

Language support

Neural Machine Translation model

The Translation API's recognition engine supports a wide variety of languages for the Neural Machine Translation (NMT) model.

These languages are specified within a recognition request using language code parameters as noted on this page. Most language code parameters conform to [ISO-639](https://wikipedia.org/wiki/ISO_639) (https://wikipedia.org/wiki/ISO_639) identifiers, except where noted.

Translations from any language to any language in this list are supported. [Romanization and transliteration](#) (`#roman`) are supported only on the Cloud Translation - Advanced API. The list is updated as new languages are added. You can also list the supported languages by using [Cloud Translation - Basic](#) (/translate/docs/basic/discovering-supported-languages) or [Cloud Translation - Advanced](#) (/translate/docs/advanced/discovering-supported-languages-v3) APIs.

Try it for yourself

If you're new to Google Cloud, create an account to evaluate how Cloud Translation performs in real-world scenarios. New customers also get \$300 in free credits to run, test, and deploy workloads.

Try Cloud Translation free (<https://console.cloud.google.com/freetrial>)

Language	ISO-639 code
Abkhaz	ab
Acehnese	ace
Acholi	ach
Afrikaans	af
Albanian	sq
Alur	alz

Language	ISO-639 code
Amharic	am
Arabic	ar
Armenian	hy
Assamese	as
Awadhi	awa
Aymara	ay
Azerbaijani	az
Balinese	ban
Bambara	bm
Bashkir	ba
Basque	eu
Batak Karo	btx
Batak Simalungun	bts
Batak Toba	bbc
Belarusian	be
Bemba	bem
Bengali	bn
Betawi	bew
Bhojpuri	bho
Bikol	bik
Bosnian	bs
Breton	br
Bulgarian	bg
Buryat	bua
Cantonese	yue

Language	ISO-639 code
Catalan	ca
Cebuano	ceb
Chichewa (Nyanja)	ny
Chinese (Simplified)	zh-CN or zh (BCP-47 (https://tools.ietf.org/html/bcp47))
Chinese (Traditional)	zh-TW (BCP-47 (https://tools.ietf.org/html/bcp47))
Chuvash	cv
Corsican	co
Crimean Tatar	crh
Croatian	hr
Czech	cs
Danish	da
Dinka	din
Divehi	dv
Dogri	doi
Dombe	dov
Dutch	nl
Dzongkha	dz
English	en
Esperanto	eo
Estonian	et
Ewe	ee
Fijian	fj
Filipino (Tagalog)	fil or tl
Finnish	fi
French	fr

Language	ISO-639 code
French (French)	fr-FR
French (Canadian)	fr-CA
Frisian	fy
Fulfulde	ff
Ga	gaa
Galician	gl
Ganda (Luganda)	lg
Georgian	ka
German	de
Greek	el
Guarani	gn
Gujarati	gu
Haitian Creole	ht
Hakha Chin	cnh
Hausa	ha
Hawaiian	haw
Hebrew	iw or he
Hiligaynon	hil
Hindi	hi
Hmong	hmn
Hungarian	hu
Hunsrik	hrx
Icelandic	is
Igbo	ig
Iloko	ilo

Language	ISO-639 code
Indonesian	id
Irish	ga
Italian	it
Japanese	ja
Javanese	jw or jv
Kannada	kn
Kapampangan	pam
Kazakh	kk
Khmer	km
Kiga	cgg
Kinyarwanda	rw
Kituba	ktu
Konkani	gom
Korean	ko
Krio	kri
Kurdish (Kurmanji)	ku
Kurdish (Sorani)	ckb
Kyrgyz	ky
Lao	lo
Latgalian	ltg
Latin	la
Latvian	lv
Ligurian	lij
Limburgan	li
Lingala	ln

