

Person Names in the Real World

Handling Real Names with CLDR Person Name formats

Michael McKenna

2023-11-07



What are the Person Name guidelines?

Unicode Technical Report #35 LDML

<https://www.unicode.org/reports/tr35/tr35-personNames.html>

Part 8, Person Names

- Formatting for person names, such as
 - ⇒ John Smith
 - ⇒ 宮崎駿
- Patterns to format name objects (for example, from a database)
 - ⇒ Formats for particular locales
- Name data has fields for the parts of people's names
 - ⇒ Example: given : "*Maria*", surname : "*Schmidt*".

Basics: data - fields

Field	Examples
title	Ms., Mr., Prof., Dr.
given	Sally, Mohan, Johnathan, Anne Marie
given-informal	Sal, Mo, John, Anne
given2	Peter, Roland, S.
surname	Schmidt, Ramamurthy, Gonzales
surname2	Rivera, Garcia
generation	Jr., Sr., III
credentials	MBA, Esq., R.N., P.E., PhD

Basics: formats - length

length	examples
long	Mr. Charles Babbage, Esq. F.R.S. Mrs. Augusta Ada Byron King, Countess of Lovelace Sir John Frederick William Herschel, Esq. Sec. R. S. 李振藩 أَبُو الْوَفَاءِ مُحَمَّدُ بْنُ مُحَمَّدِ بْنِ يَحْيَى بْنِ إِسْمَاعِيلَ بْنِ الْعَبَّاسِ الْبُورْجَانِي
medium	Charles Babbage Ada Lovelace John F. W. Herschel 李振藩 أبو الوفاء البورجاني
short	Charles Ada John 李 أبو الْوَفَاءِ

Basics: formats - usage

usage	examples
referring	Charles Babbage, Esq. F.R.S. Augusta Ada Lovelace Sir John F. W. Herschel, Esq. Sec. R. S. 李振藩 أبو الوفاء البوزجاني
addressing	Mr. Babbage Mrs. Lovelace Sir Herschel 李振藩 أبو الوفاء البوزجاني
monogram	CB AAL JFWH 李 أ.أ.

Basics: formats - formality

formality	examples
formal	Mr. Charles Babbage, Esq. F.R.S. Mrs. Augusta Ada Byron King, Countess of Lovelace Sir John Frederick William Herschel, Esq. Sec. R. S. 李振藩 أَبُو الْوَفَاءِ مُحَمَّدُ بْنُ مُحَمَّدٍ بْنِ يَحْيَى بْنِ إِسْمَاعِيلِ بْنِ الْعَبَّاسِ الْبُورْجَانِي
informal	Charles Ada John 振藩 أبو

Basics: formats - order (default from locale)

order	examples
givenFirst	Charles Babbage Ada Lovelace John F. W. Herschel 振藩李 أبو الوفاء البوزجاني
surnameFirst	Babbage Charles Lovelace Ada Herschel John F. W. 李振藩 البوزجاني أبو الوفاء
sorting	Babbage, Charles Lovelace, Ada Herschel, John F. W. 李振藩 البوزجاني, أبو الوفاء

Locale defaults

- length
- formality
- NameOrder
- separators
- patterns for initials

NameOrder For Locales	en	ja	zh
givenFirst	und en	und ☆	und en ☆
surnameFirst	ja ko vi yue zh	hu ja km ko m vi yue zh ☆	ja ko vi yue zh ☆
Default Parameters			
formality	informal	formal ☆	informal ☆
length	medium	medium ☆	medium ☆
AuxiliaryItems			
foreignSpaceReplacement		· ☆	· ☆
nativeSpaceReplacement		<EMPTY> ☆	<EMPTY> ☆
pattern-initial	{0}.	{0} ☆	{0} ☆
pattern-initialSequence	{0}{1}	{0} {1} ☆	{0} {1} ☆

Modifier: data -informal

Used to indicate that the value in a field may be used in informal cases
Only applied to *given* field to date.

given-informal

```
{ given: “Johnathan”, given-informal: “John”, surname: “Smith” }
```

```
{ given: “Александр”, given-informal: “Са́ша”, given2: “Алекса́ндрович”,  
  surname: “Серебро́в” }
```

Medium-addressing-informal:

en: *John*

ru: *Са́ша*

medium-referring-formal

en: *Jonathan Smith*

ru: *Александр Алекса́ндрович Серебро́в*

Modifier: data -prefix & -core

Used when the surname may be composed of prefix values that format differently in different contexts, e.g. Dutch

*surname-**prefix** surname-**core***

{ given: "*Daniella*", given2: "Helena:", surname-**prefix**: "*van den*",
surname-**core**: "*Berg*" }

Usage: Sorting, medium-formal:

Berg, Daniella H. van den

CLDR SurveyTool

<https://st.unicode.org/cldr-apps/v#/en/PersonNameFormats/>

PersonName Main Patterns For Order: GivenFirst

long-referring-formal {title} {given} {given2} {surname} {generation},
{credentials} (e)

long-addressing-formal {title} {surname} (e)

long-referring-informal {given-informal} {surname} (e)

Native name and script:

Zendaya

Irene Adler

Mary Sue Watson

Bertie Wooster

Foreign name and native script:

Sinbad

Käthe Müller

Zäzilia Stöber

Neele von Brühl

Foreign name and script:

Маруся Петренко

けいこ佐藤さん

{given-informal} {surname}

Native name and script:

Zendaya

I. Adler

M.S.H. Watson

B.W.H.R. Wooster

Foreign name and native script:

Sinbad

K. Müller

Z.H. Stöber

A.C.C.M. von Brühl

Foreign name and script:

проф. А. М. Б. З. Петренко доц.

