Towards Open and Reproducible Genomic Research: Lessons from OpenScienceKE

Caleb Kibet [1,3,*], Author 2[2], Author 3[1], Author 4[1], Author 5[2], Author 6[2], OpenScienceKE[3]

1 International Center of Insect Physiology and Ecology
2 Affiliation B
3 Members of the OpenScienceKE community.

*** ckibet@icipe.org**

Abstract

Open, collaborative and reproducible research – Open Science – has a great potential for advancing science. However, the training in our local universities does not equip students with the tools to practice open science. However, to work in the open and collaborate, your collaborators should be equipped to use the tools that you use. The main barrier to working open, therefore, is the lack of awareness of the collaboration tools and the skills required to utilize these tools. Therefore, to fill the gap through an open science community, funded by a Mozilla Mini-grant – OpenScienceKE, we are promoting open science among bioinformatics students and researchers in the Nairobi area by training using this model: sensitize, train, hack and collaborate. This model first sensitizes on open science practices through seminars, trains on open science tools through workshops, facilitates hands-on application of the tools through hackathons, and finally fosters a community of open science enthusiasts through meetups.

OpenScienceKE sought to address the following problems: the lack of awareness of open science practices and tools within the Bioinformatics community in Kenya; the poor adoption of open science practices in Bioinformatics; and the absence of research to establish the state of affairs in adopting open science in Kenya. From the OpenScienceKE hackathon, we managed to create an open resource that the students could use to figure out where they can cost-effectively publish open access. In addition, through literature search and data mining, we observed a growing interest in open science practices in Kenya but the lack of awareness and skills hinder the adoption. The use of preprints for research dissemination haven't caught up in Kenya; out of the 20,069 papers downloaded from bioRXiv, only 18 have Kenyan authors, a majority of which are as a result of international (16) collaborations. We also observed a lack of incentives and policy in academic and research institutions to support open science. The fear of being scooped and the competitive spirit within the scientific community are also major barriers to working in the open.

The first iteration of the model which focused on academic institutions set the foundation for next phase: promote the open and reproducible science in research institutions. This model provides the framework for the adoption of open science practices within the institution and others in the future. As genomic research data generated in Africa grows, there is a need for the adoption of open science practices in data storage, reproducible pipelines and collaborative research. We propose this approach, which develops the necessary infrastructure within research institutions, and

builds human capacity through the model: sensitize, train, hack and collaborate. Promotion of open science in Africa recognizes the future direction of research and OpenScienceKE is growing the culture and practice in the research active region.

Introduction

Our Core message is to present a clear outlook on the adoption of Open Science practice by African-based scientists, starting with Kenya.

Add an introduction of the Open Science. To cite an article, use [2]. All the bibliographies should be added to `library.bib` in the BibTeX format. See the example in `library.bib`.

Since this will more or less be like a review article, we will need to identify the various subsections based on the topics we need to cover in the review. See this article for some tips [1].

Placeholder Text for the Background

Open, collaborative and reproducible research – Open Science – has a great potential for advancing science. However, the training in our local universities does not equip students with the tools to practice open science. However, to work in the open and collaborate, your collaborators should be equipped to use the tools that you use. The main barrier to working open, therefore, is the lack of awareness of the collaboration tools and the skills required to utilize these tools. Therefore, to fill the gap through an open science community, funded by a Mozilla Mini-grant – OpenScienceKE, we are promoting open science among bioinformatics students and researchers in the Nairobi area by training using this model: sensitize, train, hack and collaborate. This model first sensitizes on open science practices through seminars, trains on open science tools through workshops, facilitates hands-on application of the tools through hackathons, and finally fosters a community of open science enthusiasts through meetups.

