World Language Mapping System
Version 19

In This Document

Other Documents

Introduction

The World Language Mapping System (WLMS) has been a project of and Global Mapping International (GMI) and several contributors. At this release, all of the living languages listed in 2016 version of ISO 639-3 standard are represented as points in the geographic files, and over seventy-five percent of the languages are represented as boundary polygons representing the region within each country, which is the traditional homeland of each indigenous language. Languages, which are very widely dispersed within a country, are generally not represented with polygons, but are included in special language point layer.

Features

- Polygons delineating the linguistic homelands of many of the language-in-country entries in the 19th edition (2016) ISO 639-3 standard, www-01.sil.org/iso639-3/default.asp. Exceptions include extinct or nearly extinct languages and languages whose use in a country is too diffuse to be mapped. These language areas are made to fit with coastlines and international borders corresponding to the Digital Chart of the World (VMAP level 0, edition 5).

- Polygon overlays for mixed-language areas.

- Point locations of all living 19th edition (2016) ISO 639-3 standard languages by country.

- Layer files providing default symbology for ArcGIS.

- Sample map project (.mxd) files illustrating use in ArcGIS.

- XML metadata compatible with ArcGIS.

- Vector country boundaries and ECW-compressed raster elevation colors and hillshades to provide geographic context. Additional geographic context can be added using GMI's Seamless Digital Chart of the World or other forms of the Digital Chart of the World (VMap Level 0) base map.

What's New in Version 19

- Point and polygon data have been updated to correspond to the 19th edition (2016) ISO 639-3 standard.
- Due to licensing issues this version does not have most of the attribute information that was in previous versions of the WLMS. If you would like this data, contact the Ethnologue, https://www.ethnologue.com/products.

Installation

The World Language Mapping System data is compressed in setup.exe in the downloadable zip file or CD-ROM, which must be executed on a Windows system. Setup.exe will prompt for an installation password, which is provided by the WorldGeoDatasets.com store on the download page and confirmation email. If the WLMS is being installed for use in conjunction with GMI's Seamless Digital Chart of the World data product, the installation directory chosen should be the same as that chosen for the DCW; the two products may be installed in either order. For normal first-time installation, choose the Complete installation option. If you need to restore a fresh copy of a part of the data at some later time, run setup.exe and select Custom installation. This allows you to select individual feature classes to be reinstalled.

Licensing

This data set is covered by an [End-User License](), and is subject to both copyright and license restrictions. In particular, redistribution of the data is not allowed except as specifically permitted in the End User License Agreement or by explicit, prior, written permission of Global Mapping International.

Release History

For changes in current release, see [What's New in Version 19]()

World Language Mapping System version numbers prior to version 16 correspond to the version numbers of the GMMS DCW Data Supplement from which they are derived. Changes in major and minor version levels did not reflect the extent of changes in the WLMS.

Changes from Version 17 and Version 16

- Point, polygon, and attribute data have been updated to correspond to the 17th Edition of the *Ethnologue.*

- Most languages have been given a vitality estimate using [EGIDS (Expanded Graded Intergenerational Disruption Scale)](). This data is available as a field in all feature classes and most tables. Two sample Language vitality (points).lyr and Language vitality (poly).lyr have been added. One of these is for language points and the other for language polygons. A sample mxd file, Vitality of Languages along Bolivia-Brazil border.mxd, has also been added.

- A file geodatabase, WLMS.gdb, is included which contains all files except raster elevation colors and hillshades, as well as three additional tables ([immigrant languages](), [extinct/dormant languages](), [sign languages]()). The file geodatabase is located in the \data folder. Future versions of the WLMS may not include shapefiles and dBase tables.

- The rother.shp, lang_other_pt.shp, and lang_extinct_pt.shp shapefiles have been discontinued. The first two didn't contain many languages and many languages formerly in them are now included in other feature classes. The immigrant language table will have any remaining languages which would have been in these first two shapefiles. The WLMS file geodatabase contains a table of extinct languages corresponding to lang_extinct_pt.shp.

