



# **Formex Version 4**

## **Formalized Exchange of Electronic Documents**

**Physical specifications**

***Version: 6.02.4***

***20260721***

---

### **Contents**

#### **Character encoding**

##### **Unicode**

##### **Quotation marks**

##### **Use of quotation marks in the different languages**

##### **Punctuation, parenthesis**

##### **Other text symbols**

##### **Currency symbols**

##### **Mathematical operators and symbols**

##### **Digits**

##### **Languages, scripts and letters**

##### **Cyrillic letters**

## **Greek letters**

## **Latin letters**

## **Arrows**

## **Processing instructions**

## **Filenames**

## **Fragmentation**

## **Exchange of data**

## **Image formats**

## **DOI for Official Journal**

## **Formex-4-Catalog**

## **Differences between versions**

---

# **Character encoding**

## ***0. Unicode***

Formex version 4 is based on Unicode 6.0. The encoding follows the specifications of UTF-8. It covers the official languages of the European Union as well as languages used in the European Economic Area and of some candidate Member States. The latter are included just for technical reasons and do not have any political value.

The following chapters give an overview about the different characters which are used in a specific context or for a given language. If for particular purposes other characters have to be integrated in a document the corresponding Unicode character has to be used.

Character entities may only be used under these circumstances:

- (1) numeric values corresponding to the Unicode code such as '&#x00A0;' for 'no-break-space';
- (2) special characters having a specific function within an XML instance: '&lt;', '&gt;', '&quot;', '&amp;' and '&apos;'.

All other character entities are forbidden. As Formex 4 is based on XML Schema, there is no possibility to define an entity; any character entity other than those mentioned above, will lead to a parsing error.

The creation of text entities is only possible in the context of fragment handling.

Information on the use of special characters can be found in the "Interinstitutional style guide" [<http://publications.europa.eu/code/en/en-000300.htm>].

## **Character ranges**

0020-007E, 00A0-00A3, 00A5-00A7, 00A9, 00AB-00AC, 00AE, 00B0-00B1, 00B5, 00B7, 00BB-00FF, 0100-0107, 010A-0113, 0116-011B, 011E-0123, 0126-0127, 012A-012B, 012E-0133, 0136-0137, 0139-0148, 014A-014D, 0150-015B, 015E-0165, 016A-016B, 016E-0173, 0178-017E, 0192, 01C4-01CC, 0218-021B, 0374, 037E, 0386-038A, 038C, 038E-03A1, 03A3-03CE, 0401-0405, 0407-042A, 042C, 042E-044A, 044C, 044E-044F, 0451-0455, 0457-045F, 0490-0491, 2010, 2013-2014, 2018-2022, 2026, 202F, 2030, 2032-2034, 2039-203A, 203E,

20AC, 2116, 2122, 2153-2154, 215B-215E, 2190-2199, 21D0-21D9, 21E6-21E9, 2213, 221A, 221E, 222B, 2243, 2245, 2248, 2260-2267, 2276-2277, 227C-227D, 22D5, 22DE-22DF, 2300, 2326, 2329-232B, 25A0-25A1, 25C7, 2605-2606, 2640, 2642, F106-F10A, F1B1, FB01-FB02.

*Nota bene:* The characters 00B5 and 00B7 are integrated because of legacy reasons, they should not be used for the creation of new instances.

*Nota bene:* The character 2026 (horizontal ellipsis) has also to be used as a place-holder for a dotted helpline. In this case, it should be used only once without regarding the number of dots effectively needed.

---

## 1. Quotation marks

|    |         |                                                                            |
|----|---------|----------------------------------------------------------------------------|
| "  | Name    | quotation mark                                                             |
|    | Unicode | 0022                                                                       |
|    | UTF-8   | 22                                                                         |
| '  | Name    | apostrophe/quote (single quotation mark)                                   |
|    | Unicode | 0027                                                                       |
|    | UTF-8   | 27                                                                         |
| «  | Name    | left pointing guillemet (ISO: left-pointing double angle quotation mark)   |
|    | Unicode | 00AB                                                                       |
|    | UTF-8   | C2AB                                                                       |
| »  | Name    | right pointing guillemet (ISO: right-pointing double angle quotation mark) |
|    | Unicode | 00BB                                                                       |
|    | UTF-8   | C2BB                                                                       |
| ‘  | Name    | single turned comma quotation mark                                         |
|    | Unicode | 2018                                                                       |
|    | UTF-8   | E28098                                                                     |
| ,’ | Name    | single comma quotation mark                                                |
|    | Unicode | 2019                                                                       |
|    | UTF-8   | E28099                                                                     |
| ‚  | Name    | low single comma quotation mark                                            |
|    | Unicode | 201A                                                                       |
|    | UTF-8   | E2809A                                                                     |
| ’  | Name    | single reversed comma quotation mark                                       |
|    | Unicode | 201B                                                                       |
|    | UTF-8   | E2809B                                                                     |

|   |                |                                                                                   |
|---|----------------|-----------------------------------------------------------------------------------|
| “ | <b>Name</b>    | double turned comma quotation mark                                                |
|   | <b>Unicode</b> | 201C                                                                              |
|   | <b>UTF-8</b>   | E2809C                                                                            |
| ” | <b>Name</b>    | double comma quotation mark                                                       |
|   | <b>Unicode</b> | 201D                                                                              |
|   | <b>UTF-8</b>   | E2809D                                                                            |
|   | <b>Unicode</b> | 201E                                                                              |
|   | <b>UTF-8</b>   | E2809E                                                                            |
|   | <b>Unicode</b> | 201F                                                                              |
|   | <b>UTF-8</b>   | E2809F                                                                            |
| < | <b>Name</b>    | left pointing single guillemet (ISO: single left–pointing angle quotation mark)   |
|   | <b>Unicode</b> | 2039                                                                              |
|   | <b>UTF-8</b>   | E280B9                                                                            |
| > | <b>Name</b>    | right pointing single guillemet (ISO: single right–pointing angle quotation mark) |
|   | <b>Unicode</b> | 203A                                                                              |
|   | <b>UTF-8</b>   | E280BA                                                                            |

## 2. Use of quotation marks in the different languages

| Language    | Level 1   | Level 2   | Level 3 ( ) |
|-------------|-----------|-----------|-------------|
| Bulgarian   | „...“     | „...“     | „...“       |
| Croatian    | „...“     | „...“     | „...“       |
| Czech       | „...“     | „...“     | »...«       |
| Danish      | »...«     | »...«     |             |
| Dutch       | „...“     | „...“     |             |
| English     | \ ... /   | “ ... ”   |             |
| Estonian    | „...“     | „...“     | „...“       |
| Finnish ( ) | “ ... ”   | / ... /   |             |
| French      | «...»     | “ ... ”   | \ ... /     |
| German      | „...“     | „...“     |             |
| Greek       | «...»     | “ ... ”   | \ ... /     |
| Hungarian   | „...“     | »...«     | / ... /     |
| Icelandic   | „...“ ( ) | „...“ ( ) |             |
| Irish       | “ ... ”   | \ ... /   | “ ... ”     |

|            |       |       |       |
|------------|-------|-------|-------|
| Italian    | «...» | "..." | \.../ |
| Lithuanian | „...“ | „...“ | „...“ |
| Latvian    | "..." | "..." | "..." |
| Maltese    | "..." | \.../ | "..." |
| Polish     | „...“ | »...« | \.../ |
| Portuguese | «...» | "..." | \.../ |
| Romanian   | „...“ | «...» | „...“ |
| Slovak     | „...“ | ,...` | »...« |
| Slovenian  | „...“ | ,...` | ,...` |
| Spanish    | «...» | "..." | \.../ |
| Swedish    | "..." | '...' |       |
| Turkish    | ()    | ()    | ()    |

- ( ) The third level is not always specified. In general, only two levels are supported, for all deeper levels the markup for the second level is re-used.
- ( ) Not yet available.
- ( ) Not yet confirmed.
- ( ) It has to be taken into account that after the two expressions 'jäljempänä' and 'Tässä ... tarkoitetaan' the quotation marks of the second level have to be used instead of those for the first level.

