HIS Registry of Language Varieties - Documentation

Date: 2022-12-14

Status: Adopted

Abstract: Documents the Registry of Language Varieties (ROLV) for the Harvest Information Standards (HIS). This registry contains a set of varieties of Living Languages.

Steward: Global Recordings Network

Editor: Allan Starling, (allanstarling@globalrecordings.net)

1. Overview

The function of the Registry of Language Varieties (ROLV) is to:

- Identify and verify specific varieties of given languages defined by ISO 639-3
- Provide unique, standardized codes for these varieties.

2. Contents of the Registry

The Registry contains a set of varieties of living languages. A code in this set represents a unique variety of a language.

By definition the scope of a language variety code is always a smaller group of speakers than the group represented by the assigned language as a whole.

Two codes are provided:

- ROLV Code. This is a standardized five-digit code (including leading zeros, when necessary) for uniquely referring to a particular variety.
- BCP-47 code. IETF Language Tags defined in BCP-47 identify not only the language variety, but also script and spelling conventions. This scheme enables users to tap into a wider set of language research and tools.

3. Changes in the Registry

Language varieties may be added or retired from the registry as a result of the following:

- The International Organization for Standards (ISO) has dropped, divided, merged, or retired an ISO code.
- The ISO has generated a new code.
- An item is determined to be a duplicate of another item.
- A new language variety has been identified

4. Verification of language-related data

Data can be verified as language varieties by any of the following:

- Has audio/video recordings.
- Has literature in any form of media, including Bible translations.
- Is described in the Ethnologue, but not just in a list of names.
- Is adequately described in Wikipedia, Joshua Project, Glottolog, Google Search, or similar references.
- Has a documented percentage intelligibility with another variety.
- An individual or organization working among the language variety has made a documented request for an ROLV code.
- Changes are reported by qualified field workers.

Language Varieties may have differences in any or all of vocabulary, grammatical construction, idioms, or marked accents. Differences may also be marked by religious or social prejudices.

5. Downloads

The ROLV consists of three lists:

5.1. Code List

Column	Format	Description
Language code	3 characters	The ISO code of the language of which this is a variety
Dialect code	5 digits	A unique identifier of this language variety
Language Tag	20 characters	A unique identifier of this language variety
Country code	2 characters	The ISO code of the country in which this variety is predominantly spoken
Variety name	75 characters	The name of the unique variety of this language
Language name	75 characters	The name of the language of which this is a variety
Location name	250 characters	The name of the location in which this variety is primarily spoken.

Download the Code List in Tab-delimited or JSON format.

5.2. Alternate Name List

Column	Format	Description
ROLV Code	5 digits	The unique identifier for the language variety
Language Tag	30 characters	A unique identifier of this language variety

Alternate Name	75 characters	An alternate name or spelling for the variety including names in other languages and scripts.

Download the Alternate Name List in Tab-delimited or JSON format.

5.3. Changes List

Column	Format	Description
Language variety Code	5 digits	The unique identifier for the language variety
Language Tag	30 characters	A unique identifier of this language variety
Date	yyyy-mm-dd	The change date
Change Type	1 character	A=Added – The code is newly created. M=Moved – The variety has been assigned to a different language. U=Updated – The definition of the dialect has extended contracted or changed in some other way. R=Retired – The dialect code should no longer be used
Previous Language Code	3 characters	The ISO code of the language to which the variety was previously assigned
Explanation	text	A more detailed description of what has changed and why.

Download the Changes List in Tab-delimited or JSON format.

6.0 Updates

A Request Form is provided on the GRN website for identifying new varieties and/or requesting a code.