Research Gate started to operate in the month of May of 2008 [14].

A study on Research Gate investigated whether or not the Research Gate Score (RGS) for a user could be used as evidence of academic scholarly reputation. This was

done as the RGS is often used for the evaluation of a researcher during recruitment and promotion processes. The results showed that a higher RGS primarily came from asking and answering questions on the site. This study also concluded that it is impossible to get a higher score by only listing publications. Thus, the study concluded that a RGS should not be primarily used as an indication of academic scholarly reputation [15].

Another study in 2017 investigated whether the number of citations for newer articles on Research Gate is comparable to other citation indexes. This was done as researchers could upload preprints to Research Gate to provide early impact evidence for new papers. This study used 2,675 recently-published library and information science articles. The results showed that Research Gate found less citations than Google Scholar, but found more than Web of Science and Scopus [6].

Finally, a study investigated in what ways academics use Research Gate and what they think of such websites. A largely American and European sample was used in the study that examined motives, use and career-related outcomes. This study found that most academics who have a Research Gate account do not use it very much, as they could not derive many benefits from it and could not relate it to their career-related outcomes [16].

3. Study context

The Central University of Technology (CUT) is located in the Central Region of SA (Central University of Technology, 2017), being one of 6 universities of technology in the country [17]. It was originally designated as the Free State Technikon in 1981 with the purpose of offering diplomas and certificate courses that were geared towards the needs of industry [18]. It was re-designated as a university of technology in 2004, being called CUT. It now offers a wide range of qualifications according to the National Qualifications Framework (NQF). The lowest NQF qualification in the Faculty of Engineering, Built Environment and Information Technology (FEBIT) is a Higher Certificate in Renewable Energy (NQF level 6) with the highest qualification being a Doctoral in Engineering (NQF level 10). This type of university has a strong vocation-driven teaching mandate with a strong applicable research focus [19]. The number of faculties and students present at CUT in 2018 is shown in Table 1.

The overview of the four faculties at CUT indicates that the Faculty of Health and Environmental Sciences (FHES) has the least number of academic staff (55) while the Faculty of Engineering, Built Environment and Information Technology (FEBIT) has the largest number (86). The FEBIT also has the largest number of undergraduate students and focusses primarily of the disciplines of science, technology, engineering and mathematics (STEM). For this reason, this article focusses only on academic staff in this faculty.

CUT identified three main research focus areas, or clusters, in 2018. Most of the research and publications of academic staff at CUT over the past decade has been related to these clusters, which include:

1. Technologies and Innovations for Sustainable Development: To investigate and apply technologies and/or innovation to foster and promote sustainable development.
2. Quality of Health and Living: To apply scientific research in different disciplines to improve on the quality of health and living standard of humans, animals and plants.
3. Socio-economic and Entrepreneurship Development: To do scientific research that empowers society for an invaluable contribution to sustainable socio-economic development.

Table 1. Academic staff and students at CUT in 2018 [20].

Faculty Name	Number of departments	Number of academic staff	Number of under-graduate students	Staff to student ratio	Number of post-graduate students
Engineering, Built Environment and Information Technology (FEBIT)	6	86	4465	51.9	216
Humanities (FH)	5	74	3645	49.2	588
Health and Environmental Sciences (FHES)	4	55	1642	29.8	38
Management Sciences (FMS)	6	80	4419	55.2	83

4. Quantitative research methodology

To generate the graphs, as seen in the results section, data first had to be collected from Google Scholar and Research Gate. This was done in the month of April 2019. Current data on the databases would not correlate to the results presented in this article, as this

type of data is not static, but rather dynamic, as new citations to existing publications are constantly recorded. MS Excel spreadsheets were used to manually capture the data. Data such as the total amount of citations per author and the highest citation number for a sole author publication was collected. The author names were originally collected from a group email that is regularly circulated among all staff members in the FEBIT at CUT. The search feature available on Google Scholar and Research Gate was used to find the required data for each staff member.

Before the graphs could be generated, the data first had to be processed in MS Excel. This involved manually checking each Google Scholar author to see if they had a Research Gate profile. This was done as there were less authors on Google Scholar than on Research Gate. Authors present on both databases were considered for this research. The next step was to sort the authors alphabetically and then represent them with numerical numbers to ensure anonymity and ethical compliance. The number of authors present on both databases were then contrasted with regard to their citations. The h-index and RGS of the total number of authors present on both databases were then contrasted. Finally, the number of authors present on Google Scholar were contrasted with regard to their highest citation count for a sole publication.