### 3. Punctuation, parenthesis

|   |                |                                                      |
|---|----------------|------------------------------------------------------|
|   | <b>Name</b>    | period (ISO: full stop)                              |
| ▪ | <b>Unicode</b> | 002E                                                 |
|   | <b>UTF-8</b>   | 2E                                                   |
|   | <b>Name</b>    | comma                                                |
| , | <b>Unicode</b> | 002C                                                 |
|   | <b>UTF-8</b>   | 2C                                                   |
|   | <b>Name</b>    | semicolon (NB: Different from "Greek question mark") |
| ; | <b>Unicode</b> | 003B                                                 |
|   | <b>UTF-8</b>   | 3B                                                   |
|   | <b>Name</b>    | colon                                                |
| : | <b>Unicode</b> | 003A                                                 |
|   | <b>UTF-8</b>   | 3A                                                   |
| ? | <b>Name</b>    | question mark                                        |
|   | <b>Unicode</b> | 003F                                                 |

|   |                |                                                     |
|---|----------------|-----------------------------------------------------|
| ¿ | <b>UTF-8</b>   | 3F                                                  |
|   | <b>Name</b>    | inverted question mark                              |
|   | <b>Unicode</b> | 00BF                                                |
| ! | <b>UTF-8</b>   | C2BF                                                |
|   | <b>Name</b>    | exclamation mark                                    |
|   | <b>Unicode</b> | 0021                                                |
| ¡ | <b>UTF-8</b>   | 21                                                  |
|   | <b>Name</b>    | inverted exclamation mark                           |
|   | <b>Unicode</b> | 00A1                                                |
| ; | <b>UTF-8</b>   | C2A1                                                |
|   | <b>Name</b>    | Greek question mark (NB: Different from semicolon)  |
|   | <b>Unicode</b> | 037E                                                |
| ; | <b>UTF-8</b>   | CDBE                                                |
|   | <b>Name</b>    | Greek middle dot (NB: Used as semicolon)            |
|   | <b>Unicode</b> | 0387                                                |
| . | <b>UTF-8</b>   | CE87                                                |
|   | <b>Name</b>    | middle dot (NB: Should not be used instead of 0387) |
|   | <b>Unicode</b> | 00B7                                                |
| ( | <b>UTF-8</b>   | C2B7                                                |
|   | <b>Name</b>    | opening parenthesis (ISO: left parenthesis)         |
|   | <b>Unicode</b> | 0028                                                |
| ) | <b>UTF-8</b>   | 28                                                  |
|   | <b>Name</b>    | closing parenthesis (ISO: right parenthesis)        |
|   | <b>Unicode</b> | 0029                                                |
| [ | <b>UTF-8</b>   | 29                                                  |
|   | <b>Name</b>    | opening square bracket (ISO: left square bracket)   |
|   | <b>Unicode</b> | 005B                                                |
|   | <b>UTF-8</b>   | 5B                                                  |

]

**Name** closing square bracket (ISO: right square bracket)

**Unicode** 005D

**UTF-8** 5D

{

**Name** opening curly bracket (ISO: left curly bracket)

**Unicode** 007B

**UTF-8** 7B

}

**Name** closing curly bracket (ISO: right curly bracket)

**Unicode** 007D

**UTF-8** 7D

<

**Name** left pointing angle bracket

**Unicode** 2329

**UTF-8** E28CA9

>

**Name** right pointing angle bracket

**Unicode** 232A

**UTF-8** E28CAA

---

#### 4. Other text symbols

#

**Name** space

**Unicode** 0020

**UTF-8** 20

**Name** non-breaking space

**Unicode** 00A0

**UTF-8** C2A0

**Name** narrow non-breaking space (thin hard space)

**Unicode** 202F

**UTF-8** 0xE2 0x80 0xAF

**Name** number sign

**Unicode** 0023

**UTF-8** 23

%

**Name** percent sign

**Unicode** 0025

|   |         |                                    |
|---|---------|------------------------------------|
| ‰ | UTF-8   | 25                                 |
|   | Name    | per mille sign                     |
|   | Unicode | 2030                               |
| & | UTF-8   | E280B0                             |
|   | Name    | ampersand                          |
|   | Unicode | 0026                               |
| * | UTF-8   | 26                                 |
|   | Name    | asterisk                           |
|   | Unicode | 002A                               |
| / | UTF-8   | 2A                                 |
|   | Name    | slash                              |
|   | Unicode | 002F                               |
| @ | UTF-8   | 2F                                 |
|   | Name    | commercial at                      |
|   | Unicode | 0040                               |
| \ | UTF-8   | 40                                 |
|   | Name    | backslash                          |
|   | Unicode | 005C                               |
| — | UTF-8   | 5C                                 |
|   | Name    | spacing underscore (ISO: low line) |
|   | Unicode | 005F                               |
| — | UTF-8   | 5F                                 |
|   | Name    | overline                           |
|   | Unicode | 203E                               |
| ~ | UTF-8   | E280BE                             |
|   | Name    | tilde (spacing)                    |
|   | Unicode | 007E                               |
| § | UTF-8   | 7E                                 |
|   | Name    | section sign                       |
|   | Unicode | 00A7                               |
| © | UTF-8   | C2A7                               |
|   | Name    | copyright sign                     |
|   | Unicode | 00A9                               |



|     |         |                                                   |
|-----|---------|---------------------------------------------------|
| ®   | UTF-8   | C2A9                                              |
|     | Name    | registered trade mark sign (ISO: registered sign) |
|     | Unicode | 00AE                                              |
| ™   | UTF-8   | C2AE                                              |
|     | Name    | trade mark sign                                   |
|     | Unicode | 2122                                              |
| °   | UTF-8   | E284A2                                            |
|     | Name    | degree sign                                       |
|     | Unicode | 00B0                                              |
| –   | UTF-8   | C2B0                                              |
|     | Name    | hyphen                                            |
|     | Unicode | 2010                                              |
| —   | UTF-8   | E28090                                            |
|     | Name    | en dash                                           |
|     | Unicode | 2013                                              |
| —   | UTF-8   | E28093                                            |
|     | Name    | em dash                                           |
|     | Unicode | 2014                                              |
| •   | UTF-8   | E28094                                            |
|     | Name    | bullet                                            |
|     | Unicode | 2022                                              |
| …   | UTF-8   | E280A2                                            |
|     | Name    | horizontal ellipsis                               |
|     | Unicode | 2026                                              |
| /   | UTF-8   | E280A6                                            |
|     | Name    | prime                                             |
|     | Unicode | 2032                                              |
| //  | UTF-8   | E280B2                                            |
|     | Name    | double prime                                      |
|     | Unicode | 2033                                              |
| /// | UTF-8   | E280B3                                            |
|     | Name    | triple prime                                      |
|     | Unicode | 2034                                              |

|   |         |                             |
|---|---------|-----------------------------|
|   | UTF-8   | E280B4                      |
|   | Name    | vertical bar                |
|   | Unicode | 007C                        |
| ^ | UTF-8   | 7C                          |
|   | Name    | circumflex accent (spacing) |
|   | Unicode | 005E                        |
| ` | UTF-8   | 5E                          |
|   | Name    | grave accent (spacing)      |
|   | Unicode | 0060                        |
| ∅ | UTF-8   | 60                          |
|   | Name    | empty set                   |
|   | Unicode | 2205                        |
| ⊘ | UTF-8   | E28885                      |
|   | Name    | diameter sign               |
|   | Unicode | 2300                        |
| ☐ | UTF-8   | E28C80                      |
|   | Name    | telephone recorder          |
|   | Unicode | F1B1                        |
| ◼ | UTF-8   | EF86B1                      |
|   | Name    | black square                |
|   | Unicode | 25A0                        |
| ★ | UTF-8   | E296A0                      |
|   | Name    | black star                  |
|   | Unicode | 2605                        |
| ☆ | UTF-8   | E29885                      |
|   | Name    | white star                  |
|   | Unicode | 2606                        |
| ♀ | UTF-8   | E29886                      |
|   | Name    | female sign                 |
|   | Unicode | 2640                        |
| ♂ | UTF-8   | E29980                      |
|   | Name    | male sign                   |
|   | Unicode | 2642                        |

|    |         |                                  |
|----|---------|----------------------------------|
| ☐  | UTF-8   | E29982                           |
|    | Name    | ballot box                       |
|    | Unicode | F106                             |
| ☐  | UTF-8   | EF8486                           |
|    | Name    | telephone                        |
|    | Unicode | F107                             |
| ☐  | UTF-8   | EF8487                           |
|    | Name    | fax                              |
|    | Unicode | F108                             |
| ☐  | UTF-8   | EF8488                           |
|    | Name    | (e-)mail                         |
|    | Unicode | F109                             |
| ☐  | UTF-8   | EF8489                           |
|    | Name    | url (universal resource locator) |
|    | Unicode | F10A                             |
| ◇  | UTF-8   | EF848A                           |
|    | Name    | white diamond                    |
|    | Unicode | 25C7                             |
| †  | UTF-8   | E29787                           |
|    | Name    | dagger                           |
|    | Unicode | 2020                             |
| ‡  | UTF-8   | E280A0                           |
|    | Name    | double dagger                    |
|    | Unicode | 2021                             |
| fi | UTF-8   | E280A1                           |
|    | Name    | Latin small ligature fi          |
|    | Unicode | FB01                             |
| fl | UTF-8   | EFAC81                           |
|    | Name    | Latin small ligature fl          |
|    | Unicode | FB02                             |
| /  | UTF-8   | EFAC82                           |
|    | Name    | Greek numeral sign               |
|    | Unicode | 0374                             |

|                                                                                   |         |              |
|-----------------------------------------------------------------------------------|---------|--------------|
|  | UTF-8   | CDB4         |
|                                                                                   | Name    | White Square |
|                                                                                   | Unicode | 25A1         |
|                                                                                   | UTF-8   | E296A1       |

Nota bene: Some characters in this table are not standard characters. They are defined in the private use area and therefore cannot be presented.

