

# Balinese Script Reference Label Generation Rules for Second-Level Domain Names

---

*Date:* 2024-07-22

*Document version:* 1.13 (after F2F meeting)

*Authors:* Balinese Script Reference LGR Generation Panel (BRGP)

## 1. Overview

This document lays down the Balinese Script Label Generation Rule for the Second-Level Domain Names. Three main components of the Balinese Script LGR i.e., Code point repertoire, Variants and Whole Label Evaluation Rules have been stated in detail here.

The formal specification of the Balinese script LGR can be found in the accompanying XML document:

`proposed-lgr-second-level-balinese-script-03jul24-en.html`

Labels for testing can be found in the accompanying text document:

`balinese-script-test-labels-03jul24-en.txt`

## 2. Script for which the Second Level LGR is proposed.

ISO 15924 Code: Bali

ISO 15924 Key N°: 360

ISO 15924 English Name: Balinese

SNI: 9047:2021

Unicode: [Balinese – 1B00-1B7F](#)

Latin transliteration of native script name: Bali

Balinese Native name of the script: ꦱꦶꦭꦶ

## 3. Background

The development of the Internet has mainly taken place in one language namely English leading to language barriers for non-English speakers. The Internet was mostly designed based on the simple English alphabet of the 26 Latin letters, the 10 “Arabic” digits (0-9), and the hyphen (plus, of course, the dots).

IETF has brought out the implementation standards for non-Latin and non-Roman characters, how the domain name is converted into puny code and how it resolves in the root zone files. The process of supporting multilingual script and other linguistic and cultural needs on the Internet is generally known as Internationalization. Internationalized Domain Names (IDNs) are domain names or web addresses represented in local language characters.

IDNs in Balinese Language/Script is one of the effective ways of further promoting the Internet among the Balinese Populace. For the proliferation and preservation of heritage, culture and content creation in multiple languages, it is essential to have the domain names in their scripts. The process of Internationalization involves the identification of the Character Set in the respective Script for each language, the identification of the Variant Set (they have the same sound and they can be perceived as the same by the script users) and a Language-based implementation rule set. This is then followed by the process of normalization and coding for inclusion in the Domain Name System. An interface for Registrars for the issuance of Domain Names is then developed for the launch of Domain Name Registrations.

IDN is a societal issue as much as a technical challenge. Considering that a large number of users are not scholars of the language and hence can be easily cheated by homographs, spoofing, pharming as well as phishing will occur to a large extent in the Balinese language/script. This calls for great care and caution in supporting local languages and scripts in the domain names.

### 3.1 History of the Balinese Language and Script

Language is the most important part of human life because, with language, humans can communicate more perfectly. Language and script are media for expressing thoughts, feelings (emotional), spirituality, behaviour, ethnic origins, civilization, art and so on.

Bali is one of the thousands of islands in Indonesia. Bali and Indonesia are recognized internationally for their rich, diverse, and deep cultural heritage. The population of Bali in 2023 will be 4.3 million people, an average growth of 1.01% per year. In 2023, the number of elementary, middle school and high school/ vocational/ SLB students will reach 758,174 people with nearly 3200 schools available on this

island. The Balinese language is a mandatory subject to be taught at any school in Bali as a local content curriculum. The Balinese language curriculum delivers both in Latin and in Balinese script. The students in Bali start practicing the Balinese language at school from the elementary to high school level. The number of students who use Balinese names is 595,931 people (79%), and students 162,243 people use non-Balinese names (21%). The Bali Province Central Statistics Agency released the percentage of the Balinese population who use the Balinese language in their families and neighbours or relatives based on generations. 78.82% of the Post Gen Z generation or those who are currently nine years old use Balinese languages at home and 75.89% use Balinese languages among relatives or neighbors. Meanwhile, 88.07% of Generation Z or those currently aged 10-25 years use Balinese languages at home and 84.20% use regional languages among relatives or neighbours [103].

Education in the Balinese language existed during the time of the kingdoms; however, its nature was still traditional and in the family environment. The first school was founded in 1875 in Singaraja [101]. Moreover, the first textbook for school was published in 1913 with the title *Tjatoer Perenidanâ, Peratamaning tjakepan pâpeladjahan sang mâmanah maoeroek mâmaos aksarâ Belanda* [113].

In Indonesia, there are more than 600 ethnic groups and 719 mother-tongue languages spoken. The latest study by the Summer Institute of Linguistics (SIL) found that a significant 13 Indonesian mother-tongue languages have vanished and been forgotten as they are no longer used for daily communication [118]. In a mother-tongue language, the beliefs, philosophical values, rules, and traditions of a community can be understood and learned by the next generations and others [105]. When a language is forgotten, the loss is not only of the structural aspect of the language, which is the main focus of the linguistic domain but also the cultural and historical knowledge attached to the language.

The Balinese language is one of 719 living languages spoken in Indonesia and is mainly used on the islands of Bali and Lombok. People in Bali practice three different levels of Balinese language: high tongues (Basa Bali Alus), common tongues (Basa Bali Madya) and low tongues (Base Bali Sor). In the Balinese social system, there are four hierarchy levels that are highly related to how the Balinese language is used: common people (Sudra), traders or government officials (Waisya), the royal family (Kshatriya), and high priests (Brahmana). "The high and low tongues are distinct, unrelated languages with separate roots, different words, and extremely dissimilar characters" [104].

Up until today, many of the Balinese people have learned and transferred their cultural knowledge through socialization, where the Balinese let others experience the culture to understand it [110]. This can be seen in every traditional practice, such as the banjar adat and costume village (desa adat) ceremonies, which always involve a large number of people from the community [111].

The Balinese language, including its script and literary elements, is the intellectual property of the Balinese people which animates every aspect of life, both in cultural and religious activities. Balinese language with dignity and grammar becomes a medium reliable for building the character and identity of Balinese people. Bali Gubernatorial Regulation Number 80 of 2018 mandates that on Thursday, full moon, no moon (tilem), and Bali Provincial Government Anniversary must use the Balinese language and Balinese dress. Also, this regulation concerns Balinese literature and Balinese Script. For instance, the use of Balinese script in institutional markers, roads, government official letterhead, and information media as seen in Figure 1, Figure 2, and Figure 3.



