

Geographic Names Standardization Policy for Tajikistan

United States Board on Geographic Names
Foreign Names Committee



June 2021

1. Introduction

This geographic names standardization policy has been prepared as an aid to those geographic names experts who are the working staff for the United States Board on Geographic Names (BGN) in the standardization of the geographic names of Tajikistan for United States Government use. The BGN and its staff work to effect consistent treatment of geographic name spellings in U.S. Government databases, publications, maps, and charts. The country policies are intended to satisfy, in part, the statutory requirements levied upon the BGN in Public Law USC 80-242 to develop principles, policies and procedures for geographic names standardization, and to promulgate decisions with respect to the principles of geographic nomenclature and orthography.

2. Languages and Language Policy

a. *Demographics*

Tajikistan is a mountainous, landlocked country and the smallest of the former Soviet Central Asian republics. Its population of approximately 8.9 million¹ persons is comprised of 80% ethnic Tajiks, 16.5% Uzbeks, with Russian, Kyrgyz, and other ethnic groups² making up the remainder.

b. *Languages*

The official language of Tajikistan is Tajik (ISO 639-3 *tgk*), an Indo-European language stemming from Persian and often considered its dialect. Tajik is sometimes referred to as Tajiki Persian or simply Tajiki. Russian (*rus*), although no longer an official language, remains a major language in the country as a legacy of Tajikistan's Soviet past and is still frequently used in business, informal, and inter-ethnic communication. The presence of minority ethnic communities from neighboring Central Asian states ensure that Kyrgyz, Pashto, Persian, and Uzbek languages are prevalent in Tajikistan as well. Parya (*paq*), Shughni (*sgn*), Wakhi (*wbl*), and Yaghobi (*yai*) languages are spoken in smaller pockets around the country.

c. *Geographic Names Standardization*

The Tajikistan Agency on Statistics (TAJSTAT) is the body responsible for geographic names standardization across the country and for other tasks such as demographic surveys. The agency publishes a periodical informational bulletin with information on development across the country's four main administrative regions and the city of Dushanbe.³ Tajikistan also participates in the United Nations Group of Experts on Geographic Names (UNGEGN) under the Asia, Southwest Division⁴.

¹ Шумораи Аъоли – Численность Населения, Page 9. Accessible at http://stat.www.tj/files/cislennost_naselenia_na_1.01.2018.pdf

² Update: Law of the Republic of Tajikistan NYU Hauser Global Law School Program. <https://www.nyulawglobal.org/globalex/Tajikistan1.html>

³ http://stat.www.tj/13b992a6b19499d948c865b3572cb4a1_1510580431.pdf

⁴ <https://unstats.un.org/unsd/geoinfo/UNGEGN/divisions.html#>

3. Toponymic Policies

a. *Script and/or Orthography*

Tajik is an Indo-European language closely related to Persian and considered by some to be a dialect of Persian. The Tajik language uses Cyrillic script, written left to right, and includes six characters not found in the Russian alphabet. Historically, Tajik was written in the Perso-Arabic script until the 1920s. The Soviet Union then introduced the Latin script for a short period from 1927 to 1930 as a means of promoting literacy, simplifying writing, better capturing vowel sounds, and distancing the population from Islamic influences in Central Asia⁵. The Cyrillic script was then introduced in the late 1930s and remains in use today, with some modifications to characters in 1998.

Following standard English lexical practice and BGN precedent, names and terms appearing in lists and tables henceforward are listed in English alphabetical order.

b. *Romanization*

BGN/PCGN developed the Romanization System of Tajik⁶ in 1994. The system was later updated with guidance from the Academy of Sciences of Tajikistan, which reformed script orthography in 1998 and designated four characters (ц, щ, ь, and ы) obsolete. Although Russian is not an official language in Tajikistan, because the Soviet Union produced many legacy Tajik map sources, they are in the Russian language and remain an important reference. Russian names collected for features in Tajikistan should be considered Variant names only and romanized in accordance with the 1974 BGN/PCGN Romanization System for Russian.

c. *Diacritics*

Romanized Tajik uses the following diacritical marks:

Cyrillic Character	Romanized Character	Character Name	Unicode Value ⁷
Ӣ	í	Latin capital letter I with acute	0205
ӣ	í	Latin lower-case letter i with acute	0237
Ӯ	Ŭ	Latin capital letter u with breve	0364

