Frequently Asked Questions about the HIS Registry of Peoples

What is the ROP?

The Registry of Peoples (ROP) is a standard reference database of all peoples of the world, and is part of the Harvest Information System. The ROP serves as a resource to other databases by providing a unique code and standard recommended name for each "ethnic entity" (also called people group, people or ethnolinguistic entity) in the world.

The purpose of the ROP code is to facilitate comparing and exchanging information without the need to merge everybody's information into one huge database.

While the ROP does not maintain population or other demographic information, a supplemental table of population from some major sources is included, which helps in identifying and comparing the entities. This should be helpful particularly when first applying ROP codes to a database.

The ROP references each people with the same code in all countries where they are found. While the heart of the ROP is the code table tblROP3people, other reference tables are included in the database to facilitate the linking of each ROP People to other Harvest Information System registries including the Registry of Geography (ROG), Registry of Languages (ROL), and the Candidate Registry of Religions (ROR). Reference information included for each ROP3People code serves to further "define" the people group so that users of the code set can accurately apply the codes.

How does the ROP get its data?

The ROP began as a compilation of major world databases, with ongoing contributions by various field research sources and agency database updates. Some contributing databases were full sets of world data, while some were discrete sets from certain countries or regions. For the sake of linking people groups from various sources, the ROP provides codes for people groups as they are submitted from various clients.

Why is my agency's list of peoples different from the ROP listing?

Various researchers or agencies vary in the way they view ethnicity or analyze what they discover about cultures. Because of the differences between various databases, due to different agency strategy needs, specific data structure, purposes of the agency and the range of data carried, different databases may account for the same ethnic entities differently.

This may introduce duplicates or overlapping entries. The editing process attempts to reconcile these as closely as possible. Verification and correlation, including updates from sources, is ongoing to discover such duplicates or overlaps.

ROP records are extensively annotated and cross-referenced to provide clearer identification of the entries. All edits which affect the meaning of a code or change a name are documented in the Change History table of the ROP.

How do the codes assist in comparison and exchange of information?

The ROP code serves as a standard reference that will relate entities in various databases as a basis for comparing and exchanging related information. Typically this will be on a one-to-one basis, entry by entry, with a direct code match.

What if the seemingly related people groups don't match in the different databases?

If one database has several smaller related groups as separate entities where the ROP has one entry for a larger defined entity which includes those several smaller ones, the data partner can request that additional ROP codes be created. In cases where individual people group lists do not agree, ROP attempts to give direction to partner in how to apply codes. For example, the memo field may contain a statement such as: "use this code for *people name* undifferentiated by dialect subgroup". Where this comment is found, ROP will include additional codes for cases where people group lists have differentiated additional subgroups based on dialect or other factors.

Greater flexibility is included when assigning language codes from the ROL. ROP codes and ROL codes are separately and independently assigned. This allows a more precise designation, identifying combinations of ethnicity characteristics and language to more accurately determine ethnic identities.

A database may have inadvertently assigned the wrong ROP code to an entity. When ROP codes for what is thought to be the same entity in two databases do not match, it provides an opportunity for both to compare internal records to find duplicates or overlaps possibly causing the discrepancy. Where potential discrepancies are noted, the ROP Editor will typically inform the partner maintaining the people group list so that code assignments can be reviewed.

How does the ROP deal with overlaps and duplications caused by these differences in databases?

ROP attempts to account uniquely for every identifiable ethnicity, ideally counting each individual only once and listing each unique ethnicity only once. Where complex cases are difficult to fully reconcile, memos indicate possibilities and common approaches in sources. However, the Registry of Peoples is a composite of partner people group lists, not a stand-alone people group list. The ROP catalogs and cross-references peoples from all partner lists. In cases where partner lists classify peoples in different ways, there is likely to be some level of overlap and/or duplication.

The editor reviews the ROP records and verifies entities. This involves a full manual verification of every entity originally reported to the ROP by all contributing databases in the original compilation. In this process, overlaps and duplicates are being identified and rectified. Thus, major variations involving similarities and differences among major databases are gradually documented in the ROP itself.

Annotations in the Memo fields provide history and tracking clarification. The ROP attempts to represent the "real world." We sometimes find that the edges and boundaries of ethnicities are fuzzy. We deal with data fields and categories. Real life is quite dynamic, continually challenging our categories and eliciting new ones to account for what we find.