Figure 1. The use of Balinese script in institution marker/name



Figure 2. Balinese script utilized to write KFC at the American fast food in Bali



பிணிக்ஞுரூபாபிக்ஞி தாமி  
**PEMERINTAH PROVINSI BALI**  
 பிக்குறாறுளிக்ஞி தாமி  
**SEKRETARIAT DAERAH**

தலைநகரம் பிணிக்ஞுரூபாபிக்ஞி தாமி (பிணிக்ஞுரூபாபிக்ஞி தாமி) பிணிக்ஞுரூபாபிக்ஞி தாமி  
 JALAN BASUKI RAHMAT DENPASAR – BALI (80235), TELEPON (0361) 224671  
 Website : [www.baliprov.go.id](http://www.baliprov.go.id)

Bali, 20 Mei 2021

Kepada

Yth. Kepala Perangkat Daerah di lingkungan  
Pemerintah Provinsi Bali.

di –

Tempat

SURAT EDARAN  
 NOMOR 13409 TAHUN 2021

TENTANG  
 PENULISAN AKSARA BALI  
 PADA KOP SURAT NASKAH DINAS  
 DI LINGKUNGAN PEMERINTAH PROVINSI BALI

Menindaklanjuti Peraturan Gubernur Bali Nomor 80 Tahun 2018 tentang  
 Perlindungan Dan Penggunaan Bahasa, Aksara, dan Sastra Bali serta  
 Penyelenggaraan Bulan Bahasa Bali, maka perlu dilakukan penyesuaian pelaksanaan  
 Tata Naskah Dinas di lingkungan Pemerintah Provinsi Bali, khususnya pada bagian  
 Kop Surat yang wajib menggunakan tulisan Aksara Bali di atas tulisan Latin. (contoh  
 terlampir).

Figure 3. Bali Province instructed to use Balinese script on every official letterhead

The Brāhmī script had developed approximately around 300 BC (Bühler, 1962). Brāhmī spread to northern and southern parts of India and it developed very widely geographically. There were two major waves in the spread of the Brāhmī script. The northern Brāhmī script later developed into various scripts, such as Siddhamatṛka, Kuṭiḷa, Proto-Bangla, Devanāgarī, Gujarātī, Śarada, Tibetan, Nepālākṣara/Rañjana, and so on. Meanwhile, the southern Brāhmī script itself developed into the Tamil, Kannada, Telugu, Malayalam, Sinhala, Khmer, Lao, Cham, Thai, Kawi, Batak, Sundanese, Javanese and other Indonesian archipelago scripts, including Bali.

Balinese is a Malayo-Polynesian language spoken by about 3.3 million people mainly in Bali in Indonesian. there are now more than 3 million speakers of Balinese. Around 80,000 live on the neighbouring island of Lombok while, largely due to transmigration, 40,000 speakers are found in southern Sumatra, and 60,000 in Sulawesi. According to the Bali Cultural Agency, a million or so people use Balinese in their everyday lives. However, in urban areas Balinese-speaking parents speak Indonesian to their children. The Balinese language is classified as an Austronesian language. The development of the Balinese language can be divided into three periods, namely *Bahasa Bali Kuna* (the ancient Balinese language), *Bahasa Bali Tengahan* (the middle age of the Balinese language) and *Bahasa Bali Baru* (the new Balinese language).

The Balinese language is written either in the Latin alphabet or the Balinese script. The Latin alphabet is generally used to write in the modern context, while the Balinese script is used for more traditional narratives and original manuscripts were written on palm leaves [114] . The usage of Balinese script in traditional manuscripts known as Lontar (palm leave) is still prevalent today. Lontars are still copied, read, and appreciated in Bali. Many activities utilize the Balinese script, both cultural and religious, and are taught in schools as a mandatory subject.

The Balinese script, or *Aksara Bali*, is one out of 30 scripts available in Indonesia. This script is used for writing the native Balinese language known as *Basa Bali*. The Balinese script is derived from the Pallava and Devanagari scripts of India, and it has many similarities with modern scripts of South Asia and Southeast Asia. The Balinese script is mainly used for writing Kawi, or Old Javanese, which had a strong influence on the Balinese language [106][115].

The Balinese script is called Hanacaraka and it has been used since the 11th century AD [108] . The Balinese script can be used not only to write the Balinese language but also to write any language. This is because the Balinese script is based on the phonetic sounds of words, allowing it to represent the sounds of different languages effectively. The Balinese Script of Hanacaraka has common 18 consonant characters inherent to the syllables as shown in Figure 4.





Figure 4 The basic Balinese script - Hanacaraka

Vowel characters can be attached either after, before, above or below the main script or syllable [117] as illustrated in Figure 5. However, only one vowel is attached to the syllable ha. Furthermore, there are only two punctuations, that is, coma and period (Figure 6). Finally, the script is continuous and has no spaces between the words. The complex structures and rules for writing the script are known as *pasang aksara*. Due to the complex structure of the *pasang aksara*, other academics have researched the development of digital tools for auto-correction for Balinese script on Android operation systems (Iswara et al., 2019).



Figure 5 Vowels positions (Pramartha & Dwidasmaras, 2014)

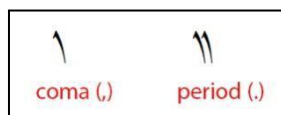


Figure 6 Balinese punctuation

Though everyday use of the script has largely been supplanted by the Latin alphabet, the Balinese script has a significant prevalence in many of the island's traditional ceremonies and is strongly associated with the Hindu religion. The script is mainly used today for copying lontar or palm-leaf manuscripts containing religious texts.

In total, there are over 185 unique characters or glyphs in the Balinese script. These are categorized into syllables, vowels, numbers, and punctuation [116]. The Balinese script is written from left to right without spaces between words or sentences, requiring readers to have a good understanding of Balinese vocabulary to comprehend the text.

## 3.2 Balinese Character Categorization

Balinese code points (Table 1) can be divided into four classifications: vowel, consonant, sign, and number.

- **Vowel** (*Aksara Suara*), there are two classifications which are both accepted for IDNs.

- Dependent vowel. Dependent vowels can be utilized to write Balinese script in any language and need to be combined with consonants or independent vowels.

Example:

- *Melali*: ꦩꦺꦭꦭꦶ (vacation)  
U+1B2B U+1B42 U+1B2E U+1B2E U+1B36
- *Perpustakaan*: ꦥꦺꦫꦸꦥꦸꦠꦱꦠꦏꦁꦠꦤ꧀ (library)  
U+1B27 U+1B42 U+1B03 U+1B27 U+1B38 U+1B32 U+1B44 U+1B22  
U+1B13 U+1B35 U+1B26 U+1B44
- Independent vowel: The independent vowel is primarily used for personal names and languages other than Balinese, such as Indonesian and English. The independent vowel is utilized in combination with *aksara wreastra*. For example, the name 'Ida Ayu' can be written in Balinese script in four different variations:

ꦶꦢꦩꦸꦪ (Ida Ayu)

U+1B07 U+1B24 U+1B05 U+1B2C U+1B38

- ꦶꦢꦩꦸꦪ (Ida Ayu) without utilizing independent vowel  
U+1B33 U+1B36 U+1B24 U+1B33 U+1B2C U+1B38
- ꦶꦢꦩꦸꦪ (Ida Ayu)  
U+1B33 U+1B36 U+1B24 U+1B05 U+1B2C U+1B38
- ꦶꦢꦩꦸꦪ (Ida Ayu)  
U+1B07 U+1B24 U+1B33 U+1B2C U+1B38

Note: According to the rules (*pasang pageh*) to write the Balinese script, there are certain spelling when using vowel ꦶ (1B0D) or vowel ꦶ (1B0B). However, for IDNs, as it is not required to restrict spelling and it may not carry meaning in the lexical sense, alternate spelling is also allowed. The examples below are similar cases of 'color' vs 'colour' in English.