- In the eth_wlms attribute table a macrolanguage class (M) has been added to the LMP_CLASS field to indicate which languages are macrolanguages. A macrolanguage is "multiple, closely related individual languages that are deemed in some usage contexts to be a single language."

- The *Ethnologue* is using a vastly different table and field structure. Thus some fields formerly in eth_wlms.dbf, all of which had data for only a few languages, will be discontinued. The data in some fields will be slightly different. We have attempted to preserve as much of the old field structure in the eth_wlms table as possible.

- Metadata for files was entered using the ISO 19139 Metadata Implementation Specification. This will include a description of data in each field. Similar information is in the [Detailed Data Documentation]() section in this document.

- Yearly updates are planned due to the yearly update schedule for the [ISO DIS 639-3]() code standard.

Changes from 3.2.1 (15ᵗʰ Edition *Ethnologue*) and Version 16 (16ᵗʰ Edition *Ethnologue*)

- **Point, polygon, and attribute data have been updated to correspond to the 16ᵗʰ Edition of the *Ethnologue.***

- **Language polygons for widespread languages have been added in nearly all populated areas where there is no minority language and there is a single, dominant national language, e.g. Spanish in Latin America, Portuguese in Brazil, etc.**

- **Layers were added to display SIL language families, Ruhlen language families, and population of languages for points and polygons. A table was added to cross reference SIL language families with Ruhlen language families.**

- **When installed with the Seamless Digital Chart of the World, language features can be controlled by the base map wizard.**

Changes from 3.2 to 3.2.1

- **Language areas have been added for English, Spanish, and French in the U.S. and Canada**

- **Languages polygons have been moved from Other Language Areas (rother.shp) to the main Language Areas (langa.shp) where the Ethnologue *database* contains a language record (usually a record for an immigrant group, designated by field IM = Y). Previously, languages were only included in langa.shp where the Ethnologue *book* (which excludes immigrant languages) included an entry for the language group.**

- **Country and language boundaries, with corresponding data attributes, have been adjusted to reflect the split of Serbia and Montenegro planned for January, 2007.**

- **A number of other minor corrections have been made to the data.**

Changes from 3.1 to 3.2

- **Point, polygon, and attribute data have been updated to correspond to the 15ᵗʰ Edition of the *Ethnologue.***

- **Unified point shapefile allows the use of classification and/or definition queries to display various classes of points (polygon centroids, widespread language points, extinct language points, etc) formerly represented only as separate shapefile layers. The separate shapefile layers are retained for back compatibility.**

- **Additional attribute data in shapefiles requires less frequent joining of the *Ethnologue* attribute table for many applications.**

- **44 additional fields of information in the *Ethnologue* data table reflect the additional scope of the published 15ᵗʰ Editon *Ethnologue,* an improved breakdown of data into separate fields, and improvements (described below) in geographic coding, language coding, name rendering, and population.**

- The *Ethnologue* attribute table now has decimal-degree latitude and longitude fields added, allowing it to be mapped directly as an XY or event table.

- Languages are coded using both the [ISO DIS 639-3](#)-compliant codes used in the 15th Edition *Ethnologue* and a legacy code compatible with earlier versions of the Ethnologue. For clarity, legacy codes are given in the traditional upper case, while ISO 639-3 codes are given in lower case. Several codes for language-in-country are provided:

- Legacy *Ethnologue* codes with upper-case ISO three-letter country codes (the previous WLMS standard for language-in-country coding).

- ISO 639-3 codes with lower-case ISO two-letter country codes (the recommended new standard for ISO-compliant language-in-country coding).

- ISO 639-3 codes with upper-case U.S. government FIPS 10-4 country codes (the recommended standard for those already using the widely-used FIPS standard for country coding, and for those doing historical research, since FIPS codes represent consistent geographic entities over time).

- Name rendering for map labels is substantially improved, with traditional upper-case ASCII (diacritic-stripped) names supplemented by mixed-case ASCII, mixed-case ANSI (ISO Latin 1) diacritical forms, and comma-free ANSI diacritical forms.

