

Online appendix: Improving Open-Source Information on African Politics, One Student at a Time

A. Description of Data and Analysis

Figures 1 and 2 are produced with data from the Wikimedia Foundation.¹ We combine two datasets to visualize the nature of content gaps in coverage of African politics. First, we employ a dataset of topic classifiers for all Wikipedia.² Wikipedia's topic classification model treats the Wikidata item statement for each entry as a 'bag-of-words'³ to predict the entry's topic labels.⁴ We isolate all entries predicted to cover 'politics or government,' further reducing this sample to English-language entries. We then combine this data with a country-based classifier that covers the majority of English-language entries.⁵ This generates a dataset of each entry classified under 'politics or government,' allowing us to examine coverage across the countries in our sample.

As noted in the main text, Figure 2 draws on a sample of eight pages for each of these countries: the (i) current executive, (ii) first executive, (iii) incumbent political parties, (iv) main opposition⁶ political parties, (v) most recent executive election, (vi) second most recent executive elections, (vii) country's constitution and (viii) entry on 'Politics of *country*.' In the case where a page does not exist (e.g., there is no entry for the main opposition party), we code the entry as zero on both dimensions. The exception is for cases where an entry is simply not possible; Eritrea has had only one president since independence, for example. In such cases, we simply omit a count for that page.

For each page, we record the number of citations or references per article as well as the page's word count. We take these as proxies for an article's quality following the suggestion of Isaac Johnson of the Wikimedia Foundation, who noted that while there is no standardized metric of article quality, word and citation counts are a salient measure of an article's quality. While word count is a more direct measure of an article's extent, citations are an important indicator of an article's verifiability as well as its degree of representativeness (e.g., *who* is being cited).⁷

Figure A1 below replicates Figure 2 but separates out each of the eight pages. As is clear, entry quality does vary significantly for some countries, such as those that had prominent first presidents, or which have well-entrenched political parties.

Of course, these figures are only illustrative, and many factors make these countries difficult to compare – it is not surprising that Ghana fairs better than say Cameroon on English Wikipedia, for example.

¹ We thank Isaac Johnson of the Wikimedia Foundation for his assistance in helping us locate this data.

² Data available at:

https://figshare.com/articles/dataset/Wikipedia_Article_Topics_for_All_Languages_based_on_article_outlinks_/12619766. Note that this dataset is dynamic as it is part of an active project and relies on a relatively strict definition of the 'political.' Some topics related to major historical events, such as civil wars or independence movements, are often coded under other categories, e.g. 'history'.

³ This process is described here: https://meta.wikimedia.org/wiki/Research:Language-Agnostic_Topic_Classification.

⁴ Topic labels in this category are listed here:

https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Council/Directory/History_and_society#Politics_and_government

⁵ Data available at: https://figshare.com/articles/dataset/Wikipedia_Articles_and_Associated_Countries/13356326. This dataset is likewise a work in progress and is not yet complete.

⁶ In non-two-party systems, we look at the entry for the party with the second highest vote share in the previous election.

⁷ Personal email correspondence, 9 December 2020.

Similarly, Eritrea and Eswatini have small populations while South Sudan has only recently gained independence. Nonetheless, the fact that even a regional powerhouse like Nigeria – an Anglophone country – has only a third as many the entries as Mexico (itself much better represented on Spanish Wikipedia) highlights the striking content gaps in African coverage.

B. Description of Assignment

We describe a term-long assignment for undergraduates conducted in collaboration with Wikiedu. Wikiedu is a non-profit organization dedicating to promoting student contributions to Wikipedia. Put simply, WikiEdu asks students to create or expand on a Wikipedia entry related to course content. This demands that students first identify a topic that they believe needs improvement and assess the feasibility of conducting research on that topic. Because instructors stipulate the assignment's timeline on the class WikiEdu dashboard, students see a clear plan for the development of their article over the course of the semester. Accompanying trainings guide students through the construction of their articles, honing the development of necessary technical skills to evaluate⁸ and edit articles on Wikipedia, as well as citation and neutral-language writing skills. Over the course of the term, students will scaffold their skill development as they learn to write for Wikipedia's platform. For an example of how the tutorials connect with student learning: students learn about digital literacy skills by having a tutorial that they click through online. They then have a homework assignment to examine the number and quality of citations as well as the frequency with which an article has been edited. Finally, there is an in-class discussion surrounding questions of identifying high quality articles and potential biases.

Assigned peer-review exercises and the option of end-of-term student presentations create opportunities to students to receive and give feedback as well as to learn from their classmates' research.⁹ The final entry, the ultimate output for the term, is often complimented with a reflective essay. This latter assignment is particularly useful as a vehicle to encourage students to interrogate their own biases (as well as those they encounter during the course of their research) and privilege in what scholars of higher education called "metacognition" (see Bransford, Brown, and Coking 2000).

The WikiEdu platform offers instructors flexibility in how they want to structure students' experiences. For example, the assignment can be done in groups or by individuals and instructors can either assign topics or leave this up to student choice. In our courses, we have given students a free rein in their choice of topics, either for any African country *or* for an assigned country that they develop expertise in throughout the term. Recently one of us has also experimented with placing students in country pods, where a small group of students collectively gain expertise on an assigned country and use their pod-mates as sounding boards on their individual projects. Students noted this worked especially well during online teaching to improve the social component of the course during the COVID-19 pandemic, while allowing them to pursue their individual pages. Although collaborating with WikiEdu is free and generally accessible to instructors at all types of institutions, instructors working with less resourced students may need to ensure that students have sufficient internet and computing access.

C. Works Cited

- Bransford, J.D., A. L. Brown, and R.R. Coking. 2000. *How people learn: Brain, mind, experience and School*. Washington D.C.: National Academy Press.
- Cassell, Mark. 2018. "When the World Helps you Teach Your Class: Using Wikipedia to Teach Controversial Issues." *PS: Political Science and Politics* 51 (2):427-433.
- Timperley, Claire. 2020. "The Subversive Potential of Wikipedia: A Resource for Diversifying Political Science Content Online." *PS: Political Science and Politics* 53 (556-560).

⁸ See Timperley (2020) on how this helps students become critical consumers of Wikipedia.

⁹ As suggested by Cassell 2018.