- Independent vowel ꦶ (1B0D), by spelling requirement, it will be used when consonant ꦶ (1B2E) combines with vowel sign pepetꦶ (1B42).

Examples

- *lekad*: ꦭꦺꦏꦢꦺ (born)  
U+1B0D U+1B13 U+1B24 U+1B44 (correct spelling)



- □□□□□  
U+1B2E U+1B42 U+1B13 U+1B24 U+1B44 (wrong spelling, but accepted for IDNs)

- Independent vowel ꦱ (1B0B) by spelling requirement, it will be used when consonant ꦱ (1B2D) combines with vowel sign pepetꦱ (1B42).

Example

- *rereh*: ꦱꦱꦱ (find)  
U+1B0B U+1B0B U+1B04 (correct spelling)
- ꦱꦱꦱ  
U+1B2D U+1B42 U+1B2D U+1B42 U+1B04 (wrong spelling, but accepted for IDNs)

- **Consonant**, the writing system of Balinese script is the abugida system which is a segmental writing in Balinese characters in which consonant–vowel sequences are written as units. Balinese consonants have an inherent -a vowel sound. The Balinese script system consists of between 18 to 33 basic characters, depending on the language spoken. Each consonant, like in other Brahmi scripts, constitutes a syllable with an inherent vowel /a/ that could be modified by adding particular diacritics (Sound Character).
  - *Aksara Wreastra* is a Balinese character that is used to write the common Balinese language, which consists of Ha Na Ca Ra Ka Da Ta Sa Wa La Ma Ga Ba Nga Pa Ja Ya Nya
  - *Aksara Swalalita* is a Balinese character that is used to write words that derived from Sanskrit and Old Javanese, for example, in literature purposes like *Kekawin* and *Kidung*.
  - *Aksara Modre* is a Balinese character that is used to write mantras (*rerajahan*).

For the purpose of the Balinese script LGR for second-level domain names, only two types of consonant characters (*Aksara Wreastra* and *Aksara Swalalita*) are allowed as they are commonly used for everyday life in Bali.

- **Sign**, the signs (see Table 1) can be used for multiple purposes:
  - a. Two signs, 1B00 ulu ricem □ and 1B01 ulu candra □, are excluded from the repertoire because they are only used to write a prayer or holy script such as *om* □□.

For example:

- *Om sang naga raja* □□□□□□□□□□

- *Om swastiastu* □□□□□□□□□□□□□□□□ is a word to welcome and greet people in Bali.
- b. The Behavior of Ra. 1B03 Balinese Sign Surangꦱ: In traditional Balinese script rules, the sign surangꦱ typically represents a final consonant -r. For example, in the word "pasar" ꦥꦱꦱꦫ (market), is represented by the Unicode sequence U+1B27 U+1B32 U+1B03. However, in the context of Internet domain names, the character U+1B03 can be placed anywhere as long as it is used to represent 'r' and not 'ra'.
- c. In traditional Balinese script rules, the behavior of nga ꦏ (U+1B17) changes to ng ceceꦏꦺ (U+1B02) instead of nga ꦏ (U+1B17) followed by adeg adeg ꦏꦺ (U+1B44). However, in the context of Internet domain names, users can use either representation. Typically, users can differentiate between them by their distinct shapes.
- When the end of the syllable of the basic word is *ng*. For example
    - *pucung*: ꦥꦸꦏꦸꦁ (bottle)  
U+1B27 U+1B38 U+1B18 U+1B38 U+1B02
    - *barang*: ꦧꦫꦁ (goods)  
U+1B29 U+1B2D U+1B02
  - When the end of the syllable of the basic word is *ng* gets a suffix, for instance *ne*. For example
    - *siungne*: ꦱꦶꦁꦏꦺ (fangs)  
U+1B32 U+1B36 U+1B2C U+1B38 U+1B02 U+1B26 U+1B3E
    - *pedangne*: ꦥꦺꦢꦁꦏꦺ (sword)  
U+1B27 U+1B42 U+1B24 U+1B02 U+1B26 U+1B3E
  - When the end of the syllable of the basic word is *ng* gets a suffix, for instance *ne*. For example
    - *cangcang*: ꦏꦁꦏꦁ (anchor)  
U+1B18 U+1B02 U+1B18 U+1B02
    - *kungkung*: ꦏꦸꦏꦸꦁ (brackets)  
U+1B13 U+1B38 U+1B02 U+1B13 U+1B38 U+1B02
  - When we want to avoid using the below 2-base (see Figure 7). For example
    - *angklung*: ꦲꦁꦏꦭꦸꦁ (musical instrument made from two bamboo tubes)  
U+1B33 U+1B02 U+1B13 U+1B44 U+1B2E U+1B38 U+1B02
    - *sungklit*: ꦱꦸꦁꦏꦭꦶꦠ ((v) insert it at the waist)  
U+1B32 U+1B38 U+1B02 U+1B13 U+1B44 U+1B2E U+1B36  
U+1B22 U+1B44

d. The sign 1B34 *nuktā* is not included in the repertoire due to this sign seldom found being used today.

- **Numbers**, Numbers 1 2 3 4 5 6 7 8 9 0 in Balinese script). See Table 1.