---

## 5. *Currency symbols*

|    |         |                      |
|----|---------|----------------------|
| €  | Name    | euro sign            |
|    | Unicode | 20AC                 |
|    | UTF-8   | E282AC               |
| \$ | Name    | dollar sign          |
|    | Unicode | 0024                 |
|    | UTF-8   | 24                   |
| ¢  | Name    | cent sign            |
|    | Unicode | 00A2                 |
|    | UTF-8   | C2A2                 |
| £  | Name    | pound sign           |
|    | Unicode | 00A3                 |
|    | UTF-8   | C2A3                 |
| ¥  | Name    | yen sign             |
|    | Unicode | 00A5                 |
|    | UTF-8   | C2A5                 |
| f  | Name    | Florin currency sign |
|    | Unicode | 0192                 |
|    | UTF-8   | C692                 |

---

## 6. *Mathematical operators and symbols*

|   |         |            |
|---|---------|------------|
| + | Name    | plus sign  |
|   | Unicode | 002B       |
|   | UTF-8   | 2B         |
| — | Name    | minus sign |
|   | Unicode | 002D       |
|   | UTF-8   | 2D         |

|                      |                                                                                                |
|----------------------|------------------------------------------------------------------------------------------------|
| $\pm$                | <p><b>Name</b> plus-or-minus sign</p> <p><b>Unicode</b> 00B1</p> <p><b>UTF-8</b> C2B1</p>      |
| $\mp$                | <p><b>Name</b> minus-or-plus sign</p> <p><b>Unicode</b> 2213</p> <p><b>UTF-8</b> E28893</p>    |
| $\sqrt{\phantom{x}}$ | <p><b>Name</b> square root</p> <p><b>Unicode</b> 221A</p> <p><b>UTF-8</b> E2889A</p>           |
| $\infty$             | <p><b>Name</b> infinity</p> <p><b>Unicode</b> 221E</p> <p><b>UTF-8</b> E2889E</p>              |
| $\int$               | <p><b>Name</b> integral</p> <p><b>Unicode</b> 222B</p> <p><b>UTF-8</b> E288AB</p>              |
| $=$                  | <p><b>Name</b> equals sign</p> <p><b>Unicode</b> 003D</p> <p><b>UTF-8</b> 3D</p>               |
| $<$                  | <p><b>Name</b> less-than sign</p> <p><b>Unicode</b> 003C</p> <p><b>UTF-8</b> 3C</p>            |
| $>$                  | <p><b>Name</b> greater-than sign</p> <p><b>Unicode</b> 003E</p> <p><b>UTF-8</b> 3E</p>         |
| $\neq$               | <p><b>Name</b> not equal to</p> <p><b>Unicode</b> 2260</p> <p><b>UTF-8</b> E289A0</p>          |
| $\equiv$             | <p><b>Name</b> identical to</p> <p><b>Unicode</b> 2261</p> <p><b>UTF-8</b> E289A1</p>          |
| $\leq$               | <p><b>Name</b> less-than or equal to</p> <p><b>Unicode</b> 2264</p> <p><b>UTF-8</b> E289A4</p> |



**Name** greater-than or equal to

**Unicode** 2265

**UTF-8** E289A5



**Name** less-than over equal to

**Unicode** 2266

**UTF-8** E289A6



**Name** greater-than over equal to

**Unicode** 2267

**UTF-8** E289A7



**Name** greater-than or less-than

**Unicode** 2277

**UTF-8** E289B7



**Name** less-than or greater-than

**Unicode** 2276

**UTF-8** E289B6



**Name** precedes or equal to

**Unicode** 227C

**UTF-8** E289BC



**Name** succeeds or equal to

**Unicode** 227D

**UTF-8** E289BD



**Name** equal to or precedes

**Unicode** 22DE

**UTF-8** E28B9E



**Name** equal to or succeeds

**Unicode** 22DF

**UTF-8** E28B9F



**Name** asymptotically equal to

**Unicode** 2243

**UTF-8** E28983



**Name** approximately equal to

**Unicode** 2245

**UTF-8** E28985

|   |                                    |
|---|------------------------------------|
| ≈ | <b>Name</b> almost equal to        |
|   | <b>Unicode</b> 2248                |
|   | <b>UTF-8</b> E28988                |
| ≡ | <b>Name</b> equal and parallel to  |
|   | <b>Unicode</b> 22D5                |
|   | <b>UTF-8</b> E28B95                |
| × | <b>Name</b> multiplication sign    |
|   | <b>Unicode</b> 00D7                |
|   | <b>UTF-8</b> C397                  |
| ÷ | <b>Name</b> division sign          |
|   | <b>Unicode</b> 00F7                |
|   | <b>UTF-8</b> C3B7                  |
|   | <b>Name</b> broken vertical bar    |
|   | <b>Unicode</b> 00A6                |
|   | <b>UTF-8</b> C2A6                  |
| ⌋ | <b>Name</b> not sign               |
|   | <b>Unicode</b> 00AC                |
|   | <b>UTF-8</b> C2AC                  |
| ≠ | <b>Name</b> not identical to       |
|   | <b>Unicode</b> 2262                |
|   | <b>UTF-8</b> E289A2                |
| ≡ | <b>Name</b> strictly equivalent to |
|   | <b>Unicode</b> 2263                |
|   | <b>UTF-8</b> E289A3                |

---

## 7. Digits

|   |                        |
|---|------------------------|
| 0 | <b>Name</b> digit zero |
|   | <b>Unicode</b> 0030    |
|   | <b>UTF-8</b> 30        |
| 1 | <b>Name</b> digit one  |
|   | <b>Unicode</b> 0031    |
|   | <b>UTF-8</b> 31        |

2

Name digit two

Unicode 0032

UTF-8 32

3

Name digit three

Unicode 0033

UTF-8 33

4

Name digit four

Unicode 0034

UTF-8 34

5

Name digit five

Unicode 0035

UTF-8 35

6

Name digit six

Unicode 0036

UTF-8 36

7

Name digit seven

Unicode 0037

UTF-8 37

8

Name digit eight

Unicode 0038

UTF-8 38

9

Name digit nine

Unicode 0039

UTF-8 39

## 8. Languages, scripts and letters ()

| Language            | Script   | Characters                                                                                                 |
|---------------------|----------|------------------------------------------------------------------------------------------------------------|
| Basque              | Latin    | Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Ññ Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz                           |
| Bulgarian           | Cyrillic | Aa Бб Вв Гг Дд Ее Жж Зз Ии Йй Кк Лл Мм Нн Оо Пп Рр Сс Тт Уу Фф Хх Цц Чч Шш Щщ Ъъ Ььюю Яя                   |
| Spanish (Castilian) | Latin    | Aa Áá Bb Cc Dd Ee Éé Ff Gg Hh Ii Íí Jj Ll Mm Nn Ññ Oo Óó Pp Qq Rr Ss Tt Uu Úú Üü Vv Xx Yy Zz               |
| Catalan             | Latin    | Aa Àà Bb Cc Çç Dd Ee Éé Èè Ff Gg Hh Ii Íí Îî Jj Ll L·l Mm Nn Oo Óó Òò Pp Qq Rr Ss Tt Uu Úú Üü Vv Xx Yy Zz  |
| Croatian            | Latin    | Aa Bb Cc Čč Ćć Dd Đđ Džđž Ee Ff Gg Hh Ii Jj Kk Ll Lj lj Mm Nn Nj nj Oo Pp Rr Ss Šš Tt Uu Vv Ww Xx Yy Zz Žž |



