Protocol for Developing a Spatial Layer of Indigenous Population Areas using the World Language Mapping System Dataset

Summary

This protocol is designed to use the World Language Mapping System spatial dataset to develop spatial data on where predominant ethnic groups have characteristics of Indigeneity (henceforward Indigenous population areas) documented in the literature. Characteristics of Indigeneity are drawn from Section F of the United Nations Report on 'The Study of the Problem of Discrimination Against Indigenous Populations' by Special Rapporteur, Mr. José Martínez Cobo (the Martínez Cobo Study).

Background on Dataset

The World Language Mapping System is the most comprehensive and widely used spatial dataset of the world's living languages. In this protocol, we match the languages included in the World Language Mapping System dataset to the ethnic groups speaking those languages. The ethnic group name is usually, but not necessarily, the same as their language (e.g. Pgaz K'Nyau People speak Pgaz K'Nyau). Many ethnic groups have multiple names, which our protocol is designed to include.

Preparation Prior to Screening

The Martínez Cobo Study of the United Nations, also known as the Study of the Problem of Discrimination Against Indigenous Populations by Special Rapporteur, Mr. José Martínez Cobo, developed a 'multi-criterion approach' for Indigeneity, arguing that singular approaches serve exclusionary purposes. This report formed the underlying principles of the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP), including self-identification and external recognition. Special attention should be paid to Chapter V: Definition of indigenous populations E/CN.4/Sub.2/1982/2/Add.6, particularly Section F - The Multiple Criterion (paragraphs 236 – 239).

The Multiple Criterion approach is advantageous due to numerous shortcomings of singular approaches for defining Indigeneity. A significant portion of Chapter V of the Martínez Cobo Study elaborates on the pitfalls of singular approaches. The Multiple Criterion approach recognizes that Indigenous identities are complex. Rather than assuming a simplistic definition, the Multiple Criterion approach works by considering multiple criteria to not unfairly exclude any ethnic group by considering only one aspect of Indigeneity. Indigenous Pgaz K'Nyau professor, Prof. Dr. Suwichan Phatthanaphraiwan, recommended using the Martínez Cobo Study as the basis for conceptualizing Indigeneity in Southeast Asia, because the report is widely used by Indigenous activists, advocates and scholars in the region.

We have summarily rejected some historical approaches of classifying Indigenous peoples on the basis of ancestry, bloodlines or race, given the racialization, discrimination and subjugation that has accompanied these practices, in addition to the many other limitations outlined in Section B – Ancestry of Chapter V of the Martínez Cobo Study (especially paragraphs 15 – 39).

Our first screening criteria of self-identification follows paragraph 9 of Chapter V of the Martínez Cobo Study that cites one of the five principles of the World Council of Indigenous Peoples that "*the right to define what is an indigenous person is reserved for the indigenous peoples themselves. Under no circumstances should we let artificial definitions such as the Indian Act in Canada, the Queensland Aboriginal Act 1971 in Australia, etc. tell us who we are*."

The multiple criterion approach outlined in Chapter V of the Martínez Cobo Study is summarized as follows:

1. Self-identification: the group self identifies as Indigenous by participating in Indigenous organizations or has issued documents stating they are Indigenous.
2. External recognition: State or international institutions recognize the ethnic group as Indigenous.
3. Geography & Territory: Ethnic group resides in Indigenous reserves or Indigenous-managed or -governed or militarized areas (e.g. ethnic states in Myanmar). Alternatively, the ethnic group may have their own spatial organization and governance structures distinct from the ethnic majority group and the state.
4. Cultural Distinctiveness: The ethnic group is culturally distinct from the ethnic majority group with different religious and spiritual beliefs and livelihood practices (e.g. shifting cultivation, alternative staple crops, etc.).
5. Relational Non-Dominance: The ethnic group is distinct from the ethnic majority group and/or the ethnic group with majority representation in government. In addition, the ethnic group has or is experiencing asymmetric power relations with the ruling majority group, including historical or ongoing colonization, marginalization, land dispossession and discrimination that may range from overt violence to limited political control and autonomy, and assimilation.

The protocol outlines key word searches to identify whether any of the aforementioned characteristics of Indigeneity are documented in the literature for ethnic groups. If any of these characteristics of Indigeneity are documented for an ethnic group in the literature, we will consider them Indigenous based on the methods and approach in this protocol that are founded on the Martínez Cobo Study.