**Table 1. Balinese Code Point Category**

Category	Numbers	Sub-Category	Code Points
Consonants	18	<i>Wreastra</i>	ꦱ(1B33), ꦱ(1B26), ꦱ(1B18), ꦱ(1B2D), ꦱ(1B13), ꦱ(1B24), ꦱ(1B22), ꦱ(1B32), ꦱ(1B2F), ꦱ(1B2E), ꦱ(1B2B) ꦱ(1B15), ꦱ(1B29), ꦱ(1B17), ꦱ(1B27), ꦱ(1B1A), ꦱ(1B2C) ꦱ(1B1C)
	15	<i>Swalalita</i>	ꦱ(1B21), ꦱ(1B19), ꦱ(1B14), ꦱ(1B16), ꦱ(1B1B), ꦱ(1B1D), ꦱ(1B1E), ꦱ(1B1F)*, ꦱ(1B20), ꦱ(1B23), ꦱ(1B25), ꦱ(1B28), ꦱ(1B2A), ꦱ(1B30), ꦱ(1B31)
Vowel	14	Independent Vowel	ꦱ(1B05), ꦱ(1B06), ꦱ(1B07), ꦱ(1B08), ꦱ(1B09), ꦱ(1B0A), ꦱ(1B0B), ꦱ(1B0C), ꦱ(1B0D), ꦱ(1B0E), ꦱ(1B0F), ꦱ(1B10), ꦱ(1B11), ꦱ(1B12)
	15	Dependent Vowel Signs	ꦱ(1B35)ꦱ(1B36)ꦱ(1B37)ꦱ(1B38)ꦱ(1B39), ꦱ(1B3A), ꦱ(1B3B)ꦱ(1B3C), ꦱ(1B3D), ꦱ(1B3E), ꦱ(1B3F), ꦱ(1B40), ꦱ(1B41), ꦱ(1B42), ꦱ(1B43)
Signs	5	Various Signs	ꦱ(1B00)*ꦱ(1B01)*ꦱ(1B03)ꦱ(1B02), ꦱ(1B04)
	2	Signs	ꦱ(1B34)*, ꦱ(1B44)
Numbers	10	Numbers	ꦱ(1B50)*, ꦱ(1B51)*, ꦱ(1B52)*, ꦱ(1B53)*, ꦱ(1B54)*, ꦱ(1B55)*, ꦱ(1B56)*, ꦱ(1B57)*, ꦱ(1B58)*, ꦱ(1B59)*

\*Code point marked will not be included in the code point repertoire (refer to section 5.1)

### 3.3 Features and Categories of Characters

The traditional Balinese script can be written with six components (see Figure 7 (a)), namely:

1. Baseline
2. Pre-base
3. Post-base
4. Above-base
5. Below 1-base
6. Below 2-base

However, based on the agreement among the Balinese people to make it easy to write and read the script, since 2002 and forward all the Balinese script should be written without the Below 2-base [102] [120]. Moreover, IDNs also use only five components (see Figure 7 (b)) due to the limitation of presenting the Internet domain name in the Internet browser URL address.

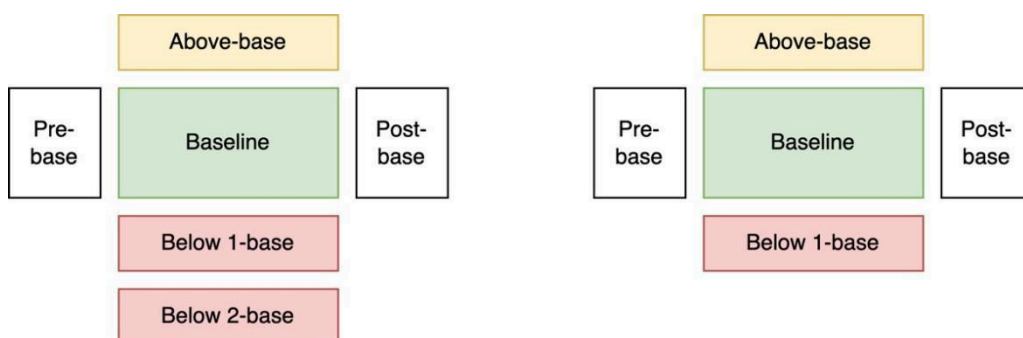


Figure 7. (a) Character position of the Balinese Script [106] , (b) Propose Character Position of Balinese script for the Internet domain name

Example of using all positions in the Balinese script

1. *Om Swastiastu* = □□□□□□□□□□□□□□ (greeting in Balinese) (type a)
2. *Dumogi becik* = □□□□□□□□□□□□□□ (hoping everything ok) (type b)
3. *Matur Suksma* = □□□□□□□□□□□□□□ (thank you) (type b)



Figure 8. Example using type (a) of the Balinese script character position on the Chrome browser

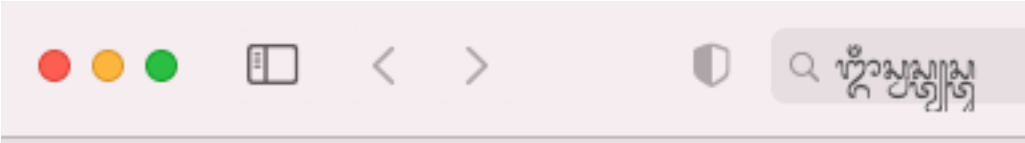


Figure 9. Example using type (a) of the Balinese script character position on the Safari browser

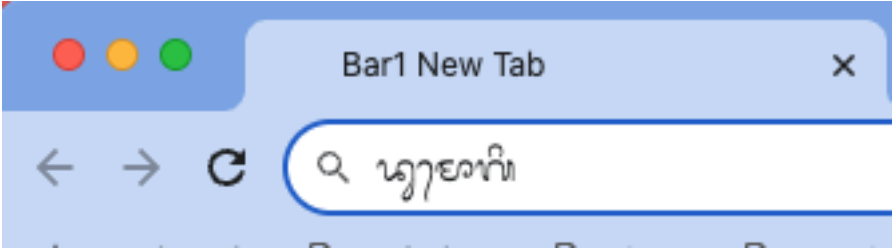


Figure 10. Example using type (b) of the Balinese script character position on the Chrome browser.

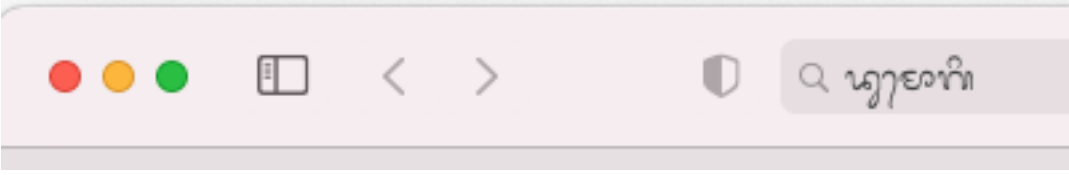


Figure 11. Example using type (b) of the Balinese script character position on the Safari browser

### 3.4 Fonts in Balinese Script

Several types of Balinese script fonts have been developed by the communities:

Developed by Aditya and David Kamholz ([source](#)):

1. [Vimala](#) = 𑄛𑄛
2. [Pustaka Bali](#) = 𑄛𑄛
3. [Kadiri](#) = 𑄛𑄛

Developed by Google:

[Noto Sans Balinese](#) = 𑄛𑄛

Developed by [I Wayan Sukanta](#) and David Kamholz:

[Natya](#) = 𑄛𑄛

[Developed by Balinese Language instructors:](#)