恵子佐藤さん

{given-initial}{given2-initial} {surname}

addressing-formal {title} {surname} (e)

referring-informal {given-informal} {surname-initial} (e)

addressing- {given-informal} (e)

Example: Passport

ICAO Doc-9303 - Machine Readable Travel Document standard

Only *TWO* fields : Primary identifier and Secondary identifier

Surname prefixes in lower case (von, de, etc.)



surname surname2

given given2

NOT a supported CLDR format

Example: Preferred Name

```
{  personID: 1234,  
  nameType: "legal",  
  title: "",  
  given: "Jun-fan",  
  given2: "",  
  surname: "Lee",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "en-US"  
}
```

```
{  personID: 1234,  
  nameType: "preferred",  
  title: "Mr.",  
  given: "Bruce",  
  given2: "",  
  surname: "Lee",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "en-US"  
}
```



Note: “personID” and “nameType” are not part of CLDR Person Name. It is included here to illustrate different object types for the same name.

Example: Preferred Name

```
{  personID: 12367,  
  nameType: "birth",  
  title: "",  
  given: "Pablo",  
  given2: "Diego José Francisco de  
Paula Juan Nepomuceno María de los  
Remedios Cipriano de la Santísima Trinidad",  
  surname: "Ruiz y",  
  surname2: "Picasso",  
  generation: "",  
  credentials: "",  
  nameLocale: "es-ES"  
}
```

```
{  personID: 12367,  
  nameType: "preferred",  
  title: "",  
  given: "Pablo",  
  given2: "",  
  surname: "Ruiz",  
  surname2: "Picasso",  
  generation: "",  
  credentials: "",  
  nameLocale: "es-ES"  
}
```



Example: Preferred Name

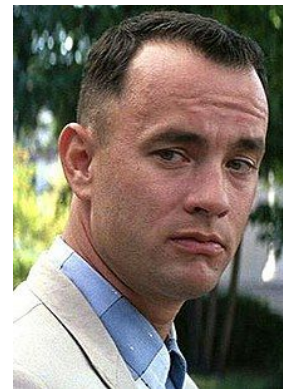
```
{  personID: 4319,  
  nameType: "legal",  
  title: "",  
  given: "Henriette",  
  given2: "Theodora",  
  surname: "Markovitch",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "fr-FR"  
}
```

```
{  personID: 4319,  
  nameType: "preferred",  
  title: "Ms.",  
  given: "Dora",  
  given2: "",  
  surname: "Maar",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "fr-FR"  
}
```



Example: Informal Name

```
{  personID: 1897,  
  nameType: "legal",  
  title: "",  
  given: "Forest",  
  given-informal: "Bubba",  
  given2: "Alexander",  
  surname: "Gump",  
  surname2: "",  
  generation: "Sr.",  
  credentials: "",  
  nameLocale: "en-US"  
}
```



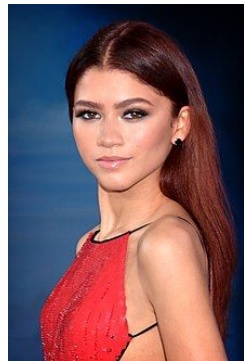
medium-referring-formal
Forest A. Gump Sr.

medium-referring-informal
Bubba Gump

Example: mononyms

```
{  personID: 4207,  
  nameType: "legal",  
  title: "Ms.",  
  given: "Zendaya",  
  given2: "Maree Stoemer",  
  surname: "Coleman",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "en-US"  
}
```

```
{  personID: 4207,  
  nameType: "preferred",  
  title: "",  
  given: "Zendaya",  
  given2: "",  
  surname: "",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "en-US"  
}
```



Example: Indonesian full names

```
{  personID: 6543,  
  nameType: "legal",  
  title: "",  
  given: "Hasan Prasetyo",  
  given2: "",  
  surname: "",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "id-ID"  
}
```



Deep Dive: "I go by my second name"



```
{  personID: 2468,  
  nameType: "legal",  
  title: "Mrs.",  
  given: "Augusta",  
  given2: "Ada Byron",  
  surname: "King",  
  surname2: "",  
  generation: "",  
  credentials: "Countess of Lovelace",  
  nameLocale: "en-GB"  
}
```

```
{  personID: 2468,  
  nameType: "preferred",  
  title: "Mrs.",  
  given: "Ada",  
  given2: "",  
  surname: "Lovelace",  
  surname2: "",  
  generation: "",  
  credentials: "",  
  nameLocale: "en-GB"  
}
```