Data Preparation and Screening Procedure

Data preparation prior to screening includes:

1. Extract list of languages from the World Language Mapping System dataset for target countries.
2. Match language to ethnic group (search on Ethnologue).
3. Record source (i.e. website name and link) that matches language to ethnic group in language tab of data compilation spreadsheet.
4. Check for alternative spellings and names of the ethnic group and add these to the spreadsheet.

Screening for Ethnic Minority Status:

1. Google country name + ethnic demographics
2. Identify the ethnic majority group and record the quote (and page number if it is a document) and the source in the Ethnic Minority tab of the data collection spreadsheet.
3. For other ethnicities, google the name of the ethnic group + country + population OR demographics. Record the population number or percentage of the total population and the source in the Ethnic Minority tab of the data collection spreadsheet.
4. Determine whether the ethnic group is an ethnic minority with a Y for Yes, N for No, or Unsure in the Ethnic Minority column of the Ethnic Minority tab.

Screening procedure for Indigeneity by criterion:

A. Criterion 2 - Self-identification
    1. Google scholar search: (the name of the ethnic group OR alternative name) AND country AND (“self-identi*” OR “self-determin*”)
    2. Check the first two pages of results to see if any sources mention self-identification or self-determination for any of the ethnic groups in our list. If so, copy and paste the relevant quote, page number and source in the correct corresponding column of the data collection spreadsheet.

3. If no relevant results are found on google scholar, check the first two pages of a regular google search using these search terms, and record the relevant quote, page number and sources accordingly.
4. Check for key Indigenous organizations representing Indigenous Peoples in the country and read their documents. If the ethnic group is included in their organization's documents, we will conclude that they self-identify as Indigenous, because they participate in Indigenous organizations representing Indigenous Peoples.

B. Criterion 1 - Recognition

Part 1 - Constitution

1. Download the English version of the national constitution. Note: there may not be an official English translation. If an English version is available, record the legal document website link under the source column in the National Constitution tab of the data collection spreadsheet.
2. Search (Ctrl + F) for 'Indigenous' and 'Indigeneity' and 'ethnic minority' and 'ethnic minorities' and 'ethnic groups' and 'tribes' and 'tribal' and 'traditional' in the document.
3. Record pages where these are mentioned in the National Constitution tab of the data collection spreadsheet. If there is no mention of any of these terms, record a 'N' in the spreadsheet. If there is no English version of the constitution available online, record 'no data.'

Part 2 – External institutional recognition

Step 1:

1. Check the International Labour Organization's Status Report of Indigenous Peoples in the World of Work in Asia and the Pacific and the United Nation's State of the World's Indigenous Peoples report for country being screened.
2. Check table 3 on page 39 of the United Nation's State of the World's Indigenous Peoples report to see whether the United Nations considers that ethnic group to be Indigenous. If so, record under the International (UN) column of the Legal Recognition tab of the data collection spreadsheet.
3. Scan the country section of each document to see whether the country offers legal protections to the ethnic group. Record Y for Yes, N for No or Unsure for Unsure, the page number and the source in each relevant column.

Step 2:

4. Google scholar search 'Indigen*' AND country AND ethnic group name AND ("legal recognition" OR rights OR legislation OR law)
5. Check the first two pages of results to see if any sources mention legal protections for any of the ethnic groups in our list. If so, record (or the page number if it is a link to a document) in the relevant excel cell and the source name, page number and link.
6. If there are no relevant sources on google scholar, search on google.

Step 3:

7. Google scholar search the name of the ethnic group AND country AND (constitution OR "cabinet resolution" OR resolution OR declaration OR law OR rights)
8. If no relevant results are found on google scholar, check the first two pages of google using these search terms.
9. Record the relevant quote, the page number (if a document) and the source under the relevant legal document column under the Recognition tab if the ethnic group is afforded special protections.

10. Record the page number and the source in the discrimination column of the Relational Non-dominance tab of the data collection spreadsheet, if the ethnic group is negatively discriminated against in any legal document.
11. If the ethnic group is not mentioned in the legal document, record 0 for 'not mentioned.'