1. 𑄛𑄛 = Bali Banat
2. 𑄛𑄛 = Bali Raung
3. 𑄛𑄛 = Mangu sastra
4. 𑄛𑄛 = Narada
5. 𑄛𑄛 = Sapuh Luih
6. 𑄛𑄛 = Sudamala
7. 𑄛𑄛 = Tutud Bali
8. 𑄛𑄛 = Urdhasastra
9. 𑄛𑄛 = Singhambara

The recommended font for the Balinese script is Natya due to this font has been approved for use as the Indonesian national standard (SNI: 9047:202)

### 3.4 The Balinese Script Computer Keyboard (Tamiang)

This invention relates to a computer input device, namely a computer keyboard with a Balinese script layout interface for writing Balinese script on a computer application with Microsoft Windows and Mac operating systems without auto-correction if there is a mistake in writing the Balinese script. This keyboard invention is comprised of a frame, a circuit board mounted within the frame, and a plurality of keys placed on top of the circuit board. In addition, each variation key on the keyboard consists of two to four related letter variations.

Furthermore, the output of variation letters is generated from the button combinations of GANTASURI, GANTAPA, and GANTACEM to display the Balinese script that the user desires. This invention is in the form of a computer keyboard with a Balinese script layout that can be used to type 93 Balinese script characters (F) and the resulting document can be copied and shared without changing the appearance of the document that was previously created (Pramartha et al., 2021).

This keyboard was introduced to the public by the governor of Bali in 2021<sup>1</sup> and has been distributed to many schools around Bali. More information about this product can be found at <https://tamiang.oss.web.id>.

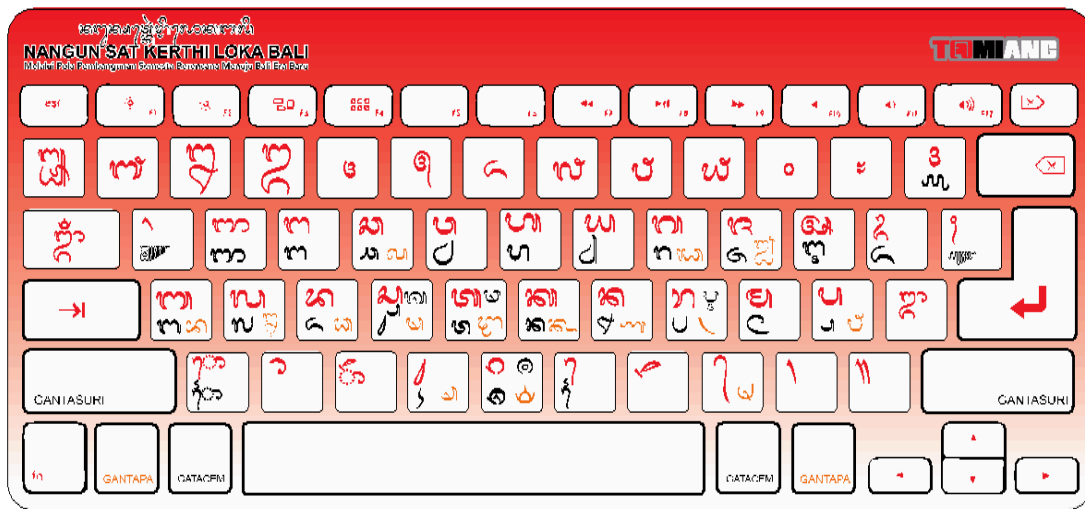


Figure 12. The Balinese script keyboard (Tamiang)

At the beginning of developing the Balinese keyboard, the glyphs for 1B11 ☐ (Okara) and 1B53 ☐ (digit 3) have the same shape/glyphs (user view) and the developer doesn't want to confuse users about the same shape with different purposes. However, the 1B11 ☐ will be included in the next shipping of the keyboard

## 4. Overall Development Process and Methodology

The Label Generation Rules (LGR) detailed in this document were meticulously developed by the members of the Balinese Script Reference LGR Generation Panel (BRGP). This process involved extensive consultations and multiple open discussions with experts in the field as well as with various community groups to ensure that the rules are comprehensive, culturally accurate, and widely accepted.

## 5. Repertoire

The Balinese script letters and ASCII-Hyphen are permitted in the same label

- o `www.☐☐☐☐-☐☐☐☐.id` : [www.dumogi-becik.id](http://www.dumogi-becik.id)
- o `www.☐-☐☐☐☐.id` : [www.i-gusti.id](http://www.i-gusti.id)
- o `☐-☐☐☐☐.id` : [i-gusti.id](http://i-gusti.id)

Mixed-script (beyond Balinese script and ASCII-Hyphen) is not allowed. The digits in both ASCII and Balinese are not allowed. The digit in Balinese script is not included

<sup>1</sup> <https://bmc.baliprov.go.id/news/title/gubernur-bali-luncurkan-keyboard-aksara-bali>



because some of them share a similar shape, which could pose a security risk. For example: Balinese digit two ꦒ (1B52) and letter la lenga ꦭ (1B0D, Balinese digit eight ꦒ (1B58) and letter pa kapal ꦥ (1B28).

Punctuation in Balinese script will NOT be allowed in internet domain names to ensure compatibility and stability across various systems and platforms. Allowing punctuation could lead to technical complications, such as difficulties in domain name resolution and inconsistent rendering across different devices and browsers. By restricting the use of punctuation, we aim to maintain a standardized and reliable approach to domain name management, facilitating easier access and reducing the risk of errors or misinterpretations [112] [116] [117] [119] [120].

If users wish to express numbers, they should use words for the representation. For instance, one (siki in Balinese) ꦱꦶꦏꦶ (U+1B32 U+1B36 U+1B13 U+1B36) instead of 1 ꦱ (U+1B51).