- Populations of most languages are now provided as a numeric field in addition to the earlier text discussion of population research.

- The Widespread Languages, Offset for Display point coverage, which allows ranged-size point symbols placed in ocean or sparsely populated land areas to substitute for polygon symbology, has been substantially enhanced.

- Additional and updated sample maps illustrate a number of possibilities for using new population, widespread language, and name data.

Changes from 3.01 to 3.1

Adjustments of some language borders to align with DCW country borders (one or two languages of Russia, Oman, Saudi Arabia, Ethiopia, Djibouti, Sudan and some islands which were associated with wrong country).

Limitations

Polygons depict the traditional linguistic homelands of each language; populations away from their homelands (e.g. in cities, refugee populations, etc.) are not mapped. The World Language Mapping System does not attempt to map immigrant languages.

Polygons are often not provided for languages which are considered widespread in a country, e.g. where the language polygon would be essentially identical to the country boundary. In a few cases (e.g. French in Canada), languages have been classified as widespread on the basis of a strong official policy of nationwide multilingualism.

A very small number of languages are classified as "Unknown Location" and are not mapped as polygons pending further research.

Points corresponding to Widespread and Unknown Location languages are located arbitrarily within the appropriate country.

Language polygons were originally digitized relative to the World Databank II (WDBII) base map. Language polygons coincident with geographic features other than coastlines and international boundaries (particularly rivers) may follow the WDBII versions of these features.

Sample Map Projects

Sample map projects in ArcGIS project format (.mxd) are included in the *<installdir>*\projects\Samples-lang directory to illustrate use of the various language layers.

- **Africa Language Families.mxd – Language families of Africa.**

- **Kenya and Tanz Language.mxd – All languages of Kenya and Tanzania classified by language family. Languages labeled.**

- **Languages in Southern Africa.mxd – Language family map zoomed in to southern Africa region.**

- **Languages of Papua New Guinea.mxd – Papua New Guinea languages randomly colored.**

- **Myanmar and Thailand.mxd – All languages of Myanmar and Thailand, includes mixed language areas.**

- **Myanmar-Thailand-1_1mil_langonly.mxd – Some Myanmar and Thailand languages. Languages labeled.**

- **SE Papua New Guinea zoom to 1_5mill_langonly.mxd – Papua New Guinea map zoomed in to southeast.**

- **World Languages Point.mxd – Language points, classified by type.**

Additional projects illustrate the use of the data in combination with GMI's Seamless Digital Chart of the World product are included in *<installdir>*\projects\Samples-dcwlang. Note that the Seamless Digital Chart of the World needs to be installed for all feature classes to be present in these mxds.

- **Kurds.mxd – Locations of four Kurdish groups.**

- **Myanmar-Thailand-1_1mil_dcw.mxd – Some Myanmar and Thailand languages including additional DCW data for geographic context.**

- **SE Papua New Guinea zoom to 1_5mill dcw.mxd – Papua New Guinea map zoomed in to southeast including additional DCW data for geographic context.**

Data Organization

When uncompressed during installation, data will be placed in these subdirectories under a specified installation directory (*<installdir>*, by default c:\gmi_geodata). Click links below for detailed documentation of data in each subdirectory.

<installdir>\bnd – Boundary: Country boundaries

<installdir>\citation – Text for source notes on sample mxds

<installdir>\data – File geodatabase, WLMS.gdb containing all feature classes and tables in WLMS except rasters

<installdir>\lang – Language points, polygons, and attributes

<installdir>\docs – Documentation files (including this document)

<installdir>\topogrid – Raster topography grids which display elevation and hillshading. Due to the resolution of the datasets, they will look best on small scale maps (scales greater than 1:3,000,000).