Language	ISO-639 code
Lithuanian	lt
Lombard	lmo
Luo	luo
Luxembourgish	lb
Macedonian	mk
Maithili	mai
Makassar	mak
Malagasy	mg
Malay	ms
Malay (Jawi)	ms-Arab
Malayalam	ml
Maltese	mt
Maori	mi
Marathi	mr
Meadow Mari	chm
Meiteilon (Manipuri)	mni-Mtei
Minang	min
Mizo	lus
Mongolian	mn
Myanmar (Burmese)	my
Ndebele (South)	nr
Nepalbhasa (Newari)	new
Nepali	ne
Northern Sotho (Sepedi)	nso
Norwegian	no

Language	ISO-639 code
Nuer	nus
Occitan	oc
Odia (Oriya)	or
Oromo	om
Pangasinan	pag
Papiamentu	pap
Pashto	ps
Persian	fa
Polish	pl
Portuguese	pt
Portuguese (Portugal)	pt-PT
Portuguese (Brazil)	pt-BR
Punjabi	pa
Punjabi (Shahmukhi)	pa-Arab
Quechua	qu
Romani	rom
Romanian	ro
Rundi	rn
Russian	ru
Samoan	sm
Sango	sg
Sanskrit	sa
Scots Gaelic	gd
Serbian	sr
Sesotho	st

Language	ISO-639 code
Seychellois Creole	crs
Shan	shn
Shona	sn
Sicilian	scn
Silesian	szl
Sindhi	sd
Sinhala (Sinhalese)	si
Slovak	sk
Slovenian	sl
Somali	so
Spanish	es
Sundanese	su
Swahili	sw
Swati	ss
Swedish	sv
Tajik	tg
Tamil	ta
Tatar	tt
Telugu	te
Tetum	tet
Thai	th
Tigrinya	ti
Tsonga	ts
Tswana	tn
Turkish	tr

Language	ISO-639 code
Turkmen	tk
Twi (Akan)	ak
Ukrainian	uk
Urdu	ur
Uyghur	ug
Uzbek	uz
Vietnamese	vi
Welsh	cy
Xhosa	xh
Yiddish	yi
Yoruba	yo
Yucatec Maya	yua
Zulu	zu

Romanization and transliteration support

Preview

This feature is subject to the "Pre-GA Offerings Terms" in the General Service Terms section of the [Service Specific Terms \(/terms/service-terms#1\)](/terms/service-terms#1). Pre-GA features are available "as is" and might have limited support. For more information, see the [launch stage descriptions \(/products#product-launch-stages\)](/products#product-launch-stages).

[Romanization \(/translate/docs/advanced/romanize-text\)](/translate/docs/advanced/romanize-text) and [transliteration \(/translate/docs/advanced/translating-text-v3#transliteration\)](/translate/docs/advanced/translating-text-v3#transliteration) support the following languages.

Language	ISO-639 code	Romanization	Transliteration
Arabic	ar	✓	✓
Amharic	am	✓	✗

Language	ISO-639 code	Romanization	Transliteration
Bengali	bn	✓	✓
Belarusian	be	✓	✗
Gujarati	gu	✗	✓
Hindi	hi	✓	✓
Japanese	ja	✓	✓
Kannada	kn	✗	✓
Myanmar	uk	✓	✗
Russian	ru	✓	✓
Serbian	sr	✓	✗
Tamil	ta	✗	✓
Telugu	te	✗	✓
Ukrainian	uk	✓	✗

Custom models

For Cloud Translation - Advanced, AutoML Translation supports the following language pairs for custom models. A supported language means that Google has an existing neural machine translation (NMT) model for that language, which AutoML Translation uses as a base to create a custom model during training.

Note: Same language translation—from en-US to en-GB, for example—is not supported.