|                |       |                                                                                                                                               |
|----------------|-------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| Czech          | Latin | Aa Áá Bb Cc Čč Dd Ďď Ee Ěě ěě Ff Gg Hh Ii Íí Jj Kk Ll Mm Nn Ňň Oo Óó Pp Rr Řř Ss Šš Tt Ťť Uu Úú Ůů Vv Yy Ýý Zz Žž                             |
| Danisch        | Latin | Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz Ææ Øø Åå                                                        |
| Dutch          | Latin | Aa Àà Áá Ää Ââ Bb Cc Dd Ee Èè Éé Êê Ëë Ff Gg Hh Ii Ìì Îî Ïï Íí Iij Jj Kk Ll Mm Nn Oo Òò Óó Öö Ôô Pp Qq Rr Ss Tt Uu Ùù Úú Üü Ûû Vv Ww Xx Yy Zz |
| English        | Latin | Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz                                                                 |
| Estonian       | Latin | Aa Bb Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Rr Ss Šš Zz Žž Tt Uu Vv Yy Öö Ää Õõ Üü                                                           |
| Finnish        | Latin | Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz Åå Ää Öö [Šš Žž]                                                |
| French         | Latin | Aa Àà Ââ Bb Cc Çç Dd Ee Éé Èè Êê Ëë Ff Gg Hh Ii Îî Ïï Jj Kk Ll Mm Nn Oo Ôô Œœ Pp Qq Rr Ss Tt Uu Ùù Ûû Üü Vv Ww Xx Yy Ýý Zz                    |
| Galician       | Latin | Aa Áá Bb Cc Dd Ee Éé Ff Gg Hh Ii Íí Jj Ll Mm Nn Ññ Oo Óó Pp Qq Rr Ss Tt Uu Úú Vv Xx Yy Zz                                                     |
| German         | Latin | Aa Ää Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Öö Pp Qq Rr Ss ß Tt Uu Üü Vv Ww Xx Yy Zz                                                      |
| Greek          | Greek | Aa 'Αά Ββ Γγ Δδ Εε 'Εέ Ζζ Ηη 'Ηή Θθ Ιι 'Ιί Ίί Ή ή Κκ Λλ Μμ Νν Ξξ Οο 'Οό Ππ Ρρ Σσ ς Ττ Υυ 'Υύ Ύϋ ΰ Φφ Χχ Ψψ Ωω 'Ωώ                             |
| Hungarian      | Latin | Aa Áá Bb Cc Dd DZDzdz Ee Éé Ff Gg Hh Ii Íí Jj Kk Ll Mm Nn Oo Óó Öö Őő Pp Rr Ss Tt Uu Úú Üü Űű Vv Ww Xx Yy Zz                                  |
| Icelandic      | Latin | Aa Áá Bb Dd Đđ Ee Éé Ff Gg Hh Ii Íí Jj Kk Ll Mm Nn Oo Óó Pp Rr Ss Tt Uu Úú Vv Xx Yy Ýý Þþ Ææ Öö                                               |
| Irish (Gaelic) | Latin | Aa Áá Bb Cc Dd Ee Éé Ff Gg Hh Ii Íí Jj Ll Mm Nn Oo Óó Pp Qq Rr Ss Tt Uu Úú Vv Ww Xx Yy Zz                                                     |
| Italian        | Latin | Aa Àà Bb Cc Dd Ee Èè Éé Ff Gg Hh Ii Ìì Íí Jj Kk Ll Mm Nn Oo Òò Óó Pp Qq Rr Ss Tt Uu Ùù Úú Vv Xx Zz                                            |
| Latvian        | Latin | Aa Āā Bb Cc Čč Dd Ee Ēē Ff Gg Ģģ Hh Ii Īī Jj Kk Ķķ Ll Ļļ Mm Nn Ņņ Oo Ūū Pp Rr Ēē Ss Šš Tt Uu Ūū Vv Yy Zz Žž                                   |
| Lithuanian     | Latin | Aa Aą Bb Cc Čč Dd Ee Eę Èè Ff Gg Hh Ii Įį Yy Jj Kk Ll Mm Nn Oo Pp Rr Ss Šš Tt Uu Ųų Ūū Vv Žž                                                  |
| Maltese        | Latin | Aa Àà Bb Cc Ċċ Dd Ee Èè Ff Ġġ Gg Hh Ħħ Ii Ìì Jj Kk Ll Mm Nn Oo Òò Pp Qq Rr Ss Tt Uu Ùù Vv Ww Xx Żż Zz                                         |
| Norwegian      | Latin | Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz Ææ Øø Åå                                                        |
| Polish         | Latin | Aa Aą Bb Cc Ćć Dd Ee Eę Ff Gg Hh Ii Jj Kk Ll Łł Mm Nn Ńń Oo Óó Pp Rr Ss Śś Tt Uu Ww Yy Zz Żż Źź                                               |
| Portuguese     | Latin | Aa Áá Àà Ââ Ãã Bb Cc Çç Dd Ee Éé Êê Ff Gg Hh Ii Íí Ll Mm Nn Oo Óó Ôô Õõ Pp Qq Rr Ss Tt Uu Úú Vv Xx Zz                                         |
| Romanian       | Latin | Aa Ââ Ăă Bb Cc Dd Ee Ff Gg Hh Ii Îî Jj Ll Mm Nn Oo Pp Rr Ss Șș Tt Țț Uu Vv Xx Zz                                                              |
| Slovak         | Latin | Aa Áá Ää Bb Cc Čč Dd Ďď Ee Éé Ff Gg Hh Ii Íí Jj Kk Ll Lí Ľľ Mm Nn Ňň Oo Óó Ôô Pp Rr Řř Ss Šš Tt Ťť Uu Úú Vv Ww Xx Yy Ýý Zz Žž                 |
| Slovenian      | Latin | Aa Bb Cc Čč Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Rr Ss Šš Tt Uu Vv Zz Žž                                                                    |
| Swedish        | Latin |                                                                                                                                               |

Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu  
Vv Ww Xx Yy Zz Åå Ää Öö

Turkish Latin Aa Ââ Bb Cc Çç Dd Ee Ff Gg Ğğ Hh İı Îî Jj Kk Ll Mm Nn Oo Öö Pp  
Rr Ss Şş Tt Uu Ûû Üü Vv Yy Zz

- () These lists only specify characters which are originally used in a given language. That means that other characters which might appear in foreign or imported lexical entities are not listed.

## 9. Cyrillic letters

|   |             |                            |
|---|-------------|----------------------------|
| А | Name        | cyrillic capital letter A  |
|   | Unicode     | 0410                       |
|   | UTF-8       | D090                       |
|   | Language(s) | Bulgarian                  |
| а | Name        | cyrillic small letter a    |
|   | Unicode     | 0430                       |
|   | UTF-8       | D0B0                       |
|   | Language(s) | Bulgarian                  |
| Б | Name        | cyrillic capital letter BE |
|   | Unicode     | 0411                       |
|   | UTF-8       | D091                       |
|   | Language(s) | Bulgarian                  |
| б | Name        | cyrillic small letter be   |
|   | Unicode     | 0431                       |
|   | UTF-8       | D0B1                       |
|   | Language(s) | Bulgarian                  |
| В | Name        | cyrillic capital letter VE |
|   | Unicode     | 0412                       |
|   | UTF-8       | D092                       |
|   | Language(s) | Bulgarian                  |
| в | Name        | cyrillic small letter ve   |
|   | Unicode     | 0432                       |
|   | UTF-8       | D0B2                       |
|   | Language(s) | Bulgarian                  |
| Г | Name        | cyrillic capital letter GE |
|   | Unicode     | 0413                       |
|   | UTF-8       | D093                       |

|   |             |                                         |
|---|-------------|-----------------------------------------|
| Г | Language(s) | Bulgarian                               |
|   | Name        | cyrillic small letter ge                |
|   | Unicode     | 0433                                    |
|   | UTF-8       | D0B3                                    |
| Ґ | Language(s) | Bulgarian                               |
|   | Name        | cyrillic capital letter GJE             |
|   | Unicode     | 0403                                    |
|   | UTF-8       | D083                                    |
| ґ | Language(s) |                                         |
|   | Name        | cyrillic small letter gje               |
|   | Unicode     | 0453                                    |
|   | UTF-8       | D193                                    |
| Ґ | Language(s) |                                         |
|   | Name        | cyrillic capital letter GHE with upturn |
|   | Unicode     | 0490                                    |
|   | UTF-8       | D290                                    |
| Ғ | Language(s) |                                         |
|   | Name        | cyrillic small letter ghe with upturn   |
|   | Unicode     | 0491                                    |
|   | UTF-8       | D291                                    |
| Д | Language(s) |                                         |
|   | Name        | cyrillic capital letter DE              |
|   | Unicode     | 0414                                    |
|   | UTF-8       | D094                                    |
| д | Language(s) | Bulgarian                               |
|   | Name        | cyrillic small letter de                |
|   | Unicode     | 0434                                    |
|   | UTF-8       | D0B4                                    |
| Д | Language(s) | Bulgarian                               |
|   | Name        | cyrillic capital letter DJE             |
|   | Unicode     | 0402                                    |
|   | UTF-8       | D082                                    |
| Ѣ | Language(s) |                                         |

|   |             |                                      |
|---|-------------|--------------------------------------|
| Ѣ | Name        | cyrillic small letter dje            |
|   | Unicode     | 0452                                 |
|   | UTF-8       | D192                                 |
|   | Language(s) |                                      |
| Е | Name        | cyrillic capital letter IE           |
|   | Unicode     | 0415                                 |
|   | UTF-8       | D095                                 |
|   | Language(s) | Bulgarian                            |
| е | Name        | cyrillic small letter ie             |
|   | Unicode     | 0435                                 |
|   | UTF-8       | D0B5                                 |
|   | Language(s) | Bulgarian                            |
| Є | Name        | cyrillic capital letter Ukrainian IE |
|   | Unicode     | 0404                                 |
|   | UTF-8       | D084                                 |
|   | Language(s) |                                      |
| Є | Name        | cyrillic small letter Ukrainian ie   |
|   | Unicode     | 0454                                 |
|   | UTF-8       | D194                                 |
|   | Language(s) |                                      |
| Ж | Name        | cyrillic capital letter ZHE          |
|   | Unicode     | 0416                                 |
|   | UTF-8       | D096                                 |
|   | Language(s) | Bulgarian                            |
| ж | Name        | cyrillic small letter zhe            |
|   | Unicode     | 0436                                 |
|   | UTF-8       | D0B6                                 |
|   | Language(s) | Bulgarian                            |
| Ѥ | Name        | cyrillic capital letter DZHE         |
|   | Unicode     | 040F                                 |
|   | UTF-8       | D08F                                 |
|   | Language(s) |                                      |

