

[Attributes:Translate vs Transliterate](#)

[Libraries used:](#)

[To be Translated:](#)

[To be Transliterated:](#)

[Attributes: Limited values vs Remaining](#)

[Limited value attrs:](#)

[Remaining attrs:](#)

[Categories](#)

[Details of Edge Cases](#)

Attributes:Translate vs Transliterate

Libraries used:

- Translation - BingTranslator
- Transliteration - deepTranslit

To be Translated:

- storyline,tagline, trivia, synopsis

To be Transliterated:

- all the others

Attributes: Limited values vs Remaining

Limited value attrs:

- duration, color, countries, languages, genre, gross, release, rated

Remaining attrs:

- all the others

Categories

- Genre సినిమాలు
- Release_year సినిమాలు

- Rated సినిమాలు
- Director దర్శకత్వం వహించిన సినిమాలు
- Writer రచన అందించిన సినిమాలు
- Composer సంగీత దర్శకత్వం వహించిన సినిమాలు
- Film_editor
- Cinematography చిత్రీకరించిన సినిమాలు
- color సినిమాలు
- countries నుండి సినిమాలు
- languages సినిమాలు
- production నిర్మాణం చేసిన సినిమాలు

Details of Edge Cases

This table contains information about edge cases observed apart from Missing values

Attribute	Edge Cases Observed	Missing Values
IMDb Id	None	0
Title	None	0
Release_Year	None	0
Duration	None	0
Rating	None	0
Rated	Can merge (unrated, not rated, approved, passed)	184
Genre	None	0
Director	None	0
Synopsis	None	62
Votes	None	0
gross	None	2009
Songs	None	974
Poster	None	0
Countries	None	2
Languages	None	7
Release	All the values do not contain all day,month and year,	2

	should be careful while wording in telugu	
Film Locations	Commas separate different countries and also address in the same country (careful use of 'and' and ',')	449
Budget	None	2323
Opening Weekend	None	2320

Cumulative	None	933
Production Company	None	12
Sound Mix	None	933
Colors	None	25
Images	None	2
Writers	None	3
Cast	None	3
Storyline	None	2
tagline	None	1170
Trivia	None	79
Producers	None	53
Composer	None	461
Cinematographer	None	253
Film editor	None	120
Casting by	None	1675
Production design	None	1229
Art design	None	1625
Set decoration	None	2349
Nominee	Remove awards which are not notable, and limit words	1337
Winner	Remove awards which are not notable, and limit words	2691

Narrative Location	None	5449
Distributed by	None	4896
Distribution format	None	5188
Part of series	None	8277
Main Subject	None	7487
stars	None	0
Based on	None	8211

Total list of attributes that need to have language checkup/ edit: [29]

Attribute value type:

- No Symbols [25] - *These attributes just have either single values or comma (,) separated values [for differentiation] but no braces*
- Have symbols [4] - *These attributes have symbols like ([] { } : ;) which are important for code and should not be removed.*

Attribute	Attribute value type	Transliterated or translated
Title	No symbols	transliterated
Director	No symbols	transliterated
Synopsis	No symbols	translated
Countries	No symbols	transliterated
Languages	No symbols	transliterated
Film Locations	No symbols	transliterated
Songs	Have Symbols	transliterated
Production Company	No symbols	transliterated
Writers	No symbols	transliterated
Cast	Have Symbols	transliterated
Storyline	No symbols	translated
Tagline	No symbols	translated
Trivia	No symbols	translated

Producer	No symbols	transliterated
Composer	No symbols	transliterated
Cinematographer	No symbols	transliterated
Film Editor	No symbols	transliterated
Casting	No symbols	transliterated
Production Design	No symbols	transliterated
Art Design	No symbols	transliterated
Set decoration	No symbols	transliterated
Winner	Have Symbols	transliterated
Nominee	Have Symbols	transliterated
Narrative Location	No symbols	transliterated
Distributed By	No symbols	transliterated
Main Subject	No symbols	transliterated
Part Of series	No symbols	transliterated
Based on	No symbols	transliterated
Stars	No symbols	transliterated