OpenScienceKE sought to address the following problems: the lack of awareness of open science practices and tools within the Bioinformatics community in Kenya; the poor adoption of open science practices in Bioinformatics; and the absence of research to establish the state of affairs in adopting open science in Kenya. From the OpenScienceKE hackathon, we managed to create an open resource that the students could use to figure out where they can cost-effectively publish open access. In addition, through literature search and data mining, we observed a growing interest in open science practices in Kenya but the lack of awareness and skills hinder the adoption. The use of preprints for research dissemination haven't caught up in Kenya; out of the 20,069 papers downloaded from bioRXiv, only 18 have Kenyan authors, a majority of which are as a result of international (16) collaborations. We also observed a lack of incentives and policy in academic and research institutions to support open science. The fear of being scooped and the competitive spirit within the scientific community are also major barriers to working in the open.

The first iteration of the model which focused on academic institutions set the foundation for next phase: promote the open and reproducible science in research institutions. This model provides the framework for the adoption of open science practices within the institution and others in the future. As genomic research data generated in Africa grows, there is a need for the adoption of open science practices in data storage, reproducible pipelines and collaborative research. We propose this approach, which develops the necessary infrastructure within research institutions, and builds human capacity through the model: sensitize, train, hack and collaborate. Promotion of open science in Africa recognizes the future direction of research and OpenScienceKE is growing the culture and practice in the research acti

Status of Open Science in Kenya: Literature search

The report from the above team will be useful for writing the introduction, as well as providing materials to be used in the discussion. In fact, we can use their resource to weave the whole paper together.

In this section, and the introduction, we will conduct a review of the status of open science in the country. - What kind of resources are available to support open science - Are there policies or incentives for open science practices? - What kind of training activities have been conducted to promote and train students and researchers on open science tools and practices? - etc

DataMining Section

This will be a data analysis section. The title of this section will depend on the results of your analysis. For example, if Kenyan researchers are not publishing open access, we will need to understand why that is the case. The solution may lie in the cost of publishing, and that is how the resource created by Open Access options team is useful.

We address questions like: - What is the publishing trend by Kenyan researchers - Are they publishing open access, and how has this changed over the years? - Are Kenyan researchers embracing pre-prints (BioRXiv, AriRXiv, ResearchGate, F1000Research). Who is driving the adoption of pre-prints? Local researchers or foreign collaborators? - What are the collaboration trends? Are Kenyan researcher collaborating locally or internationally?

But Publishing Open Access is Expensive

The article processing charge is the main barrier to publishing open access, in addition to the obsession with impact factors. However, for early career scientists and students, especially in developing countries, most publishers offer waivers and subsidies but few are aware. In this section, we explore some of the avenues to publishing open access at low cost.

To address this problem, we created a resource that can guide ECR and students on where they publish open access, and at low cost. We also provide information on how they can still be open when they publsih in paywalled journals, eg via the green route.

Figures

You can add the figures as follows:

Figure 1. Figure 1

And you can have it referenced as a figure

> **Box 1** To highlight of defining some key concepts in Open science without disrupting the flow of the articles, you can use a quote format.

Discussion

What do the results mean? How does your results fit to the current literature? How do they compare to other similar studies?

Conclusions

What is the take-home message from this article? What are the recommendations? - The need for a framework to guide the adoption of open science practices. Take note of the bariers and provide recommendations. - The need for low-cost publishing - The need for policies on open science that can be implemented by various insituitions. For example, we can provide a template that can be adopted by most instituitions

1 Acknowledgement

We acknowledge the support from ... and the contribution from Mozilla Science Lab and KENET.

Use this section to acknowledge funding and resource contributions to the project.

Acknowledgments

We thank KENET for providing us with an ample environment for our hackathon.

References

1. B. Mensh and K. Kording. Ten simple rules for structuring papers. *PLOS Computational Biology*, 13(9):1–9, 09 2017.

2. P. Schlegel, M. J. Texada, A. Miroschnikow, M. Peters, C. M. Schneider-Mizell, H. Lacin, F. Li, R. D. Fetter, J. W. Truman, A. Cardona, and M. J. Pankratz. Synaptic transmission parallels neuromodulation in a central food-intake circuit. *bioRxiv*, 2016.