|   |             |                                 |
|---|-------------|---------------------------------|
| Ѹ | Name        | cyrillic small letter dzhe      |
|   | Unicode     | 045F                            |
|   | UTF-8       | D19F                            |
|   | Language(s) |                                 |
| З | Name        | cyrillic capital letter ZE      |
|   | Unicode     | 0417                            |
|   | UTF-8       | D097                            |
|   | Language(s) | Bulgarian                       |
| з | Name        | cyrillic small letter ze        |
|   | Unicode     | 0437                            |
|   | UTF-8       | D0B7                            |
|   | Language(s) | Bulgarian                       |
| Ѕ | Name        | cyrillic capital letter DZE     |
|   | Unicode     | 0405                            |
|   | UTF-8       | D085                            |
|   | Language(s) |                                 |
| ѕ | Name        | cyrillic small letter dze       |
|   | Unicode     | 0455                            |
|   | UTF-8       | D195                            |
|   | Language(s) |                                 |
| И | Name        | cyrillic capital letter I       |
|   | Unicode     | 0418                            |
|   | UTF-8       | D098                            |
|   | Language(s) | Bulgarian                       |
| и | Name        | cyrillic small letter i         |
|   | Unicode     | 0438                            |
|   | UTF-8       | D0B8                            |
|   | Language(s) | Bulgarian                       |
| Й | Name        | cyrillic capital letter short I |
|   | Unicode     | 0419                            |
|   | UTF-8       | D099                            |
|   | Language(s) | Bulgarian                       |

|   |             |                                      |
|---|-------------|--------------------------------------|
| Й | Name        | cyrillic small letter short i        |
|   | Unicode     | 0439                                 |
|   | UTF-8       | D0B9                                 |
|   | Language(s) | Bulgarian                            |
| И | Name        | cyrillic capital letter I with grave |
|   | Unicode     | 040D                                 |
|   | UTF-8       | D08D                                 |
|   | Language(s) | Bulgarian                            |
| й | Name        | cyrillic small letter i with grave   |
|   | Unicode     | 045D                                 |
|   | UTF-8       | D19D                                 |
|   | Language(s) | Bulgarian                            |
| Ј | Name        | cyrillic capital letter JE           |
|   | Unicode     | 0408                                 |
|   | UTF-8       | D088                                 |
|   | Language(s) |                                      |
| ј | Name        | cyrillic small letter je             |
|   | Unicode     | 0458                                 |
|   | UTF-8       | D198                                 |
|   | Language(s) |                                      |
| К | Name        | cyrillic capital letter KA           |
|   | Unicode     | 041A                                 |
|   | UTF-8       | D09A                                 |
|   | Language(s) | Bulgarian                            |
| ќ | Name        | cyrillic small letter ka             |
|   | Unicode     | 043A                                 |
|   | UTF-8       | D0BA                                 |
|   | Language(s) | Bulgarian                            |
| Ќ | Name        | cyrillic capital letter KJE          |
|   | Unicode     | 040C                                 |
|   | UTF-8       | D08C                                 |
|   | Language(s) |                                      |

|   |             |                             |
|---|-------------|-----------------------------|
| Ќ | Name        | cyrillic small letter kje   |
|   | Unicode     | 045C                        |
|   | UTF-8       | D19C                        |
|   | Language(s) |                             |
| Л | Name        | cyrillic capital letter EL  |
|   | Unicode     | 041B                        |
|   | UTF-8       | D09B                        |
|   | Language(s) | Bulgarian                   |
| л | Name        | cyrillic small letter el    |
|   | Unicode     | 043B                        |
|   | UTF-8       | D0BB                        |
|   | Language(s) | Bulgarian                   |
| Љ | Name        | cyrillic capital letter LJE |
|   | Unicode     | 0409                        |
|   | UTF-8       | D089                        |
|   | Language(s) |                             |
| љ | Name        | cyrillic small letter lje   |
|   | Unicode     | 0459                        |
|   | UTF-8       | D199                        |
|   | Language(s) |                             |
| М | Name        | cyrillic capital letter EM  |
|   | Unicode     | 041C                        |
|   | UTF-8       | D09C                        |
|   | Language(s) | Bulgarian                   |
| м | Name        | cyrillic small letter em    |
|   | Unicode     | 043C                        |
|   | UTF-8       | D0BC                        |
|   | Language(s) | Bulgarian                   |
| Н | Name        | cyrillic capital letter EN  |
|   | Unicode     | 041D                        |
|   | UTF-8       | D09D                        |
|   | Language(s) | Bulgarian                   |

|   |             |                             |
|---|-------------|-----------------------------|
| Н | Name        | cyrillic small letter en    |
|   | Unicode     | 043D                        |
|   | UTF-8       | D0BD                        |
|   | Language(s) | Bulgarian                   |
| Њ | Name        | cyrillic capital letter NJE |
|   | Unicode     | 040A                        |
|   | UTF-8       | D08A                        |
|   | Language(s) |                             |
| њ | Name        | cyrillic small letter nje   |
|   | Unicode     | 045A                        |
|   | UTF-8       | D19A                        |
|   | Language(s) |                             |
| О | Name        | cyrillic capital letter O   |
|   | Unicode     | 041E                        |
|   | UTF-8       | D09E                        |
|   | Language(s) | Bulgarian                   |
| о | Name        | cyrillic small letter o     |
|   | Unicode     | 043E                        |
|   | UTF-8       | D0BE                        |
|   | Language(s) | Bulgarian                   |
| П | Name        | cyrillic capital letter PE  |
|   | Unicode     | 041F                        |
|   | UTF-8       | D09F                        |
|   | Language(s) | Bulgarian                   |
| п | Name        | cyrillic small letter pe    |
|   | Unicode     | 043F                        |
|   | UTF-8       | D0BF                        |
|   | Language(s) | Bulgarian                   |
| Р | Name        | cyrillic capital letter ER  |
|   | Unicode     | 0420                        |
|   | UTF-8       | D0A0                        |
|   | Language(s) | Bulgarian                   |



|   |             |                              |
|---|-------------|------------------------------|
| р | Name        | cyrillic small letter er     |
|   | Unicode     | 0440                         |
|   | UTF-8       | D180                         |
|   | Language(s) | Bulgarian                    |
| С | Name        | cyrillic capital letter ES   |
|   | Unicode     | 0421                         |
|   | UTF-8       | D0A1                         |
|   | Language(s) | Bulgarian                    |
| с | Name        | cyrillic small letter es     |
|   | Unicode     | 0441                         |
|   | UTF-8       | D181                         |
|   | Language(s) | Bulgarian                    |
| Т | Name        | cyrillic capital letter TE   |
|   | Unicode     | 0422                         |
|   | UTF-8       | D0A2                         |
|   | Language(s) | Bulgarian                    |
| т | Name        | cyrillic small letter te     |
|   | Unicode     | 0442                         |
|   | UTF-8       | D182                         |
|   | Language(s) | Bulgarian                    |
| Ѡ | Name        | cyrillic capital letter TSHE |
|   | Unicode     | 040B                         |
|   | UTF-8       | D08B                         |
|   | Language(s) |                              |
| ѡ | Name        | cyrillic small letter tshe   |
|   | Unicode     | 045B                         |
|   | UTF-8       | D19B                         |
|   | Language(s) |                              |
| у | Name        | cyrillic capital letter U    |
|   | Unicode     | 0423                         |
|   | UTF-8       | D0A3                         |
|   | Language(s) | Bulgarian                    |

|   |             |                                 |
|---|-------------|---------------------------------|
| у | Name        | cyrillic small letter u         |
|   | Unicode     | 0443                            |
|   | UTF-8       | D183                            |
|   | Language(s) | Bulgarian                       |
| ӱ | Name        | cyrillic capital letter short U |
|   | Unicode     | 040E                            |
|   | UTF-8       | D08E                            |
|   | Language(s) |                                 |
| ӱ | Name        | cyrillic small letter short u   |
|   | Unicode     | 045E                            |
|   | UTF-8       | D19E                            |
|   | Language(s) |                                 |
| Ѳ | Name        | cyrillic capital letter EF      |
|   | Unicode     | 0424                            |
|   | UTF-8       | D0A4                            |
|   | Language(s) | Bulgarian                       |
| ѳ | Name        | cyrillic small letter ef        |
|   | Unicode     | 0444                            |
|   | UTF-8       | D184                            |
|   | Language(s) | Bulgarian                       |
| Х | Name        | cyrillic capital letter KHA     |
|   | Unicode     | 0425                            |
|   | UTF-8       | D0A5                            |
|   | Language(s) | Bulgarian                       |
| х | Name        | cyrillic small letter kha       |
|   | Unicode     | 0445                            |
|   | UTF-8       | D185                            |
|   | Language(s) | Bulgarian                       |
| Ц | Name        | cyrillic capital letter TSE     |
|   | Unicode     | 0426                            |
|   | UTF-8       | D0A6                            |
|   | Language(s) | Bulgarian                       |