⁵ Enacademic.com <https://enacademic.com/dic.nsf/enwiki/2633724>

⁶ Romanization of Tajik, BGN/PCGN 1994. Available at <https://geonames.nga.mil/gns/html/Romanization/ROMANIZATION%20OF%20TAJIK.pdf>

⁷ Unicode values provided are Decimal format. To use on a Windows keyboard, hold the ALT key and enter in the characters listed in the Unicode Value column. For example, to enter in í, hold ALT + 0237

Cyrillic Character	Romanized Character	Character Name	Unicode Value ⁷
ӯ	ũ	Latin small letter u with breve	0365
’	’	Right single quotation mark	2019
’	’	Right single quotation mark	2019
Э	Ě	Latin capital letter e with dot above	0278
э	ě	Latin small letter e with dot above	0279

d. *Generic Terms*

A generic is a term that describes a feature. Examples include words such as ‘river,’ ‘hill,’ and ‘lake’. Some feature names have false generics—for example, a city named ‘Mount Vernon’ includes the generic term *mount*, but the name describes a populated place rather than a mountain or hill feature. Generics appearing in standardized name forms should be considered true generics, except in cases—like the above— in which the generic term does not match the feature type. Generics are not recorded separately for populated places in Tajikistan.

In Tajik, the generic term typically precedes the specific name, for example, in the feature *nohiyai Konibodom* (ноҳияи Конибодом), the term *nohiyai* (ноҳияи)— ‘district’— is the generic term. Generics are recorded in the indefinite form, so this term would be collected as *Nohiya* (Ноҳия). All generic terms in Appendix A are listed in the indefinite form. Please refer to section 3. i. *Unique Linguistic Situations* for more on the *izofa* construction used to modify the ending of the generic term.

Combined Generic Terms. Occasionally, the user will encounter approved names that combine the generic and the proper name. For example, in the feature *Roghsharvdara* (Роғшарвдара) *dara* (дара) is the generic term for ‘river.’ Provided that the term does not reference a false generic, the entire name is recorded as it appears on the source, and the generic term is also separately recorded in the Generic column. An example of a false generic in a combined name is a populated place named *Sarkokūl* (Саркоқӯл) since *kūl* (қӯл) is the generic term for ‘lake’.

With limited authoritative sources available in the Tajik language, the user may occasionally encounter a feature with a generic term only in Russian in the GeoNames Database. For example, *Ovrag Sarmych-Say*, *Ovrag Kok-Tash-Say*, or *Gora Sary-Sakh*. *Ovrag* is a Russian term for ravine, while *gora* is the Russian term for mountain: all names and generics encountered in Russian should be treated as Variants only unless ample evidence exists that the Russian generic term has been adopted into Tajik, for example in the cases of *kolkhoz* or *sovkhoz*.

e. *Hyphenation, Capitalization, and Spelling*

Hyphenated names that appear on official Tajik sources as part of the name are recorded as they appear. For instance, *Chagam-Kukabulok* or *Degi-Kham*. Frequently, the hyphenated name will be a close variant of the compound name, as in the case of *Sari-Varos* (Variant) and *Sarivaroz* (Approved), or *Shahr-e Now* (Variant) and *Shahrinav* (Approved).

The initial letter of romanized generics and hyphenated names should be capitalized when collected, even if such terms are not capitalized in official sources. However, when the hyphenation follows a primary noun as part of the *izofa* (изофа) construction— the ‘e’ sound used to connect words to show possession, modify nouns, or to connect two terms— the ‘e’ character is not separately capitalized. In the above example of *Shahr-e Now*, the modifier *-e* should not be capitalized when collected. Refer to section 3. i. *Unique Linguistic Situations* for more on the *izofa* construction.

f. Long and short forms

Short forms of geographic names in Tajikistan are collected for such features as country name (PCLI) and railroad features including stations (RSTN) and stops (RSTP). Railroad stations frequently appear on railroad atlases without generics.

Features with short names are recorded as follows: the full name is entered as an approved name with a Name Rank 2, and the short name is entered into the Short_Name field. The short name is also added as a separate approved name with Name_Rank 1.