Language Pair	Language Codes
Afrikaans <-> English	af <-> en
Albanian <-> English	sq <-> en
Arabic <-> English	ar <-> en
Azerbaijani <-> English	az <-> en

Language Pair	Language Codes
Bengali <-> English	bn <-> en
Bulgarian <-> English	bg <-> en
Catalan <-> English	ca <-> en
Chinese (Simplified) <-> English	zh-CN * <-> en
Chinese (Traditional) <-> English	zh-TW <-> en
Croatian <-> English	hr <-> en
Czech <-> English	cs <-> en
Danish <-> English	da <-> en
Dutch <-> English	nl <-> en
Estonian <-> English	et <-> en
Finnish <-> English	fi <-> en
French <-> English	fr <-> en
Galician <-> English	gl <-> en
Georgian <-> English	ka <-> en
German <-> English	de <-> en
Greek <-> English	el <-> en
Gujarati <-> English	gu <-> en
Haitian Creole <-> English	ht <-> en
Hebrew <-> English	iw <-> en
Hindi <-> English	hi <-> en
Hungarian <-> English	hu <-> en
Icelandic <-> English	is <-> en
Indonesian <-> English	id <-> en
Italian <-> English	it <-> en
Japanese <-> English	ja <-> en

Language Pair	Language Codes
Korean <-> English	ko <-> en
Latvian <-> English	lv <-> en
Lithuanian <-> English	lt <-> en
Malay <-> English	ms <-> en
Marathi <-> English	mr <-> en
Norwegian <-> English	no <-> en
Persian <-> English	fa <-> en
Polish <-> English	pl <-> en
Portuguese <-> English	pt <-> en
Punjabi <-> English	pa <-> en
Romanian <-> English	ro <-> en
Russian <-> English	ru <-> en
Serbian <-> English	sr <-> en
Slovak <-> English	sk <-> en
Slovenian <-> English	sl <-> en
Spanish <-> English	es <-> en
Swahili <-> English	sw <-> en
Swedish <-> English	sv <-> en
Thai <-> English	th <-> en
Turkish <-> English	tr <-> en
Ukrainian <-> English	uk <-> en
Urdu <-> English	ur <-> en
Vietnamese <-> English	vi <-> en
Welsh <-> English	cy <-> en

* Simplified Chinese can be specified either by zh-CN or zh.

Supported codes for language variants

The following language codes are variants of the supported languages in the previous table. You can use these codes as the source or target language when you create datasets.

Google doesn't have base NMT models for these languages. Instead, Cloud Translation uses the language variant's associated base model for training custom models.

These codes are useful, for example, when you translate content for a particular dialect or region. For example, suppose you have localized data for zh-HK that you create a custom model for. When you perform translations, you can specify the zh-HK language code, which points to your custom model and produces more accurate translations for that locale.

The following table lists the language codes, their descriptions, and their associated base models that Cloud Translation uses when training custom models.

Language code	Description	Base model
zh-HK	Hong Kong (Traditional)	zh-TW

Translation LLM supported languages

The following languages pairs are supported for text translations that use the translation LLM and for adaptive translations.

Note: If you require a language that isn't in the following list, there are additional language pairs available in private preview. To access languages in private preview, submit a [request](https://docs.google.com/forms/d/1-iCtnDJ7xeXj8ssSkXGvGNuR3g1Dh7DqVo0I6xo6FZ8) (<https://docs.google.com/forms/d/1-iCtnDJ7xeXj8ssSkXGvGNuR3g1Dh7DqVo0I6xo6FZ8>).

Language pair	Language codes
Arabic <-> English	ar <-> en
Chinese (Simplified) <-> English	zh-CN <-> en
Czech <-> English	cs <-> en
Dutch <-> English	nl <-> en
French <-> English	fr <-> en

Language pair	Language codes
German <-> English	de <-> en
Hindi <-> English	hi <-> en
Indonesian <-> English	id <-> en
Italian <-> English	it <-> en
Japanese <-> English	ja <-> en
Korean <-> English	ko <-> en
Polish <-> English	pl <-> en
Portuguese <-> English	pt <-> en
Russian <-> English	ru <-> en
Spanish <-> English	es <-> en
Thai <-> English	th <-> en
Turkish <-> English	tr <-> en
Ukrainian <-> English	uk <-> en
Vietnamese <-> English	vi <-> en

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2025-02-14 UTC.