<installdir>\projects -- Sample project files documented above

Detailed Data Documentation

Note: use Acrobat's Bookmarks pane or Find (Ctrl-F) features to find specific data in this section. Titles indicate point, line, or area (polygon) features. In the below descriptions, the shapefile location (.shp) is given first and then the file geodatabase location (e.g. data\WLMS.gdb\bnd\DCW_countries). The same fields in the shapefile and tables are also in the file geodatabase file. For polygon feature classes (shapefiles), the file geodatabase has Shape_Length and Shape_Area fields added.

Source: GMI and NIMA VMap0, Edition 5
Boundary
Country area features

bnd\DCW_countries.shp

data\WLMS.gdb\bnd\DCW_countries

Abstract: This feature class contains countries of the world and their names. It was derived by dissolving the updated first-level administrative divisions of GMI's DCW base map into countries. GMI has updated the countries to represent what is commonly considered the current political situation. Notes: Some data fields from the VPF form of polbnda have been changed to more intuitive names:

- NA2 changed to FIPS10_4
- NA2_DESCRI changed to COUNTRY
- NA3 changed to CONTCODE
- NA3_DESCRI changed to CONT

GENC (Geopolitical Entities, Names and Codes) 2 and 3 letter codes have been added in two new fields. FIPS codes are now called GEC codes. Topology has been created for this feature class to eliminate gaps and overlaps. The topology has changed some points, so the feature class will not exactly align with the original DCW.
Attributes:

Field	Field Name	Type	Width	Description
1	ID	Character	16	Unique identifier
2	FIPS10_4	Character	2	FIPS 10-4 (GEC) code
3	COUNTRY	Character	44	Country name
4	CONTCODE	Character	1	Continent/Region code
5	CONT	Character	15	Continent/Region name
6	F_CODE	Character	5	Feature code
7	F_CODE_DES	Character	19	Feature code description
8	OW_ABBREV	Character	4	Jason Mandryk, Operation World, Country Code
9	ISO_A2	Character	4	ISO 3166 2 letter country code
10	ISO_A3	Character	3	ISO 3166 3 letter country code
11	GENC_2	Character	2	Geopolitical Entities, Names, and Codes – 2 letter
12	GENC_3	Character	3	Geopolitical Entities, Names, and Codes – 3 letter
13	VMAP_ID	Character	15	Unique ID

Used in Projects: Africa Language Families.mxd, Kenya and Tanz Language.mxd, Kurds.mxd, Languages in Southern Africa.mxd, Languages of Papua New Guinea.mxd, Myanmar and Thailand.mxd, Myanmar-Thailand-1_1mil_langonly.mxd, SE Papua New Guinea zoom to 1_5mill_langonly.mxd, World Languages Point.mxd
Table Joins:

Source: GMI
Citation
Source notes for sample mxds

Citation\Basemap_note_text.shp

data\WLMS.gdb\citation\Basemap_note_text

Abstract: Text for source notes on sample mxds.
Note: Sample mxds have a dataframe where this feature class is present. The fields in the feature class are used for the text of the notes.
Attributes:

Field	Field Name	Type	Width	Description
1	OBJECTID	Number	9	Internal feature number
2	Shape	Geometry	0	Feature geometry.
3	MapID	Character	25	Name for note
3	Line1Label	Character	170	WLMS source statement
4	Line2Label	Character	170	SDCW source statement
5	Line3Label	Character	170	GMI Geodataset note
6	Line4Label	Character	50	GMI GMMS note

Field	Field Name	Type	Width	Description
7	TransLabel	Character	50	Translation Status note
8	GMMSnote	Character	50	GMMS source statement

Used in Projects: all sample mxds
Table Joins:

Source: World Language Mapping System

Language Areas

Language area features

lang\langa.shp

data\WLMS.gdb\lang\langa

Abstract: Languages areas of the world, updated to match the 2016 ISO 636-3 standard.
Note:
Attributes:

Field	Field Name	Type	Width	Description
1	ID	Character	16	Old style LMP identifier- Legacy ISO 639-3 with uppercase ISO A3 country code
2	ID_ISO_A3	Character	10	Unique identifier with ISO language code and ISO A3 country code
3	ID_ISO_A2	Character	7	Unique identifier with ISO language code and ISO A2 country code (preferred for most applications)
4	ID_FIPS	Character	7	Unique identifier with ISO language code and FIPS 10-4 (GEC) country code (preferred for historical studies due to code consistency over time)
5	NAM_LABEL	Character	40	Name in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
6	NAME_PROP	Character	48	Name in mixed-case, diacritic-stripped ASCII
7	NAME2	Character	40	Name in upper-case ASCII

8	NAM_ANSI	Character	40	Name in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
9	C1	Character	33	Country Name
10	OtherCntry	Character	254	Other countries or number of other countries where language is spoken.
11	LMP_CLASS	Character	5	Language mapping class: A=Point mapped, no polygon L=Polygon mapped; point location is polygon centroid U=Unknown location; point location based on location given in Ethnologue or arbitrary W=Widespread language (national or regional language of wider communication)
12	FAMILYPROP	Character	30	Language family; first element of language classification
13	FAMILY	Character	30	Language family; first element of language classification in upper case for compatibility with family classification in older maps
14	LMP_C1	Character	32	Country name used by WLMS (e.g. Kosovo used in WLMS but included with Serbia in source country code file)
15	LANG_ISO	Character	3	3 letter language ISO code

Used in Projects: Africa Language Families.mxd, Kenya and Tanz Language.mxd, Kurds.mxd, Languages in Southern Africa.mxd, Languages of Papua New Guinea.mxd, Myanmar and Thailand.mxd, Myanmar-Thailand-1_1mil_langonly.mxd, SE Papua New Guinea zoom to 1_5mill_langonly.mxd
Special layer files: lang family poly.lyr, lang family - Ruhlen (poly).lyr
Table Joins: ISO639_3toRuhlen.dbf

Mixed Language area features

lang\mixlanga.shp

data\WLMS.gdb\lang\\mixlanga

Abstract: Areas where languages overlap in langa.shp.
Notes: The Unique identifier (ID…) fields are created by two (if there are two language areas overlapping) or more (more than two language areas overlapping) groups consisting of a three letter ISO 639-3 language code and a hyphen. A country code (ISO 2 or 3 letter, or FIPS 10-4 (now called GEC)) is after the hyphen.
In the NAM_LABEL and NAME_PROP fields, the various language names are separated by a semi-colon (";"). When necessary, some language names have been abbreviated (or, as a last resort, omitted) for polygons with several long language names.
Attributes:

Field	Field Name	Type	Width	Description
1	ID	Character	32	Old style LMP identifier- uppercase ISO language codes with upper case ISO A3 country code
2	ID_ISO_A3	Character	32	ISO language codes with upper case ISO A3 country code.
3	ID_ISO_A2	Character	31	ISO language codes and ISO A2 country code.
4	ID_FIPS	Character	31	ISO language codes, with FIPS 10-4 (GEC) country code
5	NAM_LABEL	Character	100	Names of languages, semicolon separated, in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
6	NAME_PROP	Character	100	Names of languages, semicolon separated, in mixed-case ASCII
7	C1	Character	33	Country Name

Used in Projects: Myanmar and Thailand.mxd, Myanmar-Thailand-1_1mil_langonly.mxd
Table Joins:

Language Points
Standard Language Point Attributes

Notes: Unless otherwise indicated below, all language point feature classes have the following attributes.