|   |             |                                   |
|---|-------------|-----------------------------------|
| Ц | Name        | cyrillic small letter tse         |
|   | Unicode     | 0446                              |
|   | UTF-8       | D186                              |
|   | Language(s) | Bulgarian                         |
| Ч | Name        | cyrillic capital letter CHE       |
|   | Unicode     | 0427                              |
|   | UTF-8       | D0A7                              |
|   | Language(s) | Bulgarian                         |
| ч | Name        | cyrillic small letter che         |
|   | Unicode     | 0447                              |
|   | UTF-8       | D187                              |
|   | Language(s) | Bulgarian                         |
| Щ | Name        | cyrillic capital letter SHA       |
|   | Unicode     | 0428                              |
|   | UTF-8       | D0A8                              |
|   | Language(s) | Bulgarian                         |
| щ | Name        | cyrillic small letter she         |
|   | Unicode     | 0448                              |
|   | UTF-8       | D188                              |
|   | Language(s) | Bulgarian                         |
| Щ | Name        | cyrillic capital letter SHCHA     |
|   | Unicode     | 0429                              |
|   | UTF-8       | D0A9                              |
|   | Language(s) | Bulgarian                         |
| щ | Name        | cyrillic small letter shcha       |
|   | Unicode     | 0449                              |
|   | UTF-8       | D189                              |
|   | Language(s) | Bulgarian                         |
| Ъ | Name        | cyrillic capital letter hard sign |
|   | Unicode     | 042A                              |
|   | UTF-8       | D0AA                              |
|   | Language(s) | Bulgarian                         |

|   |             |                                   |
|---|-------------|-----------------------------------|
| ъ | Name        | cyrillic small letter hard sign   |
|   | Unicode     | 044A                              |
|   | UTF-8       | D18A                              |
|   | Language(s) | Bulgarian                         |
| Ѣ | Name        | cyrillic capital letter soft sign |
|   | Unicode     | 042C                              |
|   | UTF-8       | D0AC                              |
|   | Language(s) | Bulgarian                         |
| Ѥ | Name        | cyrillic small letter soft sign   |
|   | Unicode     | 044C                              |
|   | UTF-8       | D18C                              |
|   | Language(s) | Bulgarian                         |
| Ю | Name        | cyrillic capital letter IU        |
|   | Unicode     | 042E                              |
|   | UTF-8       | D0AE                              |
|   | Language(s) | Bulgarian                         |
| ю | Name        | cyrillic small letter iu          |
|   | Unicode     | 044E                              |
|   | UTF-8       | D18E                              |
|   | Language(s) | Bulgarian                         |
| Я | Name        | cyrillic capital letter IA        |
|   | Unicode     | 042F                              |
|   | UTF-8       | D0AF                              |
|   | Language(s) | Bulgarian                         |
| я | Name        | cyrillic small letter ia          |
|   | Unicode     | 044F                              |
|   | UTF-8       | D18F                              |
|   | Language(s) | Bulgarian                         |
| Ё | Name        | cyrillic capital letter IO        |
|   | Unicode     | 0401                              |
|   | UTF-8       | D081                              |
|   | Language(s) |                                   |

|   |             |                          |
|---|-------------|--------------------------|
| ë | Name        | cyrillic small letter io |
|   | Unicode     | 0451                     |
|   | UTF-8       | D191                     |
|   | Language(s) |                          |

|   |             |                            |
|---|-------------|----------------------------|
| Ӣ | Name        | cyrillic capital letter YI |
|   | Unicode     | 0407                       |
|   | UTF-8       | D087                       |
|   | Language(s) |                            |

|   |             |                          |
|---|-------------|--------------------------|
| ӥ | Name        | cyrillic small letter yi |
|   | Unicode     | 0457                     |
|   | UTF-8       | D197                     |
|   | Language(s) |                          |

---

## 10. Greek letters

|   |             |                            |
|---|-------------|----------------------------|
| Α | Name        | Greek capital letter alpha |
|   | Unicode     | 0391                       |
|   | UTF-8       | CE91                       |
|   | Language(s) | Greek                      |

|   |             |                          |
|---|-------------|--------------------------|
| α | Name        | Greek small letter alpha |
|   | Unicode     | 03B1                     |
|   | UTF-8       | CEB1                     |
|   | Language(s) | Greek                    |

|   |             |                                  |
|---|-------------|----------------------------------|
| Ά | Name        | Greek capital letter alpha tonos |
|   | Unicode     | 0386                             |
|   | UTF-8       | CE86                             |
|   | Language(s) | Greek                            |

|   |             |                                |
|---|-------------|--------------------------------|
| ά | Name        | Greek small letter alpha tonos |
|   | Unicode     | 03AC                           |
|   | UTF-8       | CEAC                           |
|   | Language(s) | Greek                          |

|   |         |                           |
|---|---------|---------------------------|
| Β | Name    | Greek capital letter beta |
|   | Unicode | 0392                      |
|   | UTF-8   | CE92                      |

|   |             |                                    |
|---|-------------|------------------------------------|
| β | Language(s) | Greek                              |
|   | Name        | Greek small letter beta            |
|   | Unicode     | 03B2                               |
|   | UTF-8       | CEB2                               |
| Γ | Language(s) | Greek                              |
|   | Name        | Greek capital letter gamma         |
|   | Unicode     | 0393                               |
|   | UTF-8       | CE93                               |
| γ | Language(s) | Greek                              |
|   | Name        | Greek small letter gamma           |
|   | Unicode     | 03B3                               |
|   | UTF-8       | CEB3                               |
| Δ | Language(s) | Greek                              |
|   | Name        | Greek capital letter delta         |
|   | Unicode     | 0394                               |
|   | UTF-8       | CE94                               |
| δ | Language(s) | Greek                              |
|   | Name        | Greek small letter delta           |
|   | Unicode     | 03B4                               |
|   | UTF-8       | CEB4                               |
| Ε | Language(s) | Greek                              |
|   | Name        | Greek capital letter epsilon       |
|   | Unicode     | 0395                               |
|   | UTF-8       | CE95                               |
| ε | Language(s) | Greek                              |
|   | Name        | Greek small letter epsilon         |
|   | Unicode     | 03B5                               |
|   | UTF-8       | CEB5                               |
| Έ | Language(s) | Greek                              |
|   | Name        | Greek capital letter epsilon tonos |
|   | Unicode     | 0388                               |
|   | UTF-8       | CE88                               |
|   | Language(s) | Greek                              |

|   |             |                                  |
|---|-------------|----------------------------------|
| Έ | Name        | Greek small letter epsilon tonos |
|   | Unicode     | 03AD                             |
|   | UTF-8       | CEAD                             |
|   | Language(s) | Greek                            |
| Ζ | Name        | Greek capital letter zeta        |
|   | Unicode     | 0396                             |
|   | UTF-8       | CE96                             |
|   | Language(s) | Greek                            |
| ζ | Name        | Greek small letter zeta          |
|   | Unicode     | 03B6                             |
|   | UTF-8       | CEB6                             |
|   | Language(s) | Greek                            |
| Η | Name        | Greek capital letter eta         |
|   | Unicode     | 0397                             |
|   | UTF-8       | CE97                             |
|   | Language(s) | Greek                            |
| η | Name        | Greek small letter eta           |
|   | Unicode     | 03B7                             |
|   | UTF-8       | CEB7                             |
|   | Language(s) | Greek                            |
| Ἡ | Name        | Greek capital letter eta tonos   |
|   | Unicode     | 0389                             |
|   | UTF-8       | CE89                             |
|   | Language(s) | Greek                            |
| ἥ | Name        | Greek small letter eta tonos     |
|   | Unicode     | 03AE                             |
|   | UTF-8       | CEAE                             |
|   | Language(s) | Greek                            |
| Θ | Name        | Greek capital letter theta       |
|   | Unicode     | 0398                             |
|   | UTF-8       | CE98                             |
|   | Language(s) | Greek                            |

Θ

|                    |                          |
|--------------------|--------------------------|
| <b>Name</b>        | Greek small letter theta |
| <b>Unicode</b>     | 03B8                     |
| <b>UTF-8</b>       | CEB8                     |
| <b>Language(s)</b> | Greek                    |

Ι

|                    |                           |
|--------------------|---------------------------|
| <b>Name</b>        | Greek capital letter iota |
| <b>Unicode</b>     | 0399                      |
| <b>UTF-8</b>       | CE99                      |
| <b>Language(s)</b> | Greek                     |

ι

|                    |                         |
|--------------------|-------------------------|
| <b>Name</b>        | Greek small letter iota |
| <b>Unicode</b>     | 03B9                    |
| <b>UTF-8</b>       | CEB9                    |
| <b>Language(s)</b> | Greek                   |

Ϊ

|                    |                                 |
|--------------------|---------------------------------|
| <b>Name</b>        | Greek capital letter iota tonos |
| <b>Unicode</b>     | 038A                            |
| <b>UTF-8</b>       | CE8A                            |
| <b>Language(s)</b> | Greek                           |

ϊ

|                    |                               |
|--------------------|-------------------------------|
| <b>Name</b>        | Greek small letter iota tonos |
| <b>Unicode</b>     | 03AF                          |
| <b>UTF-8</b>       | CEAF                          |
| <b>Language(s)</b> | Greek                         |