Why does the ROP not have every people we show? It appears that the ROP has as one people with only one code what shows up in our database as several smaller peoples.

ROP attempts to account for every ethnic entity in some way. It is always possible to divide any identified ethnic group into even smaller segments or subgroups. When partner people group lists subdivide what was once thought to be a single people group into multiple people groups, it is important to notify the ROP Editor of changes. This will allow for review and update of the registry.

The unique ROP code for each ROP entity will provide a standard reference for any database to show where its data correlate with those of any other database, without placing limits or imposing controls on what any other researcher or database might include.

Common differences in divisions of related ethnic groups are noted in the ROP editor's comments. These memo fields are intended to assist code users in correlating subgroups or related peoples.

Should we change the name in our database to match the ROP names? We have names organized so we can relate the group easily to our strategy teams and projects.

It is not the goal of the ROP to address the strategic goals of contributors by including or excluding entities that are on contributing sources' lists, which serve their strategic needs. The ROP attempts to account for every ethnic entity in some way, from the best information from multiple sources, representing the real-world situation.

The ROP editing process attempts to regularize name forms and provides an extensive listing of alternate names to enable data managers to correlate their data to the ROP codes. Likewise the codes provide a mechanism for comparison of data between any two or more databases that use the codes to identify their entries.

The ROP recommended name for an ethnic entity is based on several factors. The form most commonly found in sources is considered. In general the English form of a name is preferred over another European-language form.

The people's self-name is given more weight, especially in lesser-known entities or where it helps to distinguish between similar names of different ethnic entities. A western form of a common name might be used where it is already well-known and established in academic disciplines. In some cases names are descriptive to differentiate between peoples with different ethnolinguistic characteristics that use the same name for themselves, such as the Mixtec and other clusters in Mexico. Such names are referred to as "construct names."

Is the ROP a standalone database, not integrated with the other HIS registries?

The ROP codes and names or any tables can be used as standalone resources without reference to other HIS registries. However, ROP uses codes from tblROG to indicate Primary Country, tblROL to indicate Primary Language, and tblROR to indicate Primary Religion. In this way ROP supports and relies on the work of other registry editors.

What about updated ROL (SIL) language codes?

ROL codes are downloadable directly from Ethnologue: Languages of the World.

Updates are provided in a separate database, containing the updated code set, with a Change History documenting changes to codes. Users will encounter some difficulty, however, in making necessary changes to their database. Some have found that the info provided in the Change History table is insufficient to enable a proper updating of old codes.

For instance, when an old language is analyzed as two or more languages with new codes, it is difficult to tell which previous dialects of the language code are now associated with which of the new codes.

How can the ROP help see which languages are spoken by which peoples?

For most users, the ROP Geopolitical Index will be the most helpful, as it gives the fullest description and can be related to various sources either by ROL code or by the people code. This is particularly helpful in the initial correlation stage, when you are first assigning the ROP codes. As edits occur in the ROP with every new edition, this table will probably best reflect most clearly all changes to any particular entity.

Why is the ROP population sometimes different from other sources?

ROP is not a source for populations. Some representative populations from some sources are included as supplemental information. This information is helpful for comparative or identification purposes, but partner people group lists should always be consulted for the most up-to-date population estimate for any given people group.

How do I query "all people groups in Mexico >= 100,000 population" on the HIS? I didn't see population info on the ROP?

ROP is not a primary source of population for any people group. Populations from several standard sources are provided for reference in tblROP3suppData (linked to tblROP3geoIndex on the GeoID field). One or more of these sources should have a population estimate. Thus, if you want to query by population, you would specify the source. Not all sources have populations for all entities, and some will have only one source, so your query might not produce a definitive result.

The purpose of the ROP is not to maintain demographic information, but to provide a set of unique standard codes to enable data providers and managers to compare and exchange data. Demographic information would be provided by the various data providers in agencies using the ROP codes.

The ROP3 code and PeopleName in tblROP3people provides the unique "definition" of each ethnic entity along with a Primary Location descriptor and Primary Language descriptor. Entries in tblROP3geoIndex provide a more complete listing of countries where the People "defined" by each code are found. In this table you'll see that PeopleName sometimes varies by country and that Primary Language (PROL) and Primary Religion (PROR) sometimes varies by country.

ROP Documentation provides an introduction to the various tables included in the registry, which might be helpful in determining how you could best assign the correct ROP codes to your data records.