Attributes:

Field	Field Name	Type	Width	Description
1	ID	Character	7	Old style LMP identifier- Legacy ISO 639-3 with uppercase ISO A3 country code
2	ID_ISO_A3	Character	7	Unique identifier with ISO language code and ISO A3 country code
3	ID_ISO_A2	Character	7	Unique identifier with ISO language code and ISO A2 country code (preferred for most applications)
4	ID_FIPS	Character	7	Unique identifier with ISO language code and FIPS 10-4 (now called GEC) country code
5	NAM_LABEL	Character	48	Name in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
6	NAME_PROP	Character	48	Name in mixed-case, diacritic-stripped ASCII
7	NAME2	Character	48	Name in upper-case ASCII
8	NAM_ANSI	Character	40	Name in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
9	C1	Character	33	Country Name
10	OtherCntry	Character	254	Other countries or number of other countries where language is spoken.
11	LMP_CLASS	Character	5	Language mapping class: A=Point mapped, no polygon

				L=Polygon mapped; point location is polygon centroid U=Unknown location; point location based on location given in Ethnologue or arbitrary W=Widespread language (national or regional language of wider communication)
12	FAMILYPROP	Character	30	Language family; first element of language classification
13	FAMILY	Character	30	Language family; first element of language classification in upper case for compatibility with family classification in older maps
14	LMP_C1	Character	32	Country name used by WLMS (e.g. Kosovo used in WLMS but included with Serbia in source country code file)
15	LANG_ISO	Character	3	3 letter language ISO code
16	LMP_SUBCL	Character	6	Language mapping subclass, indicating point reliability; available for some class U, W, and A points: N=No reliable location; point placed arbitrarily in corresponding country R=Regional reliability; point placed in general region suggested by location description in ISO 639-3 M=Most reliable location; sufficient data for point, but not for polygon

All Language Point Features

lang\lang_all_pt.shp

data\WLMS.gdb\lang\lang_all_pt

Abstract: Point features for all 2016 version of ISO 639-3 standard languages, including widespread and unknown location languages. Point is either centroid of polygon, placed near the center of the area where the language is located (if no polygon), or placed as best as possible with limited information. Where necessary, points have been placed arbitrarily within the correct country to create a complete set.
Notes:
Attributes: See Standard Language Point Attributes
Used in Projects: World Languages - Points.mxd
Special layer files: lang family pt.lyr, lang family - Ruhlen (pt).lyr
Table Joins: ISO639_3toRuhlen.dbf

Language point features

lang\lang_points_pt.shp

data\WLMS.gdb\lang\lang_points_pt

Abstract: Points within the language areas (langa.shp).
Attributes: See Standard Language Point Attributes
Table Joins:

Widespread Language point features

lang\lang_widesprd_pt.shp

data\WLMS.gdb\lang\lang_widesprd_pt

Abstract: National or regional trade languages which cannot be properly mapped as polygons since they overlap a large number of other language areas are represented by these points, which are placed arbitrarily within the correct country or region of a country.
Notes: File included for back compatibility with older maps; for many new applications, All Language Point Features with appropriate classification and definition queries will be easier to set up and produce more understandable legends
Attributes: See Standard Language Point Attributes
Table Joins:

Point-only point features

lang\lang_only_pt.shp

data\WLMS.gdb\lang\lang_only_pt

Abstract: Points of languages lacking an associated polygon.
Notes: File included for back compatibility with older maps; for many new applications, All Language Point Features with appropriate classification and definition queries will be easier to set up and produce more understandable legends
Attributes: See Standard Language Point Attributes
Table Joins:

Unknown Location point features

lang\lang_unkwn_loc_pt.shp

data\WLMS.gdb\lang\lang_unkwn_loc_pt

Abstract: These points represent ISO 639-3 standar languages for which insufficient data was available to create a language polygon in langa.shp. Where some textual description of the language location was given in the Ethnologue, the point was located in the approximate center of the area described. Where the text description is absent or unintelligible, the point is located arbitrarily within the country.
Notes: File included for back compatibility with older maps; for many new applications, All Language Point Features with appropriate classification and definition queries will be easier to set up and produce more understandable legends
Attributes: See Standard Language Point Attributes
Table Joins:

Language attribute data
Language Attributes

lang\eth_wlms.dbf

data\WLMS.gdb\eth_wlms

Note: These tables have been dropped due to proprietary and licensing issues. Available attributes from the 2016 ISO 639-3 standard are in lang_all_pt.shp. To get other data formerly in eth_wlms.dbf go to www.ethnologue.com/products.