6. Results and discussions

Figure 1 shows a comparison of citations from Google Scholar and from Research Gate for each of the 31 authors that are present on both databases. Google Scholar is marked with black and Research Gate with grey. Author 31 has the highest number of Google Scholar citations (over 1300) and author 13 has the highest number of Research Gate citations (over 900). The average number of citations for Google Scholar is 198.16, while the average number of citations for Research Gate is 137.06 (this equates to an average difference of 61). This confirms previous research that large differences do exist between the number of citations present on both databases [7].

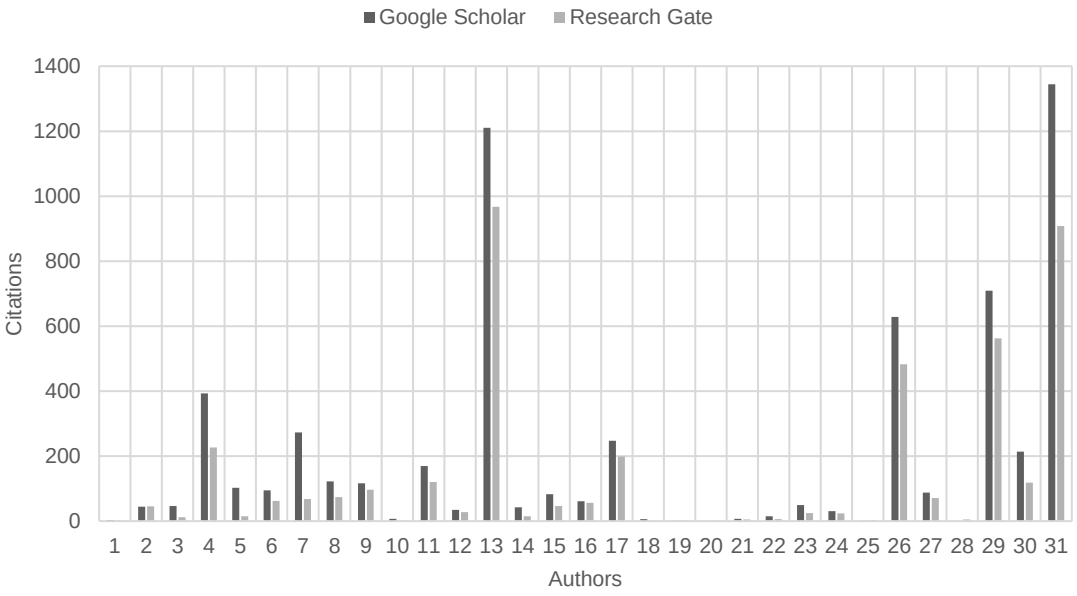

Figure 1. Google Scholar versus Research Gate in terms of total citations

Figure 2 shows the RGS and Google Scholar h-index values for each author. Again, 31 authors are listed in no specific order where author 13 has the highest RGS (being 27.59) with the highest Google Scholar h-index (being 21). The average RGS is 8.00 while the average h-index is 5.39. A Pearson correlation between the h-index and RGS was calculated to be r = 0.925 that indicates a statistical significant relationship between these two variables. This suggests that as the one variable increases, the other one will do likewise, suggesting a measure of consistency in the way these two variables are calculated on the two databases.