Deep Dive: Handling multiple scripts



```
{  personID: 1234,
  nameType: "preferred",
  title: "Mr.",
  given: "Bruce",
  surname: "Lee",
  generation: "",
  credentials: "",
  nameLocale: "en-US"
}
```

```
{  personID: 1234,
  nameType: "preferred",
  title: "",
  given: "小龍",
  surname: "李",
  generation: "",
  credentials: "",
  nameLocale: "zh-Hant"
}
```

```
{  personID: 1234,
  nameType: "pinyin",
  title: "",
  given: "Siu-lung",
  surname: "Lei",
  generation: "",
  credentials: "",
  nameLocale: "zh-Latn"
}
```

Deep Dive: Handling multiple scripts



```
{  personID: 98765,  
  nameType: "romaji",  
  title: "",  
  given: "Hayao",  
  surname: "Miyazaki",  
  generation: "",  
  credentials: "",  
  nameLocale: "ja-Latn"  
}
```

```
{  personID: 98765,  
  nameType: "preferred",  
  title: "",  
  given: "駿",  
  surname: "宮崎",  
  generation: "",  
  credentials: "",  
  nameLocale: "ja-JP"  
}
```

```
{  personID: 98765,  
  nameType: "kana",  
  title: "",  
  given: "はやお",  
  surname: "みやざき",  
  generation: "",  
  credentials: "",  
  nameLocale: "ja-Hira"  
}
```

Deep Dive: Handling declensions



Nominative (referring)

Vocative (**-vocative** modifier), used in *addressing* cases)

Genitive (**-genitive** modifier), with *referring* for possessive case with title

Example in Latvian: *Kārlis Ulmanis* (past president of Latvia)

```
{  title: "kungs",
   given: "Kārlis",
     given-vocative: "Kārli",
   given2: "Augusts Vilhelms",
  surname: "Ulmanis",
     surname-genitive: "Ulmanņa", surname-vocative: "Ulmani",
  nameLocale: "lv-LV",
  credentials: "PhD"
}
```

Deep Dive: Handling declensions

Nominative (referring)

Vocative (**-vocative** addressing)

Genitive (**-genitive** variant) possessive case

	lv-LV	format
nominative (referring)	<i>Kārlis Ulmaņa kungs</i>	{given} {surname-genitive} {title}
-vocative, addressing	Sveikas Kār li !	{given}

Deep Dive: Handling declensions

Nominative (referring)

Vocative (**-vocative** modifier), used in *addressing* cases)

Genitive (**-genitive** modifier), used for possessive case with *referring*

Example in Lithuanian: [Rolandas Gimbutis](#) (olympic swimmer)

```
{  title: "P.",
  given: "Rolandas",
  given-vocative: "Rolandai", given-genitive: "Rolando",
  surname: "Gimbutis",
  surname-vocative: "Gimbuti", surname-genitive: "Gimbutio",
  nameLocale: "lt-LT"
}
```


Deep Dive: Handling declensions

Nominative (referring)

Vocative (**-vocative** addressing)

Genitive (**-genitive** variant) possessive case

INCORRECT
not supported in this way

	en-US	lt-LT
nominative (referring)	Mr. Rolandas Gimbutis	P. Rolandas Gimbutis
-vocative, addressing	Hello Rolandas!	Sveikas Rolandai!
-genitive, referring	Rolandas Gimbutis' Olympic Medal.	Rolando Gimbutio Olimpiniis Medalis.

Future: glottal stop handling

Many Polynesian languages use the ‘okina (‘ U+02BB) or apostrophe as a glottal stop

“‘Ōlelo Hawai‘i’

‘Aikanaka Keohokālole ⇒ monogram-long-formal ⇒ **‘K** or **‘AK**?

Future: more grammatical cases?

Many more declension cases for

- Latvian
- Lithuanian
- Finnish
- etc.

Future: pronouns?

Requests for pronouns

- **Personal_pronoun** vs he, she, they
 - “She bought the book.”
- **Possessive_adjective** vs his, her, their
 - “It is her book.”
- **Possessive_pronoun** vs his, hers, theirs.
 - “The book is hers.”
- **Reflexive_pronoun** vs himself, herself, themselves.
 - “She bought the book for herself.”

	<u>SUBJECT</u>	<u>OBJECT</u>	<u>POSSESSIVE</u>	<u>PRONUNCIATION</u>
GENDERED	she	her	hers	shee, her, herz
	he	him	his	hee, him, hiz
GENDER-NEUTRAL	they	them	theirs	thā, them, therz
	ze	hir	hirs	zhee, here, heres
	ze	zir	zirs	zhee, zhere, zheres
	xe	xem	xyrs	zhee, zhem, zheres

NEO-PRONOUNS

Questions?

Thank you!