Extinct Languages

data\WLMS.gdb\eth_Extinct   **2016**

Abstract: Extinct languages from the 2016 ISO 639-3 language code standard.
Notes: In WLMS 17, this table was named eth_DormantExtinct.This table replaces the point shapefile, lang_extinct_pt.shp which was in WLMS versions before 17. There has not been a lot of interest in a point database of extinct languages and also we didn't have a complete set of point locations. These have been tagged as extinct or very nearly extinct.
Attributes:

Field	Field Name	Type	Width	Decscription
1	ID_ISO_A3	Character	7	Unique identifier with ISO language code and ISO A3 country code
2	ID_ISO_A2	Character	7	Unique identifier with ISO language code and ISO A2 country code (preferred for most applications)
3	ID_FIPS	Character	7	Unique identifier with ISO language code and FIPS 10-4 (GEC) country code (preferred for

				historical studies due to code consistency over time)
4	ISO	Character	3	ISO language Code
5	NAME_PROP	Character	40	Name in mixed-case ASCII
6	NAM_LABEL	Character	40	Name in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
7	C1	Character	32	Country Name

Sign Languages

data\WLMS.gdb\eth_SignLanguages 2016

Abstract: Sign languages from the 2016 ISO 639-3 language code standard.
Notes: Due to the large number of sign languages, it was decided to create a table for them.
Attributes:

Field	Field Name	Type	Width	Decscription
1	ID_ISO_A3	Character	7	Unique identifier with ISO language code and ISO A3 country code
2	ID_ISO_A2	Character	7	Unique identifier with ISO language code and ISO A2 country code (preferred for most applications)
3	ID_FIPS	Character	7	Unique identifier with ISO language code and FIPS 10-4 (GEC) country code (preferred for historical studies due to code consistency over time)
4	ISO	Character	3	ISO language Code
5	NAME_PROP	Character	40	Name in mixed-case ASCII
6	NAM_LABEL	Character	40	Name in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
7	C1	Character	32	Country Name

Immigrant Languages

data\WLMS.gdb\eth_Immigrant 2016

Abstract: Immigrant languages from the 2016 ISO 639-3 language code standard.

Notes: An immigrant language is a language spoken by relatively recently arrived or transient populations which do not have a well-established, multi-generational community of language users in the country. This table replaces rother.shp and lang_other_pt.shp in versions of the WLMS previous to 17. It was decided to put these languages into a table since the ISO 639-3 language code standard didn't have points for many of the languages.

Attributes:

Field	Field Name	Type	Width	Decscription
1	ID_ISO_A3	Character	7	Unique identifier with ISO language code and ISO A3 country code
2	ID_ISO_A2	Character	7	Unique identifier with ISO language code and ISO A2 country code (preferred for most applications)
3	ID_FIPS	Character	7	Unique identifier with ISO language code and FIPS 10-4 (GEC) country code (preferred for historical studies due to code consistency over time)
4	ISO	Character	3	ISO language Code
5	NAME_PROP	Character	40	Name in mixed-case ASCII
6	NAM_LABEL	Character	40	Name in mixed-case, ANSI (ISO Latin 1) character set, inverted to move modifiers to beginning and remove commas (preferred for most map labeling purposes)
7	C1	Character	32	Country Name

Source: GMI/Dr. Merritt Ruhlen

Conversion for Language Families

Conversion between ISO639-3 and Ruhlen Families

lang\ISO639_3toRuhlen.dbf

data\WLMS.gdb\ISO639_3toRuhlen 2016

Abstract: This is a simple translation table developed by GMI providing an approximate equivalency between the highest level of the ISO 639-3 genetic tree genetic tree and the smaller number of top-level classifications in Dr.