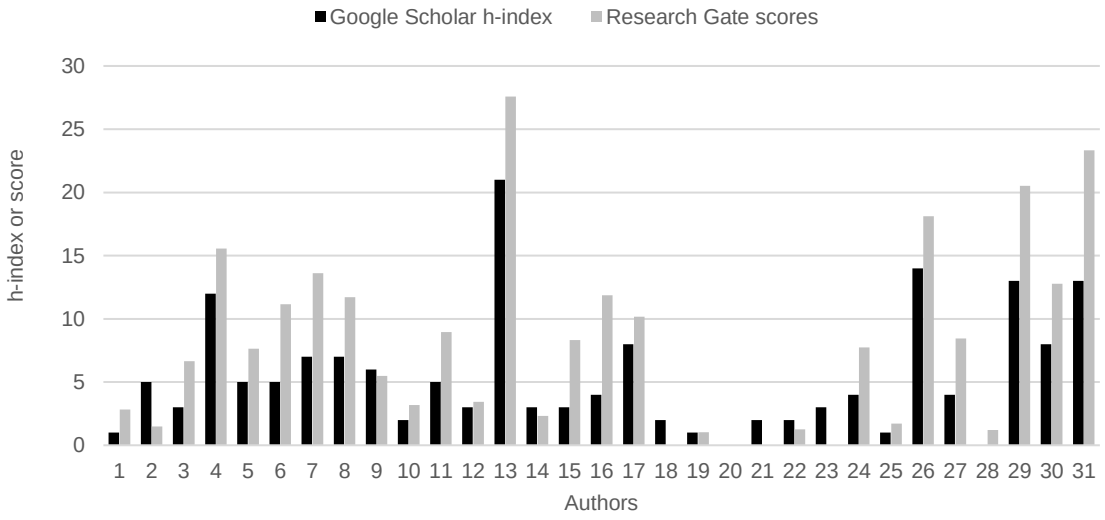

Figure 2. Google Scholar h-index versus Research Gate scores

Figure 3 shows the highest number of citations for a sole author publication (only one author listed on the publication) as listed on Google Scholar. Again, 31 authors are listed in no specific order where author 26 has the highest number of citations for a sole publication (being 74). The average number of citations for a sole publication is 8.68. Of the 31 authors, 11 have no sole authorship, that equates to a percentage of 35%.

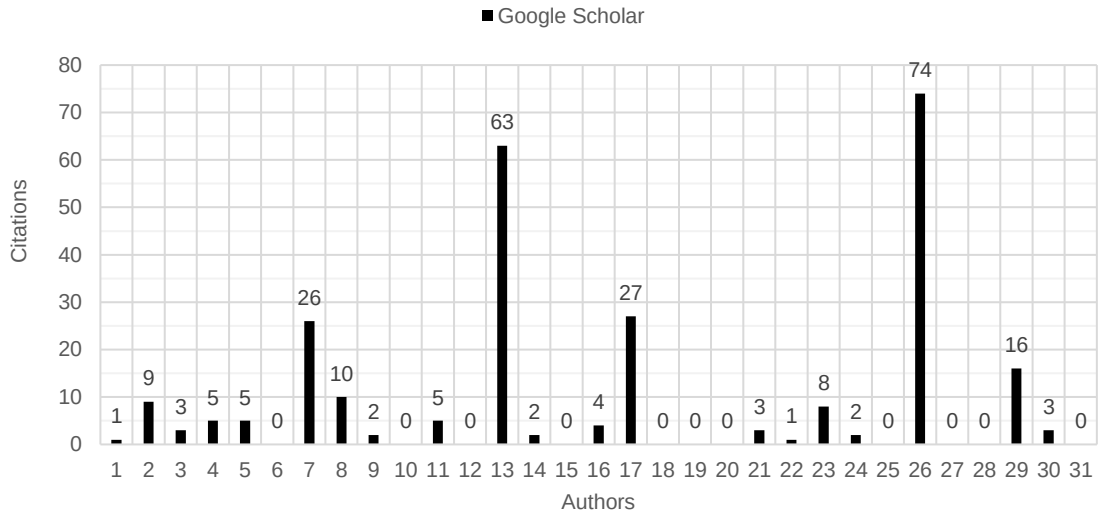

Figure 3. Google Scholar citations for a sole-author publication

A comparison between Figure 1 and Figure 3 suggests that authors 13 and 26 seem to be balancing their research between multi-author (collaborating with fellow researchers) and sole author (working on their own) publications. This is due to the fact that they are achieving a high number of citations for both types of publications. This type of collaboration is also seen in postgraduate studies [20], where a single student may have one or more supervisors with at least 2 authors (multi-author publication) listed on each publication.

It is further noteworthy that author 31 has no sole-author citation count (see Figure 3). This suggests that this author is only producing multi-authored publications (engaging only in collaborative work), as the total citation count from Google Scholar is the highest in Figure 1. It must be noted that for a researcher to be granted the status of a rated scientist in SA by the National Research Foundation (NRF) requires that the researcher have a number of sole-author publications, as it indicates the ability of the author to publish research that they have done on their own. To be recognized as a rated scientist in SA is very important, as it results in additional research funding for the researcher. It also enhances the research reputation of the researcher, making him or

her more marketable. Furthermore, some universities offer monetary incentives to their academic staff who successfully achieve such a rating.

