Mapping the computer science research landscape in sub-Saharan Africa: Bibliometric and altmetric analyses

Matthew	Harsh	matthew.harsh@concordia.ca	Canada	Concordia University
Alexandra	Meikleham		Canada	Concordia University, Montreal, QC, Canada
Lisa Maria	Negro		Canada	Concordia University, Montreal, QC, Canada
Charles C.	Onu		Canada	Concordia University, Montreal, QC, Canada
Ravtosh	Bal	ravtosh.bal@gmail.com	Canada	Concordia University, Montreal. QC, Canada

Introduction

Standard metrics that focus on journal publications, citations, and patents do not present a complete picture of research productivity in Africa. Publication counts and citations use databases like the Web of Science or Scopus are biased towards publications in the English language, placing research conducted in other languages at a disadvantage (Archambault et al.2006). A large amount of research in Africa is published in gray literature (working papers, reports, publication of non-governmental organizations, etc.) or only published as MSc or PhD thesis (Harsh & Zachary 2013). In addition, much of African research is published in local journals that remain invisible to the global scientific community (Tijssen et al. 2006). The invisibility of this research raises important concerns about the criteria and methodology for measuring scientific productivity and research capacity, evaluating research performance and understanding the impact of research in Africa.

The aim of our research is to reduce this invisibility and map the contours of computer science research activity in sub-Saharan Africa. Computer science is central to the development of information technology, a field that has already transformed commerce and has further potential to boost incomes and quality of life. Our study addresses the following research questions: (i) What is the research productivity of the sub-Saharan countries in the field of computer science? (ii) What are the patterns of research collaboration for African computer science? (iii) How do altmetrics compare with traditional bibliometrics in mapping the research landscape of computer science in sub-Saharan Africa? (iv) How can we measure the wider, societal impact of computer science research in sub-Saharan Africa?

Data and methods

Our study utilizes a mixed methodology approach based on bibliometric and altmetric data as well as survey data. Our analysis is based on a database of sub-Saharan computer science researchers constructed by our team. A list of 29 sub-Saharan African countries and 78 universities which offer computer science related programs was created by manual web searches. Each university website was manually reviewed for relevant faculty and researcher information including name, position, and contact information. This list of 317 contacts formed the basis of our author database for computer science related researchers in Africa. After successive iterations of analysis on Academia.edu., the list grew to over 1100 researchers. In the first phase of our research, we have focused only on the Anglophone countries of sub-Saharan Africa excluding South Africa which is an outlier in the region in terms of research output and productivity.

To map African computer science research productivity and visualize collaboration patterns, we utilize both bibliometric and altmetric data derived from specialized databases such as ACM, IEEE Xplore, and INSPEC and from academic networking sites such as Academia.edu. Journal articles from the region for the period from 1978 – 2015 were downloaded from these specialized databases. We also used Academia.edu, a social networking site, to download research articles, papers and other types of research outputs to construct a database that is not restricted solely to journal articles. We use co-authorship as the measure for scientific research collaboration (Katz & Martin 1997). Analysis of authors' institutional affiliation is used to determine international collaboration. VOSviewer and Gephi are used to visualize collaboration patterns. Lastly, we analyzed responses to a survey that was administered via the internet to computer science researchers in the region to understand the wider impacts of research and provide additional insights on research productivity and collaboration.

Findings and conclusions

Significantly more computer science research is happening in this region than indicated by bibliometric analyses based on international databases. The dominant type of research outputs are journal articles and conference papers. The total scientific output of computer science articles for the period from 1978 - 2015 was 4220 (INSPEC), 2061 (IEEE Xplore), and 311 (ACM). Research productivity in computer science is very unevenly distributed across the countries. For instance, based on the IEEE database, the most productive countries in computer science are Nigeria, Botswana, Cameroon, Kenya, and Mauritius producing 59% of the region's total publications.

Of a total of 3108 papers (2061 from IEEE and 1047 from Academia.edu), the total out of country collaborations were found to be 319. Of these collaborations, almost ten times more were found in papers on Academia.edu (285 collaborations) compared to IEEE (35 collaborations). Collaborations were classified as "internal-internal" (a sub-Saharan African country collaborated with a different country in the region) or as "internal-external" (a sub-Saharan African country collaborated with a country external to the region). Regional collaborations are rarer, corroborating findings from previous bibliometric analyses. Only 6 collaborations in IEEE were internal-internal (17% of IEEE collaborations), while 26 collaborations found in Academia.edu were internal-internal (9% of Academia.edu collaborations). 12 discrete country-country relationships account for 50% of internal-external collaborations found in Academia.edu.

Our findings indicate that international databases, including specialized computer science databases, do not present a complete picture of sub-Saharan African research productivity in the field of computer science. Academic networking sites such as Academia.edu can help to fill some of the gaps. But the use of these sites as a source of data is also beset with methodological challenges that require intensive manual review to overcome, such as: researchers having multiple profiles, the lack of verification of users, researchers selectively uploading papers, and the duplication of uploaded papers. Our analysis of societal impact was based on survey questions about researchers' perceptions of their influence and research impact. While our analysis was exploratory, it points to the potential that researchers' perceptions of their work can provide for a more context-relevant understanding of the pathways to societal impact.

References

Archambault, E., Vignola-gagne, E., Cote G., Lariviere, E. and Gingras, Y. (2006). Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. *Scientometrics, 68*(3), 329–342.

Harsh, M. & Zachary, G. P., (2013) Computer science research capacity as a driver of ICTD innovation: institutional factors in Kenya and Uganda. *Internet and Communication Technology for Development 2013 Proceedings*, ACM 978-1-4503-1907-2/13/12.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy, 26*, 1–26.

Tijssen, R., Mouton, J., Van Leeuwen, T. & Boshoff, N. (2006). How relevant are local scholarly journals in global science? A case study of South Africa. *Research Evaluation, 15*(3), 163-174.