C. Criterion 3 – Geography & Territory

1. Google scholar search: (the name of the ethnic group OR alternative name) AND country AND (territory OR "Indig* reservation" OR "commun* ownership" OR "commun* land" OR "collective land" OR "land tenure" OR "Indigenous Conserved Area" OR "Indigenous protected area" OR "ethnic state" OR bioregion OR "trib* lands" OR "Indigen* land")
2. Google scholar search: name of the ethnic group AND country AND (clan OR kinship OR decision-making OR consensus OR "community leader" OR "trib* leader" OR "village leader" OR "spiritual leader" OR "commun* justice" OR "traditional justice" OR "traditional legal systems" OR "trib* affiliation" OR "trib* governance")
3. Repeat the search in google if there are no relevant sources on google scholar.
4. Look through the first two pages. Record Y/N/Unsure, the relevant quote, page number and source in the corresponding column of the Geography & Territory tab of the data collection spreadsheet.

D. Criterion 4 – Cultural distinctiveness

1. Google scholar search: (the name of the ethnic group OR alternative name) AND country AND (culture OR language OR religion OR animis* OR shaman OR livelihood OR swidden OR "shifting culti*")
2. Look through the first two pages. Record Y/N/Unsure, the relevant quote, page number and source in the corresponding column of the Culture tab of the data collection spreadsheet.
3. Repeat the search on google if there are no relevant sources on google scholar.

E. Criterion 5 – Relational non-dominance

1. Google scholar search: (the name of the ethnic group OR alternative name) AND country AND (persecution OR genocide OR conflict OR violen* OR "forced conscript*" OR "forced labor" OR "indentured labor" OR "forced relocation" OR "forcible relocation"
2. Google scholar search: (the name of the ethnic group OR alternative name) AND country AND (coloniz* OR discriminat* OR "settler colonial*" OR "internal colonial*")
3. Google scholar search: (the name of the ethnic group OR alternative name) AND country AND (dispossess* OR landgrab* OR "land appropriat*")
4. Google scholar search: (the name of the ethnic group OR alternative name) AND country AND ("political representation" OR "political power")
5. Look through the first two pages. Record Y/N/Unsure, the relevant quote, page number and source in the corresponding column of the Relational Non-dominance tab of the data collection spreadsheet. Repeat the search on google if there are no relevant sources on google scholar.

Note on colonization: We are interested in relational power structures under colonialism. European colonizers would often leverage and exacerbate existing ethnic inequalities for imperial gain. It is these relational power dynamics under colonialism that we are interested in, rather than experiences of

colonialism shared with the ethnic majority group (I.e. everyone was colonized). We are also interested in how current ethnic conflicts could be a result of colonial influence.

*Note on alternate spellings of ethnic group names: As you search, record any alternate spellings for the ethnic group and the corresponding source in the Language – Ethnicity tab of the data collection spreadsheet. After completing the previous screening steps, google search the country + alternate spelling of the ethnic group. Check the first two pages of google and google scholar. If the alternate spelling is mentioned together with the most common spelling in most of the results, then no further search is needed. If the alternate spelling is frequently mentioned without the common spelling, then repeat your search for Criteria 1 – 5 with the alternate spelling of the ethnic group name that arose during your search.

Final Steps to Summarize Findings

After completing your search about the ethnic group for each criterion, summarize and record whether the ethnic group has characteristics of Indigeneity in the 'Indigenous' tab of the data collection spreadsheet. If any characteristic of Indigeneity is present, record 1 for that ethnic group in the Indigeneity tab of the data collection spreadsheet; if none of the Indigeneity characteristics were found in the literature search, then record an 0. If you are unsure, record 9 in that column and comment on why you are unsure.

Repeat this process for each ethnic group.

With any uncertainties about source credibility or vagaries of whether the ethnic group meets a criterion, please record your notes and comments under the 'Notes / Comments' columns of each tab. If you choose Unsure out of Yes / No / Unsure, then make a comment about why you are unsure.

Developing the Spatial Layer

Open the World Language Mapping System spatial dataset in ArcGIS Pro or QGIS. Clip the global dataset to the countries included in the study. Open the attribute table and add a new text field. You can either manually enter the final Indigeneity code (0, 1 or 9) for each ethnic group in the attribute table, or you can join a table. If you choose the second option, ensure that the ethnic group names match. Once you have entered the Indigeneity codes for each ethnic group, you can modify the symbology to see the unique values. Merge the polygons for each Indigeneity value (0, 1 and 9) to produce a final aggregated spatial layer of Indigenous population areas (areas where the predominant ethnic group has documented characteristics of Indigeneity in the literature).

Send comments or feedback on this protocol to lmz5288@psu.edu.