Being a sole author does represent a significant responsibility, as it means that one has full autonomy and accountability for producing papers worthy of publication. However, it also enables one to avoid engaging in the somewhat contested space of author order negotiation [21]. It also removes the debate of what contribution each author made to the publication as a percentage and negates the dividing of research credits in countries where the government incentivizes research publications. For example, in SA, the government awards around $8 000 to a university for each research credit produced (equates to one Scopus based journal article where all the authors are affiliated to the university). If the authors are affiliated to different universities, then the research credit, and thus research funding, needs to be equally apportioned.

Table 2 shows the comparison of the total number of authors present on both databases, along with the total number of citations for all these authors. Research Gate has the highest number of researchers (over 70) from CUT while Google Scholar has the highest number of total citations (over 10 000). This supports previous research that found that Research Gate has less citations than Google Scholar [6].

Table 2. Total number of authors and citations present on both databases in 2019

	Total Number of Authors	Total Number of Citations
Google Scholar	31	10048
Research Gate	73	4227

7. Conclusions

The purpose of this article was to contrast the major metrics of Google Scholar with that of Research Gate of a number of engineering academics employed at a university of technology in SA, in order to determine any notable differences. Although 86 academic staff members were present in the FEBIT at CUT during 2018, only 73 had a research profile on Research Gate, with even a smaller number (31) being present on Google Scholar. This last value equates to 36% of academic staff in the faculty who could also feel that they are not deriving many benefits from using Research Gate and could not relate it to their career-related outcomes, as noted in the literature review of this article. Only 4 academic staff from the FEBIT at CUT have more than 500 citations on both databases, thereby suggesting that only a small number of engineering researchers are really producing publications that merit a citation by other researchers around the globe. This equates to a percentage of only 4.6% of academic staff in the faculty. It must be noted that Google Scholar is able to identify highly-cited documents

effectively, as stated in the literature review of this article, which adds credence to this metric. The following research questions are answered.

- What difference exists in the average number of citations between Google Scholar and Research Gate for a specific list of researchers?

The difference in the average number of citations between the two databases equals 61, which lends support to published literature that large differences do exist between Google Scholar and Research Gate in terms of citation numbers

- What relationship exists between the main metrics of Google Scholar and Research Gate for a specific list of researchers?

A statistical significant relationship exists between RGS and the Google Scholar h-index values ($r = 0.925$). This suggests that RGS are influenced by the number of publications and citations, although previous research indicates that a higher RGS primarily comes from asking and answering questions on the site. Furthermore, a measure of consistency in the way these two variables are calculated on the two databases is suggested by this relationship.

- What percentage of researchers listed on Google Scholar do not engage in sole authorship?

35% of the researchers listed on the Google Scholar database have not yet produced a sole publication. This may impact on their application as a rated scientist or on their promotion as they need to substantiate their contribution to the publication which is divided among the number of authors listed.

A limitation of this study includes the small sample size that is limited to only one faculty at a university of technology in SA. However, the results of this study can serve as a baseline from which to evaluate other universities and from which to evaluate the progress of these specific engineering academics over time, by considering another snapshot of these databases in a few years-time. Another limitation relates to the fact that not all publications are automatically listed on Google Scholar or Research Gate, but requires a manual upload by the researcher. A researcher may upload a publication to one database, and not to the other, which may also impact on the metrics.

The metrics on these two databases should not be considered as the sole indicator, or pinnacle, of one's reputation, but should rather be considered as an initial stepping stone to establish the reputation. It is recommended that researchers try to find a balance between multi-authored and sole authored publications, as this is important to academics who wish to become rated scientists in their respective countries. Furthermore, it is recommended that management at universities mandate the presence of their academic staff on one of the many available databases in this regard, thereby

enhancing the visibility of the research done at the university and enabling an easier review of the achievements of staff for personal performance management purposes and for promotion.

8. References

[1] Brainy Quote, "Homepage", available online: http://www.brainyquote.com/quotes/, last visit: 10 March 2020.(**2020**).