## 5.1 Included Code Points

**Table 2. Included Code Points**

No.	Code Point	Glyph	Name	Ref.
1	002D	-	HYPHEN-MINUS	
2	1B02		BALINESE SIGN CECEK	[112][116] [117] [119] [120]
3	1B03		BALINESE SIGN SURANG	[112][116] [117] [119] [120]
4	1B04	ꦒ	BALINESE SIGN BISAH	[112][116] [117] [119] [120]
5	1B05	ꦒ	BALINESE LETTER AKARA	[112][116] [117] [119] [120]
6	1B06	ꦒ	BALINESE LETTER AKARA TEDUNG	[112][116] [117] [119] [120]
7	1B07	ꦒ	BALINESE LETTER IKARA	[112][116] [117] [119] [120]
8	1B08	ꦒ	BALINESE LETTER IKARA TEDUNG	[112][116] [117] [119] [120]
9	1B09	ꦒ	BALINESE LETTER UKARA	[112][116] [117] [119] [120]
10	1B0A	ꦒ	BALINESE LETTER UKARA TEDUNG	[112][116] [117] [119] [120]
11	1B0B	ꦒ	BALINESE LETTER RA REPA	[112][116] [117]

No.	Code Point	Glyph	Name	Ref.
				[119] [120]
12	1B0C	ꦫꦺꦥꦩꦠꦺꦁ	BALINESE LETTER RA REPA TEDUNG	[112][116] [117] [119] [120]
13	1B0D	ꦭꦺꦁꦩꦠꦺꦁ	BALINESE LETTER LA LENG TEDUNG	[112][116] [117] [119] [120]
14	1B0E	ꦭꦺꦁꦩꦠꦺꦁ	BALINESE LETTER LA LENG TEDUNG	[112][116] [117] [119] [120]
15	1B0F	ꦺꦏꦫ	BALINESE LETTER EKARA	[112][116] [117] [119] [120]
16	1B10	ꦲꦶꦏꦫ	BALINESE LETTER AIKARA	[112][116] [117] [119] [120]
17	1B11	ꦺꦏꦫ	BALINESE LETTER OKARA	[112][116] [117] [119] [120]
18	1B12	ꦺꦏꦫꦠꦺꦁ	BALINESE LETTER OKARA TEDUNG	[112][116] [117] [119] [120]
19	1B13	ꦏ	BALINESE LETTER KA	[112][116] [117] [119] [120]
20	1B14	ꦏꦩꦲꦥꦫꦤ	BALINESE LETTER KA MAHAPRANA	[112][116] [117] [119] [120]
21	1B15	ꦒ	BALINESE LETTER GA	[112][116] [117] [119] [120]
22	1B16	ꦒꦺꦫ	BALINESE LETTER GA GORA	[112][116] [117] [119] [120]
23	1B17	ꦒ	BALINESE LETTER NGA	[112][116] [117] [119] [120]
24	1B18	ꦕ	BALINESE LETTER CA	[112][116] [117] [119] [120]
25	1B19	ꦕꦭꦕ	BALINESE LETTER CA LACA	[112][116] [117] [119] [120]
26	1B1A	ꦗ	BALINESE LETTER JA	[112][116] [117] [119] [120]
27	1B1B	ꦗꦗ	BALINESE LETTER JA JERA	[112][116] [117] [119] [120]
28	1B1C	ꦚ	BALINESE LETTER NYA	[112][116] [117] [119] [120]
29	1B1D	ꦠꦭꦶꦏ	BALINESE LETTER TA LATIK	[112][116] [117] [119] [120]

No.	Code Point	Glyph	Name	Ref.
30	1B1E	ꦠꦩꦸꦫꦢ	BALINESE LETTER TA MURDA MAHAPRANA	[112][116] [117] [119] [120]
31	1B20	ꦢꦩꦸꦫꦢ	BALINESE LETTER DA MURDA MAHAPRANA	[112][116] [117] [119] [120]
32	1B21	ꦤꦫꦩꦧꦠ	BALINESE LETTER NA RAMBAT	[112][116] [117] [119] [120]
33	1B22	ꦠ	BALINESE LETTER TA	[112][116] [117] [119] [120]
34	1B23	ꦠꦠꦮ	BALINESE LETTER TA TAWA	[112][116] [117] [119] [120]
35	1B24	ꦢ	BALINESE LETTER DA	[112][116] [117] [119] [120]
36	1B25	ꦢꦩꦢꦸ	BALINESE LETTER DA MADU	[112][116] [117] [119] [120]
37	1B26	ꦤ	BALINESE LETTER NA	[112][116] [117] [119] [120]
38	1B27	ꦥ	BALINESE LETTER PA	[112][116] [117] [119] [120]
39	1B28	ꦥꦏꦥꦭ	BALINESE LETTER PA KAPAL	[112][116] [117] [119] [120]
40	1B29	ꦧ	BALINESE LETTER BA	[112][116] [117] [119] [120]
41	1B2A	ꦧꦏꦩꦧꦁ	BALINESE LETTER BA KEMBANG	[112][116] [117] [119] [120]
42	1B2B	ꦩ	BALINESE LETTER MA	[112][116] [117] [119] [120]
43	1B2C	ꦪ	BALINESE LETTER YA	[112][116] [117] [119] [120]
44	1B2D	ꦫ	BALINESE LETTER RA	[112][116] [117] [119] [120]
45	1B2E	ꦭ	BALINESE LETTER LA	[112][116] [117] [119] [120]
46	1B2F	ꦮ	BALINESE LETTER WA	[112][116] [117] [119] [120]
47	1B30	ꦱꦱꦒ	BALINESE LETTER SA SAGA	[112][116] [117] [119] [120]
48	1B31	ꦱꦱꦥ	BALINESE LETTER SA SAPA	[112][116] [117]

No.	Code Point	Glyph	Name	Ref.
				[119] [120]
49	1B32	ꦱ	BALINESE LETTER SA	[112][116] [117] [119] [120]
50	1B33	ꦲ	BALINESE LETTER HA	[112][116] [117] [119] [120]
51	1B35	ꦠꦸꦁ	BALINESE VOWEL SIGN TEDUNG	[112][116] [117] [119] [120]
52	1B36		BALINESE VOWEL SIGN ULU	[112][116] [117] [119] [120]
53	1B37		BALINESE VOWEL SIGN ULU SARI	[112][116] [117] [119] [120]
54	1B38		BALINESE VOWEL SIGN SUKU	[112][116] [117] [119] [120]
55	1B39		BALINESE VOWEL SIGN SUKU ILUT	[112][116] [117] [119] [120]
56	1B3A		BALINESE VOWEL SIGN RA REPA	[112][116] [117] [119] [120]
57	1B3B	ꦠꦸꦁꦱꦫ	BALINESE VOWEL SIGN RA REPA TEDUNG	[112][116] [117] [119] [120]
58	1B3C		BALINESE VOWEL SIGN LA LENGA	[112][116] [117] [119] [120]
59	1B3D	ꦠꦸꦁꦱꦫꦭ	BALINESE VOWEL SIGN LA LENGA TEDUNG	[112][116] [117] [119] [120]
60	1B3E	ꦠꦸꦁꦠꦭꦶꦁ	BALINESE VOWEL SIGN TALING	[112][116] [117] [119] [120]
61	1B3F	ꦠꦸꦁꦠꦭꦶꦁꦱꦫ	BALINESE VOWEL SIGN TALING REPA	[112][116] [117] [119] [120]
62	1B40	ꦠꦸꦁꦠꦭꦶꦁꦠꦸꦁ	BALINESE VOWEL SIGN TALING TEDUNG	[112][116] [117] [119] [120]
63	1B41	ꦠꦸꦁꦠꦭꦶꦁꦱꦫꦠꦸꦁ	BALINESE VOWEL SIGN TALING REPA TEDUNG	[112][116] [117] [119] [120]
64	1B42		BALINESE VOWEL SIGN PEPET	[112][116] [117] [119] [120]
65	1B43	ꦠꦸꦁꦠꦭꦶꦁꦱꦫꦠꦸꦁ	BALINESE VOWEL SIGN PEPET TEDUNG	[112][116] [117] [119] [120]
66	1B44	ꦠꦸꦁꦠꦭꦶꦁꦱꦫꦠꦸꦁꦠꦸꦁ	BALINESE ADEG ADEG	[112][116] [117] [119] [120]