Merritt Ruhlen's *A Guide to the World's Languages.* There are a small number of cases where this simple translation between top-level categories fails to produce the correct Ruhlen classification, most notably American Indian, Pacific Island languages, and Korean and Japanese.
Note: In previous WLMS version name of table was SILtoRuhlen.
Attributes:

```
Field  Field Name      Type       Width  Description
    1  FAMILY          Character    25    ISO 639-3 main
                                          language family
                                          (upper case)
    2  LANGFAMILY      Character    30    Corresponding
                                          Ruhlen language
                                          family
```

Source: GLOBE/GMI

Topography Grids
Global Elevation/Bathymetry Colors compressed raster

TopoGrid\elevation.ecw   **2003**

Abstract: Elevation.ecw is a highly-compressed, high-resolution raster image applying conventional elevation colors (shades of blue/cyan for ocean depths, dark green at low elevations, browns at mid elevations, and gray values lightening to white at high elevations). Due to the resolution of the data, it will look best on small scale maps (scales greater than 1:3,000,000).
Notes:
A. Two Digital Elevation Models were used:
> 1. Land elevations were produced from GMI's WorldDTM90 DEM. This was created as follows:
> a. We started with the GTOPO30 30 arc-second DEM (https://lta.cr.usgs.gov/GTOPO30) as assembled for latitudes above 60 degrees south and converted to Esri GRID format by Chris Deckert of Campus Crusade's Story Runner project.

> b. Additional elevation data for latitudes south of 60S were added from the GLOBE Project DEM (http://www.ngdc.noaa.gov/mgg/topo/globe.html).

> c. The Caspian Sea was converted to an ocean area by converting all cells at -28 meters elevation to NoData values. This is known to have added a few spurious NoData cells in land areas elsewhere having elevations of -28 meters, notably in areas north of the Caspian, near the Dead Sea, and in Death Valley, California.

> 2. Ocean depths (bathymetry) were produced with data from the ETOPO2 2-minute DEM (http://www.ngdc.noaa.gov/mgg/fliers/01mgg04.html).

Both DEMs were converted to ERMapper binary format, and colors from the ERMapper "elevation" color ramp were applied separately to land elevation and bathythmetric DEMs using an ERMapper algorithm file. The color ramps are using several linear segments to increase color distinctions at low land elevations and ocean depths, and reduce distinctions at very high elevations and extreme depths. The resulting ERMapper algorithm and binary data were compressed at full resolution

and 20:1 target compression using ERMapper 6.0 by Dawn Wise at San Diego State University.
Attributes: RGB color components
Used in Projects: Kurds.mxd

Global Hillshades compressed raster

TopoGrid\hillshades.ecw **2003**

Abstract: hillshades.ecw is a highly compressed, high-resolution grayscale hillshade raster image. The image was calculated using the hillshade algorithm of Esri ArcGIS spatial analyst applied to GMI's worlddtm90 grid dataset (derivation described under elevation.ecw above). Due to the resolution of the data, it will look best on small scale maps (scales greater than 1:3,000,000).
Notes:
A. Hillshades were calculated at a sun elevation of 60 degrees and vertical scale factor of .00005, chosen to avoid black saturation in all but the steepest mountainsides. The resulting grayscale grid was converted to ERMapper binary format and compressed using ERMapper 6.0 by Dawn Wise at San Diego State University.
B. The hillshade settings chosen for this image give good terrain definition in mountainous areas, but do not clearly distinguish terrain in areas of lower relief. This may be improved somewhat with the existing image by adjusting the maximum and minimum values of the color ramp. Greatly improved rendering of low relief requires recalculation with lower sun angles and higher vertical scale factors.
Attributes: RGB color components
Used in Projects: Kenya and Tanz Language.mxd, Kurds.mxd, Languages in South Africa.mxd, Languages of Papua New Guinea.mxd, SE Papua New Guinea.mxd, SE Papua New Guinea zoom to 1_5mill.mxd

Global Elevation/Bathymetry Colors+Hillshades compressed raster

TopoGrid\elev-hs.ecw **2003**

Abstract: This file is the hillshades.ecw file overlaying the elevation.ecw file at 70% transparency for use as a background in systems not supporting layer transparency.