[2] Giuliani, E, Morrison, A, Pietrobelli, C, and Rabellotti, R, "Who are the researchers that are collaborating with industry? An analysis of the wine sectors in Chile, South Africa and Italy," Research Policy, 39, 748-761. (**2010**)

[3] Pollard, V, Hains-Wesson, R, and Young, K, "Creative teaching in STEM," Teaching in Higher Education, 23, 178-193. (**2018**)

[4] Cochran, A, "How many grains of salt must we take when looking at metrics?", available online: https://scholarlykitchen.sspnet.org/2017/02/08/49186/, last visit: 17 September 2019.(**2017**).

[5] Ma, L and Ladisch, M, "Scholarly communication and practices in the world of metrics: an exploratory study," in Proceedings of the 79th ASIS&T Annual Meeting: Creating Knowledge, Enhancing Lives through Information & Technology, 2016, p.^pp. 132. (**2016**).

[6] Thelwall, M and Kousha, K, "ResearchGate versus Google Scholar: Which finds more early citations?," Scientometrics, 112, 1125-1131. (**2017**)

[7] Orduna-Malea, E and Delgado López-Cózar, E, "Performance behavior patterns in author-level metrics: a disciplinary comparison of Google Scholar Citations, ResearchGate, and ImpactStory," Frontiers in Research Metrics and Analytics, 2, 14. (**2017**)

[8] Hasheminejad, SH-A, Zahedi, M, and Hasheminejad, SMH, "A hybrid clustering and classification approach for predicting crash injury severity on rural roads," International journal of injury control and safety promotion, 25, 85-101. (**2018**)

[9] Elhadad, E, "Insurance Business Enterprises' Intelligence in View of Big Data Analytics," (**2021**. (21-13, 26, مجلة الجمعية المصرية لنظم المعلومات وتكنولوجيا الحاسبات

[10] Delfino, D, "'What is Google Scholar?': What you need to know about Google's database for students, researchers, and other curious minds", available online: https://www.businessinsider.com/what-is-google-scholar?IR=T, last visit: 19 September 2019.(**2019**).

[11] Haddaway, NR, Collins, AM, Coughlin, D, and Kirk, S, "The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching," PLOS ONE, 10, e0138237. (**2015**)

[12] Martin-Martin, A, Orduna-Malea, E, Harzing, A-W, and Delgado López-Cózar, E, "Can we use Google Scholar to identify highly-cited documents?," Journal of Informetrics, 11, 152-163. (**2017**)

[13] Martín-Martín, A, Orduna-Malea, E, Thelwall, M, and Delgado López-Cózar, E, "Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories," Journal of Informetrics, 12, 1160-1177. (**2018**)

[14] University of Melbourne, "Researcher Profiles, Identifiers and Social Networks: Maximise your Impact", available online: https://unimelb.libguides.com/researcher_profiles/rg, last visit: 19 September 2019.(**2019**).

[15] Orduna-Malea, E, Martín-Martín, A, Thelwall, M, and López-Cózar, ED, "Do ResearchGate Scores create ghost academic reputations?," Scientometrics, 112, 443-460. (**2017**)

[16] Muscanell, N and Utz, S, "Social networking for scientists: an analysis on how and why academics use ResearchGate," Online information review, 41, 744-759. (**2017**)

[17] Swart, AJ, "Student usage of a learning management system at an open distance learning institute – A case study in electrical engineering," IJEEE, International Journal of Electrical Engineering Education, 52, 142-154. (**2015**)

[18] Sehoole, C and Ojo, EO, "Challenges and opportunities for new faculty in South African higher education," Young faculty in the twenty-first century: International perspectives, 253-283. (**2015**)

[19] Pitso, T, "Status of the scholarship of teaching and learning in South African universities," South African Journal of Higher Education, 27, 196-208. (**2013**)

[20] Swart, AJ, "An Analysis of Master Dissertations: A Case Study of Central University of Technology, South Africa," African Journal of Library, Archives and Information Sciences, 28, 211-223. (**2018**)

[21] Merga, MK, "Thesis by publication in education: An autoethnographic perspective for educational researchers," Issues in Educational Research, 25, 291-308. (**2015**)

Authors

James Swart received his DTech: Electrical: Engineering degree in 2011 from the Vaal University of Technology. His research interests include engineering education development and energy monitoring of PV modules. He is currently an Associate Professor at the Central University of Technology.