For example, a railroad feature (RSTP) that appears near the populated place of *Regor* (Perop) is collected as *Regor* and given Name Rank 1; *Stantsiyai Regor* is also collected and given Name Rank 2. For this name, *Regor* is then recorded as the Short Name, and *Stantsiyai* is recorded as a generic term.

g. Numbers

If cardinal and/or ordinal numbers occur as integral parts in geographic names, they should be recorded as they appear on the most authoritative source.

Railroad stops (RSTP) and sidings (RSD) in Tajikistan frequently appear on maps as generic terms and a number which indicates distance in kilometers, particularly on Soviet-produced maps in Russian, for example *раз. 10 км*. These labels are considered map information and are not collected as geographic names.

h. Abbreviations

The following abbreviations and acronyms are commonly found on Tajik sources:

<i>Abbreviation on source</i>	<i>Expanded Term</i>	<i>Definition</i>	<i>Romanization</i>
аҶб.	аҶба	<i>mountain pass</i>	<i>aghba</i>

<i>Abbreviation on source</i>	<i>Expanded Term</i>	<i>Definition</i>	<i>Romanization</i>
айл.	айлок	<i>encampment, pasture</i>	<i>aylok</i>
чаш.	чашмеш	<i>spring</i>	<i>chashmesh</i>
фер.	ферма	<i>farm</i>	<i>ferma</i>
к.	куҳ	<i>mountain</i>	<i>kuh</i>
қат.	қаторкӯҳ	<i>mountain range</i>	<i>qatorkuh</i>
кул.	куллаи	<i>peak</i>	<i>qullai</i>
НТҚ	Ноҳияҳои Тобеи Ҷумҳурии	<i>Districts Under Republic Administration</i>	<i>Nohiyahoi Tobei Jumhuri</i>
об.	обанбор	<i>reservoir</i>	<i>obanbor</i>
ш.	шаҳр	<i>city</i>	<i>shahr</i>

i. Unique linguistic situations

Izofa Construction. Similar to Farsi and some Turkic languages, the Tajik language employs the *izofa* (изофа) construction. This is a short vowel sound that connects two words and can be used to show possession or to modify a noun. While in Farsi this sound is generally only pronounced and not written as a separate character when in Persian script, because Tajik is written in Cyrillic script and indicates more vowels than does Farsi or Dari, the *izofa* is expressed with *и* or *е* characters and romanized as *i* or *e*. The *izofa* may be written either jointly with the proper noun or in a hyphenated construction when it appears as such on the source.

For example, the term *aghba* (ағба, or mountain pass) uses the *izofa* in conjunction with the proper name *Zardolu* to become *Aghbai Zardolu*. Likewise, *nohiya* (ноҳия, or district) uses the construction in conjunction with the proper name *Shahritus* to become *Nohiyai Shahritus*.

4. Political Geography Policy

a. Country Name

Conventional Short Form

Tajikistan

Conventional Long Form	Republic of Tajikistan
Tajik Form – Non-Roman Script	Тоҷикистон
Tajik Short Form – Romanized	Tojikiston
Tajik Long Form – Non-Roman Script	Ҷумҳурии Тоҷикистон
Tajik Long Form – Romanized	Jumhurii Tojikiston

b. Capital Name

Tajik Form – Non-Roman Script	Душанбе
Tajik Form – Romanized	Dushanbe

c. First-Order Administrative Divisions

Tajikistan is divided into five first-level administrative divisions, including the capital district of Dushanbe and a conglomerate of 13 districts directly under federal administration -- Nohiyahoi Tobei Jumhuri (translated as Districts Under Republic Administration). With the exception of this conglomerate and the capital district, the other first-level administrative divisions are referred to as *Viloyat* (Вилоят) which translates to 'Province'.

In addition to the first-order administrative divisions outlined above, the BGN recognizes the following structure of lower-level administrative divisions in Tajikistan: nohiya (ноҳия) as second-level administrative divisions (ADM2/districts) and jamoat (ҷамоат) as third-level administrative divisions (ADM3/municipalities.)

