More Than Just a Policy - Day to Day Effects of Data Governance on the Data Scientist

Assoc. Prof. Vukosi Marivate

Context

The rise of Data Science and its use in public and private spheres also comes with policy considerations that make for an environment that both allows for the creation of data products while taking into account human factors. The reality is that there is a gap in the understanding of policy makers of the decisions, and their impacts, made by data scientists that may ultimately impact people. At the same time policy makers have held off in some instances to create an environment with clear data governance that allows for the use of data that can lead to better delivery of services and growing the local information economy.

The problem

Effective data governance requires the understanding of both policy makers and data scientists. Policy makers may not have a good understanding of the data science and pipeline. This makes it harder for policy makers to set data governance that balances between an enabling environment while protecting citizens. The African Union Convention on Cyber Security and Personal Data Protection (known as the Malabo Convention)was adopted by AU member states in 2014. It sets out to provide protections for cyber infrastructure, protection of personal information, cyber security and the necessary foundations to enable an information economy across the African continent. Even though ratified in 2014, only 8 countries had ratified the convention by 2020. This example illustrates some of the disconnects between those working to build data tools (who seek clear policies and expectations from societies) and the policy makers who may not understand the expediency of providing both protections and an opportunity to guide how the countries navigate the information economy. On the other hand, data scientists need to understand that even the smallest decisions on the data science pipeline may have unintended serious effects on people. This has been shown on how systems could discriminate due to data bias and models that are not adequately audited or tested for negative outcomes. As such this study presents a literature review that looks into the data science pipeline and how human factors permeate all parts of the pipeline. Offering definitions of typical steps taken in the data science process, the study aims to demystify areas which may be a challenge to understand for those who are not experts in data science. At the same time the study works to bridge the gap between the policy maker and the data scientists by looking into data governance and its current status on the African continent. In summary, I work to answer the questions: What are the salient points that Data Scientists should be aware of when it comes to Data Governance within organisations?; Do the current policies or mechanisms on the African continent provide a coherent view that can be used by Data Scientists to navigate and respond appropriately to the needs of the organisation/ society? Finally I work to look at the ICT for Development community for lessons learned from their practice as a resource for both the data science community and policy makers.

Background

Data science Is a field that uses scientific modelling techniques (typically from a diverse set of scientific disciplines) to extract patterns/information/knowledge from a wide variety of data

Research results

The work highlights the current understanding of human factors in the data science pipeline as well as the data governance work that works to mitigate any harm as well as enable an environment to benefit from the availability and creation of data products. On the governance side, large gaps still exist between data scientists and the policies that are available across African countries. One such gap being an uneven policy landscape where each African country has different data policies (if any) and are still to ratify the Malabo convention. For the data scientist, work has been done within the fair, equitable and ethical use of data science, machine learning and artificial intelligence. This work is ongoing and needs engagement from both data scientists and policy makers. Mitigating the harm of data products is a vital part of our practice and it requires more than just leaving it to the data scientist to make the decision.

Implications for Policy Makers: As the information economy dominates the growth in most western economies, especially through big tech, we will be facing many decision points on how countries deal with protecting data of citizens and mitigating harm of data products while enabling an environment where data scientists can work. A first step would be African countries ratifying the Malabo convention. This would just be the first step. More steps are needed to get a joint understanding between the public and data science practitioners on the role of data in society and the information economy. This will lead to better policy making as well as regulation. Lastly, the data science community also has to work to bridge their gaps with policy makers and make themselves aware of the local laws as well as fair and ethical international best practice.

Mission

To strengthen local capacity for conducting independent,
rigorous inquiry into the problems facing the management of economies in sub-
Saharan Africa.

The mission rests on two basic premises: that development is more likely to
occur where there is sustained sound management of the economy, and that such
management is more likely to happen where there is an active, well-informed group of
locally based professional economists to conduct policy-relevant research.

www.aercafrica.org

Learn More

www.facebook.com/aercafrica

www.instagram.com/aercafrica_official/

twitter.com/aercafrica

www.linkedin.com/school/aercafrica/

Contact Us
African Economic Research Consortium
Consortium pour la Recherche Economique en Afrique
Middle East Bank Towers,
3rd Floor, Jakaya Kikwete Road
Nairobi 00200, Kenya
Tel: +254 (0) 20 273 4150
communications@aercafrica.org