## 5.2 Excluded Code Points

These 14 Balinese scripts are excluded from current consideration to prioritize the implementation and stabilization of the most widely used characters in the Balinese script. This approach allows for a more manageable and focused initial deployment. However, their exclusion does not preclude their future inclusion. As the system matures and we gather more feedback from users and experts, we will evaluate the feasibility and necessity of incorporating these additional scripts in subsequent phases of development.

**Table 3. Excluded Code Points**

No.	Code Point	Glyph	Name	Category
1	1B1F	ꦢꦩꦸꦫꦢꦱꦤ꧀	BALINESE LETTER DA MURDA ALPAPRANA	Consonant
2	1B00	ꦲꦶꦫꦺꦴ	BALINESE SIGN ULU RICEM	Dependent vowel
3	1B01	ꦲꦶꦫꦺꦴꦩꦤ꧀	BALINESE SIGN ULU CANDRA	Signs
4	1B34	ꦲꦶꦫꦺꦴꦫꦺꦴꦏꦤ꧀	BALINESE SIGN REREKAN	Signs
5	1B50	ꦲꦶꦫꦺꦴꦠꦺꦴ	BALINESE DIGIT ZERO	Digits
6	1B51	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴ	BALINESE DIGIT ONE	Digits
7	1B52	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT TWO	Digits
8	1B53	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT THREE	Digits
9	1B54	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT FOUR	Digits
10	1B55	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT FIVE	Digits
11	1B56	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT SIX	Digits
12	1B57	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT SEVEN	Digits
13	1B58	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT EIGHT	Digits
14	1B59	ꦲꦶꦫꦺꦴꦠꦺꦴꦫꦺꦴꦫꦺꦴꦫꦺꦴ	BALINESE DIGIT NINE	Digits

## 5.3 Normalization Forms

[Unicode Normalization Forms](#) are formally defined normalizations of Unicode strings which make it possible to determine whether any two Unicode strings are equivalent to each other. Depending on the particular Unicode Normalization Form, that equivalence can either be a canonical equivalence or a compatibility equivalence. IDNs require Normalization Form C (Canonical Composition). Therefore, the code points in NFC form are included in the repertoire, and the equivalent NFD form will be disallowed by the rules.

**Table 4. Normalization Forms of Balinese Code Points**

Source	NFC	NFD
1B06 ☐	1B06 ☐	1B05 ☐ 1B35 ☐
1B08 ☐	1B08 ☐	1B07 ☐ 1B35 ☐
1B0A ☐	1B0A ☐	1B09 ☐ 1B35 ☐
1B0C ☐	1B0C ☐	1B0B ☐ 1B35 ☐
1B0E ☐	1B0E ☐	1B0D ☐ 1B35 ☐
1B12 ☐☐	1B12 ☐☐	1B11 ☐ 1B35 ☐
1B3B ☐	1B3B ☐	1B3A ☐ 1B35 ☐
1B3D ☐	1B3D ☐	1B3C ☐ 1B35 ☐
1B40 ☐	1B40 ☐	1B3E ☐ 1B35 ☐
1B41 ☐	1B41 ☐	1B3F ☐ 1B35 ☐
1B43 ☐	1B43 ☐	1B42 ☐ 1B35 ☐

## 6. Variants

### 6.1 In-Script Variants

The Aksara Swalalita are included as variants of Aksara Wreastra because they produce the same sound and are perceived as equivalent by script users, as explained in Section 3.2.

**Table 5. In-Script Variants**

Aksara Wreastra		Sound	Aksara Swalalita	
Code Point	Glyph		Code Point	Glyph
1B14	☐	Ka	1B13	☐
1B16	☐	Ga	1B15	☐
1B19	☐	Ca	1B18	☐
1B1B	☐	Ja	1B1A	☐
1B1D	☐	Ta	1B22	☐
1B1E	☐	Ta	1B22	☐
1B25	☐	Da	1B24	☐
1B20	☐	Da	1B24	☐
1B21	☐	Na	1B26	☐
1B23	☐	Ta	1B22	☐
1B28	☐	Pa	1B27	☐
1B2A	☐	Ba	1B29	☐
1B30	☐	Sa	1B32	☐
1B31	☐	Sa	1B32	☐

## 6.2 Cross-Script Variants

In addition to scrutinizing distinctions within the Balinese script variants, the analysis was extended to encompass a cross-script variant check, including comparisons with the Sinhala, Myanmar, and Javanese scripts.

Based on the comprehensive investigation, to the best of our knowledge of the BRGP members, revealed no discernible similarities in shape that could compromise the security of Balinese IDNs from the user's perspective. This thorough examination provides reassurance regarding the visual distinctiveness and security integrity of Balinese IDNs.

There is no Cross-script variant required.

## 7. Whole Label Evaluation Rules





- pesu: ꦥꦱꦸ + ꦱꦸ = ꦥꦱꦸꦱꦸ  
U+1B27 U+1B42 U+1B32 U+1B38
  - inget: ꦲꦶꦁꦺꦠꦺ + ꦶꦁꦺꦠꦺ + ꦶꦁꦺꦠꦺ = ꦶꦁꦺꦠꦺꦶꦁꦺꦠꦺ  
U+1B33 U+1B36 U+1B17 U+1B42 U+1B22 U+1B44
- Independent vowels, e.g., ꦲ ꦶ ꦸ ꦺ ꦺ (a, i, u, e, o) can appear either before or after a consonant or another independent vowel.  
Example
    - I gusti: ꦲꦶꦁꦸꦱꦠꦶ + ꦶꦁꦸꦱꦠꦶ + ꦶꦁꦸꦱꦠꦶ = ꦶꦁꦸꦱꦠꦶꦶꦁꦸꦱꦠꦶ  
U+1B07 U+1B15 U+1B38 U+1B32 U+1B44 U+1B22 U+1B36
    - aget: ꦲꦶꦁꦺꦠꦺ + ꦶꦁꦺꦠꦺ = ꦶꦁꦺꦠꦺꦶꦁꦺꦠꦺ  
U+1B33 U+1B15 U+1B42 U+1B22 U+1B44

## Signs Restriction

### U+1B35 (ꦶ) Restriction

As listed in Section 5.3, 1B35 cannot follow a set of code points to prevent the NFD form.