<i>Name</i>	<i>GEC</i>	<i>GENC</i>	<i>Seat</i>
Dushanbe <i>Душанбе</i>	TI04	TJ-DU	Dushanbe <i>Душанбе</i>
Kūhistoni Badakhshon <i>Кӯҳистони Бадахшон</i>	TI01	TJ-GB	Khorugh <i>Хоруғ</i>
Sughd <i>Суғд</i>	TI03	TJ-SU	Khujand <i>Хуҷанд</i>
Khatlon <i>Хатлон</i>	TI02	TJ-KT	Bokhtar <i>Бохтар</i>
Nohiyahoi Tobei Jumhuri <i>Ноҳияҳои Тобеи Ҷумҳурии</i>	TI05	TJ-SU	n/a

d. Conventional and/or Anglicized Variant Names

The following Conventional and Anglicized Variant names are found in Tajikistan:

Conventional Names

<i>Name</i>	<i>UFI</i>	<i>UNI</i>	<i>Designation</i>
Amu Darya	-2575215	-3543885	STM
Fergana Valley	-2331758	-3218017	VAL
Tajikistan	-2557092	14076297	PCLI
Republic of Tajikistan	-2557092	-3521168	PCLI

Anglicized Variant Names

<i>Name</i>	<i>UFI</i>	<i>UNI</i>	<i>Designation</i>
Districts Under Republic Administration	1150323	13712773	ADM1
Gorno-Badakhshan	-2555718	-3519092	ADM1

e. Unique Geopolitical Situations

Tajikistan is a landlocked country and shares borders with Afghanistan, China, Kyrgyzstan, and Uzbekistan. Border demarcation and concerns over exclaves remain issues to monitor in Tajikistan.

Between 2011 and 2019, Tajikistan signed agreements with China and Uzbekistan to delineate disputed parts of their shared borders.^{8, 9}

Tajikistan has two exclaves within Kyrgyzstan—Vorukh (Sughd Province) and Kayragach (Khatlon Province), as well as one exclave within Uzbekistan—Sarvan (Sughd Province).¹⁰ Vorukh border demarcation has caused tension between Tajikistan and Kyrgyzstan.¹¹

For the latest country-specific boundary dispute information, consult the U.S. Department of State’s Office of the Geographer and Global Issues at internationalboundaries@state.gov

5. Source Material

Please note, native authoritative sources are extremely limited. Below are the preferred sources:

Agentii Tojikoinot, First Order Administrative Division Map series, scaled 1:300,000 -1:600,000 2005-2006

Kharitasoz (Jumhuri Tojikistan Geodezia va Kharitazosi) National Map, “Jumhurii Tojikiston”

⁸ Uzbekistan, Tajikistan to Sign Border Demarcation Agreement, 2 March 2018. Caravanserai (https://central.asia-news.com/en_GB/articles/cnmi_ca/newsbriefs/2018/03/02/newsbrief-01) Accessed 28 May 2020

⁹ China, Tajikistan Sign Border Agreement. 13 January 2011 The Hindu (<https://www.thehindu.com/news/international/China-Tajikistan-sign-border-agreement/article15517984.ece>) Accessed 28 May 2020

¹⁰ Gabdulhakov, Rashid. Geographical Enclaves of the Ferghana Valley: Do Good Fences Make Good Neighbors? http://osce-academy.net/upload/Policy_briefs/Policy_Brief_14.pdf

¹¹ Small Exclave Spells Big Problems for Kyrgyzstan, Tajikistan, 16 January 2014. RadioFreeEurope (<https://www.rferl.org/a/kyrgyzstan-tajikistan-exclaves-vorukh-tensions/25232311.html>) Accessed 28 May 2020

Appendix A. Glossary of Generic Terms

The following are common generics the user may encounter on Tajik maps. Note that this chart lists generics in the indefinite form, although they may appear in a different form on the native source. Please refer to section 3.d. *Generic Terms* for an explanation of this.

Generic (Romanized)	Generic (Native Script)	Feature Designation Name	Feature Designation Code
aghba	ағба	mountain pass	PASS
aĭlok	айлок	encampment, pasture	CMP
chashmesh	чашмеш	spring	SPNG
dara	дара	stream	STM
ferma	ферма	farm	FRM
kharoba	хароба	ruins	RUIN
kūh	кӯҳ	mountain, hill	MT, HLL
kūl	кӯл	lake, salt lake	LK, LKN
obanbor	обанбор	reservoir	RSV
piryakh	пирях	glacier	GLCR
qatorkūh	қаторкӯҳ	mountains, mountain range	MTS
qulla	қулла	peak	PK
shahr	шаҳр	city	PPL ¹²

¹² Generic terms are not separately recorded for PPLs