Ϊ̈

|                    |                                     |
|--------------------|-------------------------------------|
| <b>Name</b>        | Greek capital letter iota diaeresis |
| <b>Unicode</b>     | 03AA                                |
| <b>UTF-8</b>       | CEAA                                |
| <b>Language(s)</b> | Greek                               |

ϊ̈

|                    |                                   |
|--------------------|-----------------------------------|
| <b>Name</b>        | Greek small letter iota diaeresis |
| <b>Unicode</b>     | 03CA                              |
| <b>UTF-8</b>       | CF8A                              |
| <b>Language(s)</b> | Greek                             |

Ϊ̈́

|                    |                                         |
|--------------------|-----------------------------------------|
| <b>Name</b>        | Greek small letter iota diaeresis tonos |
| <b>Unicode</b>     | 0390                                    |
| <b>UTF-8</b>       | CE90                                    |
| <b>Language(s)</b> | Greek                                   |



|   |             |                             |
|---|-------------|-----------------------------|
| Κ | Name        | Greek capital letter kappa  |
|   | Unicode     | 039A                        |
|   | UTF-8       | CE9A                        |
|   | Language(s) | Greek                       |
| κ | Name        | Greek small letter kappa    |
|   | Unicode     | 03BA                        |
|   | UTF-8       | CEBA                        |
|   | Language(s) | Greek                       |
| Λ | Name        | Greek capital letter lambda |
|   | Unicode     | 039B                        |
|   | UTF-8       | CE9B                        |
|   | Language(s) | Greek                       |
| λ | Name        | Greek small letter lambda   |
|   | Unicode     | 03BB                        |
|   | UTF-8       | CEBB                        |
|   | Language(s) | Greek                       |
| Μ | Name        | Greek capital letter mu     |
|   | Unicode     | 039C                        |
|   | UTF-8       | CE9C                        |
|   | Language(s) | Greek                       |
| μ | Name        | Greek small letter mu       |
|   | Unicode     | 03BC                        |
|   | UTF-8       | CEBC                        |
|   | Language(s) | Greek                       |
| Ν | Name        | Greek capital letter nu     |
|   | Unicode     | 039D                        |
|   | UTF-8       | CE9D                        |
|   | Language(s) | Greek                       |
| ν | Name        | Greek small letter nu       |
|   | Unicode     | 03BD                        |
|   | UTF-8       | CEBD                        |
|   | Language(s) | Greek                       |

Ξ

|             |                         |
|-------------|-------------------------|
| Name        | Greek capital letter xi |
| Unicode     | 039E                    |
| UTF-8       | CE9E                    |
| Language(s) | Greek                   |

ξ

|             |                       |
|-------------|-----------------------|
| Name        | Greek small letter xi |
| Unicode     | 03BE                  |
| UTF-8       | CEBE                  |
| Language(s) | Greek                 |

Ο

|             |                              |
|-------------|------------------------------|
| Name        | Greek capital letter omicron |
| Unicode     | 039F                         |
| UTF-8       | CE9F                         |
| Language(s) | Greek                        |

ο

|             |                            |
|-------------|----------------------------|
| Name        | Greek small letter omicron |
| Unicode     | 03BF                       |
| UTF-8       | CEBF                       |
| Language(s) | Greek                      |

Ό

|             |                                    |
|-------------|------------------------------------|
| Name        | Greek capital letter omicron tonos |
| Unicode     | 038C                               |
| UTF-8       | CE8C                               |
| Language(s) | Greek                              |

ό

|             |                                  |
|-------------|----------------------------------|
| Name        | Greek small letter omicron tonos |
| Unicode     | 03CC                             |
| UTF-8       | CF8C                             |
| Language(s) | Greek                            |

Π

|             |                         |
|-------------|-------------------------|
| Name        | Greek capital letter pi |
| Unicode     | 03A0                    |
| UTF-8       | CEA0                    |
| Language(s) | Greek                   |

π

|             |                       |
|-------------|-----------------------|
| Name        | Greek small letter pi |
| Unicode     | 03C0                  |
| UTF-8       | CF80                  |
| Language(s) | Greek                 |

Ρ

|             |                          |
|-------------|--------------------------|
| Name        | Greek capital letter rho |
| Unicode     | 03A1                     |
| UTF-8       | CEA1                     |
| Language(s) | Greek                    |

ρ

|             |                        |
|-------------|------------------------|
| Name        | Greek small letter rho |
| Unicode     | 03C1                   |
| UTF-8       | CF81                   |
| Language(s) | Greek                  |

Σ

|             |                            |
|-------------|----------------------------|
| Name        | Greek capital letter sigma |
| Unicode     | 03A3                       |
| UTF-8       | CEA3                       |
| Language(s) | Greek                      |

σ

|             |                          |
|-------------|--------------------------|
| Name        | Greek small letter sigma |
| Unicode     | 03C3                     |
| UTF-8       | CF83                     |
| Language(s) | Greek                    |

ς

|             |                                |
|-------------|--------------------------------|
| Name        | Greek small letter final sigma |
| Unicode     | 03C2                           |
| UTF-8       | CF82                           |
| Language(s) | Greek                          |

Τ

|             |                          |
|-------------|--------------------------|
| Name        | Greek capital letter tau |
| Unicode     | 03A4                     |
| UTF-8       | CEA4                     |
| Language(s) | Greek                    |

τ

|             |                        |
|-------------|------------------------|
| Name        | Greek small letter tau |
| Unicode     | 03C4                   |
| UTF-8       | CF84                   |
| Language(s) | Greek                  |

Υ

|             |                              |
|-------------|------------------------------|
| Name        | Greek capital letter upsilon |
| Unicode     | 03A5                         |
| UTF-8       | CEA5                         |
| Language(s) | Greek                        |

|     |             |                                            |
|-----|-------------|--------------------------------------------|
| ϵ   | Name        | Greek small letter epsilon                 |
|     | Unicode     | 03C5                                       |
|     | UTF-8       | CF85                                       |
|     | Language(s) | Greek                                      |
| Υ   | Name        | Greek capital letter epsilon tonos         |
|     | Unicode     | 038E                                       |
|     | UTF-8       | CE8E                                       |
|     | Language(s) | Greek                                      |
| ϵ̇  | Name        | Greek small letter epsilon tonos           |
|     | Unicode     | 03CD                                       |
|     | UTF-8       | CF8D                                       |
|     | Language(s) | Greek                                      |
| Υ̇  | Name        | Greek capital letter epsilon diaeresis     |
|     | Unicode     | 03AB                                       |
|     | UTF-8       | CEAB                                       |
|     | Language(s) | Greek                                      |
| ϵ̈  | Name        | Greek small letter epsilon daeresis        |
|     | Unicode     | 03CB                                       |
|     | UTF-8       | CF8B                                       |
|     | Language(s) | Greek                                      |
| ϵ̈́ | Name        | Greek small letter epsilon diaeresis tonos |
|     | Unicode     | 03B0                                       |
|     | UTF-8       | CEB0                                       |
|     | Language(s) | Greek                                      |
| Φ   | Name        | Greek capital letter phi                   |
|     | Unicode     | 03A6                                       |
|     | UTF-8       | CEA6                                       |
|     | Language(s) | Greek                                      |
| φ   | Name        | Greek small letter phi                     |
|     | Unicode     | 03C6                                       |
|     | UTF-8       | CF86                                       |
|     | Language(s) | Greek                                      |

|    |             |                                  |
|----|-------------|----------------------------------|
| Χ  | Name        | Greek capital letter chi         |
|    | Unicode     | 03A7                             |
|    | UTF-8       | CEA7                             |
| χ  | Language(s) | Greek                            |
|    | Name        | Greek small letter chi           |
|    | Unicode     | 03C7                             |
| Ψ  | UTF-8       | CF87                             |
|    | Language(s) | Greek                            |
|    | Name        | Greek capital letter psi         |
| ψ  | Unicode     | 03A8                             |
|    | UTF-8       | CEA8                             |
|    | Language(s) | Greek                            |
| Ψ  | Name        | Greek small letter psi           |
|    | Unicode     | 03C8                             |
|    | UTF-8       | CF88                             |
| Ω  | Language(s) | Greek                            |
|    | Name        | Greek capital letter omega       |
|    | Unicode     | 03A9                             |
| ω  | UTF-8       | CEA9                             |
|    | Language(s) | Greek                            |
|    | Name        | Greek small letter omega         |
| 'Ω | Unicode     | 03C9                             |
|    | UTF-8       | CF89                             |
|    | Language(s) | Greek                            |
| Ω̇ | Name        | Greek capital letter omega tonos |
|    | Unicode     | 038F                             |
|    | UTF-8       | CE8F                             |
| ω̇ | Language(s) | Greek                            |
|    | Name        | Greek small letter omega tonos   |
|    | Unicode     | 03CE                             |
|    | UTF-8       | CF8E                             |
|    | Language(s) | Greek                            |

A

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter A                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 0041                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 41                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

a

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter a                                                                                                                                                                                                                                             |
| <b>Unicode</b>     | 0061                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 61                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