## Summary of the WLE rules

In summary, the WLE Rules for Balinese script are:

1. “ꦶ” (U+1B44) is used at the end of a word, it cannot be at the beginning of a label, therefore:
  - a. it cannot be at the starting position of a label
  - b. it cannot follow a hyphen
  - c. it cannot follow itself
  - d. it cannot follow a dependent vowel
2. A dependent vowel, except U+1B35, must follow a consonant or an independent vowel.
3. U+1B35 must follow a consonant.

## 8. Contributors

- **Cokorda Rai Adi Pramatha** ([cokorda@unud.ac.id](mailto:cokorda@unud.ac.id))  
The Center for Interdisciplinary Research on the Humanities and Social Sciences ([CIRHSS](#)), Udayana University, Bali, Indonesia
- **I Wayan Kasu Wardana** ([kasuwardana860@gmail.com](mailto:kasuwardana860@gmail.com))  
The Balinese Community
- **Dewa Ayu Carma Citrawati** ([da.carma@unud.ac.id](mailto:da.carma@unud.ac.id))  
The Faculty of Humanities, Udayana University, Bali, Indonesia.

- **Ida Bagus Ary Indra Iswara** ([indraiswara@instiki.ac.id](mailto:indraiswara@instiki.ac.id))  
The Balinese Cultural Digitalization Study Centre, Institut Bisnis dan Teknologi  
Indonesia

## 9. References

- [101] Agung, A. A. G. P., & Musta, I. N. (1991). *Sejarah Pendidikan Daerah Bali*: Departemen Pendidikan dan Kebudayaan, Direktorat Jendral Kebudayaan.
- [102] Bali, D. K. P. (2002). *Pedoman Pasang Aksara Bali*.
- [103] Candrawati, I. A. (2023, 20/02/2023). Bahasa Bali di Era Digitalisasi. *Balipost.com*. Retrieved from <https://www.balipost.com/news/2023/02/20/324115/Bahasa-Bali-di-Era-Digitalisasi.html>
- [104] Covarrubias, M. (2008). *Island of Bali*. Singapore: Periplus Editions (HK) Limited.
- [105] Dixon, R. M. (1997). *The rise and fall of languages*. Cambridge: Cambridge University Press.
- [106] Everson, M. S., I Made. (2005). *Proposal for encoding the Balinese script in the UCS*. Retrieved 02 July from <https://www.unicode.org/L2/L2005/05008-n2908-balinese.pdf>
- [106] Habibi, I. (2005). *Pemrosesan Teks Berbasis Standar Unicode Aksara Bali* Institut Teknologi Bandung]. Bandung. [https://informatika.stei.itb.ac.id/~rinaldi.munir/TA/Makalah\\_TA%20Imam%20Habibi%20\(Indonesian\).pdf](https://informatika.stei.itb.ac.id/~rinaldi.munir/TA/Makalah_TA%20Imam%20Habibi%20(Indonesian).pdf)
- [107] Iswara, I. B. A. I., Santika, P. P., & Wijaya, I. N. S. W. (2019, 9-11 Oct. 2019). An Algorithm for Auto-Correction in PaTik Bali Using Pasang Pageh Aksara Wianjana. 2019 5th International Conference on New Media Studies (CONMEDIA),
- [108] Nala, N. (2006). *Aksara Bali dalam USADA*. Paramita.
- [109] Pramatha, C., & Dwidasmaria, I. B. G. (2014, 1-4 June 2014). The composition approach non-QWERTY keyboard for Balinese script. Humanitarian Technology Conference - (IHTC), 2014 IEEE Canada International, Montreal, Canada.
- [110] Pramatha, C., & Davis, J. G. (2016). Digital Preservation of Cultural Heritage: Balinese Kulkul Artefact and Practices. In M. Ioannides, E. Fink,

- A. Moropoulou, M. Hagedorn-Saupe, A. Fresa, G. Liestøl, V. Rajcic, & P. Grussenmeyer (Eds.), *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: 6th International Conference, EuroMed 2016, Nicosia, Cyprus, October 31 – November 5, 2016, Proceedings, Part I* (pp. 491-500): Springer International Publishing.
- [111] Pramatha, C., Davis, J. G., & Kuan, K. K. Y. (2018). A Semantically-Enriched Digital Portal for the Digital Preservation of Cultural Heritage with Community Participation. In *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection: 7th International Conference, EuroMed 2018, Nicosia, Cyprus, October 29 – November 3, 2018, Proceedings*: Springer International Publishing.
- [112] Pramatha, C., Iswara, I. B. A. I., Suputra, I. P. G. H., & Dwidasmaria, I. B. G. (2021). Digital Humanities: Prototype Development for Balinese Script. In M. Ioannides, E. Fink, L. Cantoni, & E. Champion (Eds.), *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection* (pp. 205-214). Springer International Publishing.  
[https://doi.org/10.1007/978-3-030-73043-7\\_17](https://doi.org/10.1007/978-3-030-73043-7_17)
- [113] Purnama, I. G. G., & Yasa, I. P. E. G. (2020). Melacak Jejak Kepengarangan Sastrawan Bali Modern Pra-Kemerdekaan. *Humanis*(4), 457-463%V 424. doi:10.24843/JH.2020.v24.i04.p15
- [114] Rita Widiadana, N. K. E. (2011). Ancient 'lontar' manuscripts go digital. *The Jakarta Post*.  
<http://www.thejakartapost.com/news/2011/01/29/ancient-'lontar'-manuscripts-go-digital.html>
- [115] Sudewa, I. B. A. (2003). "Contemporary use of the Balinese script."  
[Http://www.unicode.org/L2/L2003/03118-balinese.pdf](http://www.unicode.org/L2/L2003/03118-balinese.pdf).
- [116] Suwija, I. N. (2012). *Ngiring Nulis Bali*. Wineka Media.
- [117] Tinggen, I. N. (1993). *Pasang aksara Bali: celah-celah kunci*.
- [118] Widiyanto, N. (2018, 24 July). Badan Bahasa Petakan 652 Bahasa Daerah di Indonesia. Retrieved from  
<https://www.kemdikbud.go.id/main/blog/2018/07/badan-bahasa-petakan-652-bahasa-daerah-di-indonesia>
- [119] The Unicode Standard, Balinese. Range: 1B00–1B7F Version 15.1. Retrieved from <https://www.unicode.org/charts/PDF/U1B00.pdf>