Plan Overview

A Data Management Plan created using DMPonline.be

Title: Post-colonial path dependency, was Congo destined to be a failed state? A closer look at forced cultivation and concession policies

Creator: Justin Mutambeshya

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Project abstract:

Despite being one of the largest and most populous countries globally with abundant geographical blessings, the Democratic Republic of the Congo (DRC) continues to grapple with a tragic postcolonial path marked by enduring political chaos, endemic violence, and economic precarity. This research proposal aims to quantify and uncover the channels of persistence through which the forced cotton cultivation policy ("cultures obligatoires") instituted by the Belgian government in 1917 impacted the Congolese development trajectory. The first working package involves constructing a digital map delineating the regions subjected to the “cultures obligatoires” policy between 1917 and 1935 in Colonial Congo. Subsequently, the second working package will leverage this map to uncover the effects of the “cultures obligatoires” policy on outcomes like natality, gender ratio and population densities on the short- to midterm utilizing census and survey data from 1955, 1970 and 1984. Finally, the third working package utilizing Demographic and Health Surveys (DHS) data from 2007 and 2014 and self-collected survey results, will analyze the long-term outcomes and delve into the channels through which the impacts of the forced cotton cultivation policy persist.

ID: 214064

Start date: 11-01-2024

End date: 11-01-2028

Last modified: 31-03-2025

Post-colonial path dependency, was Congo destined to be a failed state? A closer look at forced cultivation and concession policies

Application DMP

Questionnaire

The questions in this section should only be answered if you are currently applying for FWO funding.
Are you preparing an application for funding?

Question not answered.

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

Question not answered.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

Question not answered.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

Question not answered.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

Question not answered.

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

Question not answered.

For whom might your data be useful outside of the research project, e.g. researchers or other stakeholders? How will you share this data?

Question not answered.

Post-colonial path dependency, was Congo destined to be a failed state? A closer look at forced cultivation and concession policies

DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- No

Post-colonial path dependency, was Congo destined to be a failed state? A closer look at forced cultivation and concession policies

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Yes

Post-colonial path dependency, was Congo destined to be a failed state? A closer look at forced cultivation and concession policies

FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
(public) Archival data and maps from state archives and Africa Museum (originating in the Belgian Congo era)	see dataset name	Generate new data	Physical, but my work will digitizalize some parts of it	aggregated data	.csv	• <100MB	
(public) HYDE database	History database of the Global Environment	Reuse existing data	Digital	compiled data	.csv	• <100MB	
(public) Night time satellite data from Earth Observation Group	see dataset name	Reuse existing data	Digital	observational	.csv	• <100MB	
(public) Survey data of the Democratic Republic of Congo from 1955-1957	see dataset name	Generate new data	Physical, but my work will digitizalize some parts of it	observational	.csv	• <100MB	

census data for 1970-1984	see dataset name	Reuse existing data	Physical, but my work will digitizalize some parts of it	observational	.csv	• <100MB	
(public) Demographic and Health Surveys (DHS) data from 2007 and 2014	see dataset name	Reuse existing data	Digital	observational	.csv	• <100MB	
(confidential) Own Survey of a Congolese village	survey about political beliefs of villagers	Generate new data	Digital	experimental	.csv	• <100MB	
(public) Caloric suitabilty index (Galor & Ozak)	see dataset name	Reuse existing data	Digital	compiled	.csv	• <100MB	
(public) cotton suitablity index (FAO GAEZ DB)	see dataset name	Reuse existing data	Digital	compiled	.csv	• <100MB	
(public) Murdock Ethnographic data	see dataset name	Reuse existing data	Digital	observational	.shp	• <100MB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

- https://eogdata.mines.edu/products/vnl/
- https://www.dhsprogram.com/data/available-datasets.cfm?utm_source=chatgpt.com
- https://ozak.github.io/Caloric-Suitability-Index/
- https://gaez.fao.org/
-

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data

The survey regarding a Congolese village will entail questions about ethnic group, village of origin and birth, migration, politics and values. Those answers will have to be kept confidential. Therefore I will take the following steps:
1. Establish a clear data collection procedure and avoid collecting personally identifiable information
2. Educate myself on correct data handling procedures
3. Use strong authentication methods such as multi-factor authentication for accessing sensitive data
4. Encrypt survey data both during transmission and while stored on servers or local devices
5. Store survey data on secure servers

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes

The survey regarding a Congolese village will entail questions about ethnic group, village of origin
and birth, migration, politics and values. Those answers will have to be kept confidential as they are quite personal. The data type will be .csv

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

Some maps used in my work come from the Africa MUSEUM and are therefore their property and cannot be shared with other without their permission

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

To ensure that the data I compile and process remains understandable and usable both during and after the research process, I will prioritize comprehensive and accessible documentation. Specifically, I will create detailed README.txt files to accompany each dataset, providing contextual information about the data sources, cleaning steps, merging procedures, variable descriptions, and any other relevant notes.
Additionally, I will produce Codebook.tsv files wherever necessary to systematically describe the variables included in the merged database. These codebooks will include variable names, definitions, units of measurement, coding schemes, and any transformations applied.
This documentation will be especially important for the database I am constructing for my research on the long-term impact of forced cotton policies in Congo. By maintaining clear, consistent, and standardized documentation practices, I aim to ensure that the dataset remains transparent and reusable for myself and other researchers in the future.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

Maybe in the future I will consider adding some metadata though.

3. Data storage & back-up during the research project

Where will the data be stored?

On my work laptop and the KU Leuven Onedrive server.
The confidential survey data from the Congolese village will be stored on separate and secured harddrive.

How will the data be backed up?

The KU Leuven Onedrive servers already have multiple back ups of the files created on it.
I will have the separate drive for the survey backed up to other storage places as well.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

The KUL Onedrive servers provide each user 1TB of space if I am not mistaken. That is more than enough for my research purposes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

I will only work on my work laptop on the data which is already secured with bitlocker and a personal password. Additionally the KU Leuven Onedrive servers are ofcourse protected by Microsoft protocols.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

0 euro. If any costs would come up, these could be covered by the FWO working costs allowance.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All the historical data that I will digitize, will be made available to the broader public. No data will be erased after the end of the project.

Where will these data be archived (stored and curated for the long-term)?

TBA

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

TBD

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

If access is restricted, please specify who will be able to access the data and under what conditions.

Access will not be restricted

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- Yes, Privacy aspects
- Yes, Ethical aspects

The survey regarding the political beliefs of villagers in a (to be determined) Congolese village will be confidential data and will not be shared with other because of privacy and ethical concerns.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

TBA

When will the data be made available?

TBA

Which data usage licenses are you going to provide? If none, please explain why.

TBA

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

Yes, but I am not at that stage yet.

What are the expected costs for data sharing? How will these costs be covered?

If there any costs those will be covered by the yearly FWO working costs allowance.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Justin Mutambeshya

Who will manage data storage and backup during the research project?

Justin Mutambeshya

Who will manage data preservation and sharing?

Justin Mutambeshya

Who will update and implement this DMP?

Justin Mutambeshya