Á

|                    |                                                                                  |
|--------------------|----------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter A acute                                                     |
| <b>Unicode</b>     | 00C1                                                                             |
| <b>UTF-8</b>       | C381                                                                             |
| <b>Language(s)</b> | Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Portuguese, Slovak, Spanish |

á

|                    |                                                                                  |
|--------------------|----------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter a acute                                                       |
| <b>Unicode</b>     | 00E1                                                                             |
| <b>UTF-8</b>       | C3A1                                                                             |
| <b>Language(s)</b> | Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Portuguese, Slovak, Spanish |

À

|                    |                                                      |
|--------------------|------------------------------------------------------|
| <b>Name</b>        | Latin capital letter A grave                         |
| <b>Unicode</b>     | 00C0                                                 |
| <b>UTF-8</b>       | C380                                                 |
| <b>Language(s)</b> | Catalan, Dutch, French, Italian, Maltese, Portuguese |

à

|                    |                                                      |
|--------------------|------------------------------------------------------|
| <b>Name</b>        | Latin small letter a grave                           |
| <b>Unicode</b>     | 00E0                                                 |
| <b>UTF-8</b>       | C3A0                                                 |
| <b>Language(s)</b> | Catalan, Dutch, French, Italian, Maltese, Portuguese |

Â

|             |                                              |
|-------------|----------------------------------------------|
| Name        | Latin capital letter A circumflex            |
| Unicode     | 00C2                                         |
| UTF-8       | C382                                         |
| Language(s) | Dutch, French, Portuguese, Romanian, Turkish |

â

|             |                                              |
|-------------|----------------------------------------------|
| Name        | Latin small letter a circumflex              |
| Unicode     | 00E2                                         |
| UTF-8       | C3A2                                         |
| Language(s) | Dutch, French, Portuguese, Romanian, Turkish |

Ã

|             |                              |
|-------------|------------------------------|
| Name        | Latin capital letter A tilde |
| Unicode     | 00C3                         |
| UTF-8       | C383                         |
| Language(s) | Portuguese                   |

ã

|             |                            |
|-------------|----------------------------|
| Name        | Latin small letter a tilde |
| Unicode     | 00E3                       |
| UTF-8       | C3A3                       |
| Language(s) | Portuguese                 |

Ä

|             |                                                   |
|-------------|---------------------------------------------------|
| Name        | Latin capital letter A diaeresis                  |
| Unicode     | 00C4                                              |
| UTF-8       | C384                                              |
| Language(s) | Dutch, Estonian, Finnish, German, Slovak, Swedish |

ä

|             |                                                   |
|-------------|---------------------------------------------------|
| Name        | Latin small letter a diaeresis                    |
| Unicode     | 00E4                                              |
| UTF-8       | C3A4                                              |
| Language(s) | Dutch, Estonian, Finnish, German, Slovak, Swedish |

Å

|             |                                    |
|-------------|------------------------------------|
| Name        | Latin capital letter A ring        |
| Unicode     | 00C5                               |
| UTF-8       | C385                               |
| Language(s) | Danish, Finish, Norwegian, Swedish |

å

|             |                                    |
|-------------|------------------------------------|
| Name        | Latin small letter a ring          |
| Unicode     | 00E5                               |
| UTF-8       | C3A5                               |
| Language(s) | Danish, Finish, Norwegian, Swedish |

Ā

|             |                               |
|-------------|-------------------------------|
| Name        | Latin capital letter A macron |
| Unicode     | 0100                          |
| UTF-8       | C480                          |
| Language(s) | Latvian                       |

ā

|             |                             |
|-------------|-----------------------------|
| Name        | Latin small letter a macron |
| Unicode     | 0101                        |
| UTF-8       | C481                        |
| Language(s) | Latvian                     |

Ă

|             |                              |
|-------------|------------------------------|
| Name        | Latin capital letter A breve |
| Unicode     | 0102                         |
| UTF-8       | C482                         |
| Language(s) | Romanian                     |

ă

|             |                            |
|-------------|----------------------------|
| Name        | Latin small letter a breve |
| Unicode     | 0103                       |
| UTF-8       | C483                       |
| Language(s) | Romanian                   |

Ą

|             |                               |
|-------------|-------------------------------|
| Name        | Latin capital letter A ogonek |
| Unicode     | 0104                          |
| UTF-8       | C484                          |
| Language(s) | Lithuanian, Polish            |

ą

|             |                             |
|-------------|-----------------------------|
| Name        | Latin small letter a ogonek |
| Unicode     | 0105                        |
| UTF-8       | C485                        |
| Language(s) | Lithuanian, Polish          |

Æ

|             |                              |
|-------------|------------------------------|
| Name        | Latin capital letter AE      |
| Unicode     | 00C6                         |
| UTF-8       | C386                         |
| Language(s) | Danish, Icelandic, Norwegian |

æ

|             |                              |
|-------------|------------------------------|
| Name        | latin small letter ae        |
| Unicode     | 00E6                         |
| UTF-8       | C3A6                         |
| Language(s) | Danish, Icelandic, Norwegian |



# B

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter B                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 0042                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 42                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

# b

|                    |                                                                                                                                                                                                                                                                 |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter b                                                                                                                                                                                                                                            |
| <b>Unicode</b>     | 0062                                                                                                                                                                                                                                                            |
| <b>UTF-8</b>       | 62                                                                                                                                                                                                                                                              |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

# C

|                    |                                                                                                                                                                                                                                             |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter C                                                                                                                                                                                                                      |
| <b>Unicode</b>     | 0043                                                                                                                                                                                                                                        |
| <b>UTF-8</b>       | 43                                                                                                                                                                                                                                          |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Finnish, French, Galician, German, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

# c

|                    |                                                                                                                                                                                                                                             |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter c                                                                                                                                                                                                                        |
| <b>Unicode</b>     | 0063                                                                                                                                                                                                                                        |
| <b>UTF-8</b>       | 63                                                                                                                                                                                                                                          |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Finnish, French, Galician, German, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

# Ć

|                    |                              |
|--------------------|------------------------------|
| <b>Name</b>        | Latin capital letter C acute |
| <b>Unicode</b>     | 0106                         |
| <b>UTF-8</b>       | C486                         |
| <b>Language(s)</b> | Croatian, Icelandic, Polish  |

|   |             |                                                         |
|---|-------------|---------------------------------------------------------|
| Ć | Name        | Latin small letter c acute                              |
|   | Unicode     | 0107                                                    |
|   | UTF-8       | C487                                                    |
|   | Language(s) | Croatian, Icelandic, Polish                             |
| Ç | Name        | Latin capital letter C cedilla                          |
|   | Unicode     | 00C7                                                    |
|   | UTF-8       | C387                                                    |
|   | Language(s) | Catalan, French, Portuguese, Turkish                    |
| ç | Name        | Latin small letter c cedilla                            |
|   | Unicode     | 00E7                                                    |
|   | UTF-8       | C3A7                                                    |
|   | Language(s) | Catalan, French, Portuguese, Turkish                    |
| Č | Name        | Latin capital letter C hacek                            |
|   | Unicode     | 010C                                                    |
|   | UTF-8       | C48C                                                    |
|   | Language(s) | Croatian, Czech, Latvian, Lithuanian, Slovak, Slovenian |
| č | Name        | Latin small letter c hacek                              |
|   | Unicode     | 010D                                                    |
|   | UTF-8       | C48D                                                    |
|   | Language(s) | Croatian, Czech, Latvian, Lithuanian, Slovak, Slovenian |
| Ċ | Name        | Latin capital letter C dot                              |
|   | Unicode     | 010A                                                    |
|   | UTF-8       | C48A                                                    |
|   | Language(s) | Maltese                                                 |
| ċ | Name        | Latin small letter c dot                                |
|   | Unicode     | 010B                                                    |
|   | UTF-8       | C48B                                                    |
|   | Language(s) | Maltese                                                 |
| D | Name        | Latin capital letter D                                  |
|   | Unicode     | 0044                                                    |
|   | UTF-8       | 44                                                      |

|   |                    |                                                                                                                                                                                                                                                                  |
|---|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| d | <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|   | <b>Name</b>        | Latin small letter d                                                                                                                                                                                                                                             |
|   | <b>Unicode</b>     | 0064                                                                                                                                                                                                                                                             |
|   | <b>UTF-8</b>       | 64                                                                                                                                                                                                                                                               |
| Ď | <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|   | <b>Name</b>        | Latin capital letter D hacek                                                                                                                                                                                                                                     |
|   | <b>Unicode</b>     | 010E                                                                                                                                                                                                                                                             |
|   | <b>UTF-8</b>       | C48E                                                                                                                                                                                                                                                             |
| ď | <b>Language(s)</b> | Czech, Slovak                                                                                                                                                                                                                                                    |
|   | <b>Name</b>        | Latin small letter d hacek                                                                                                                                                                                                                                       |
|   | <b>Unicode</b>     | 010F                                                                                                                                                                                                                                                             |
|   | <b>UTF-8</b>       | C48F                                                                                                                                                                                                                                                             |
| Ð | <b>Language(s)</b> | Czech, Slovak                                                                                                                                                                                                                                                    |
|   | <b>Name</b>        | Latin capital letter ETH                                                                                                                                                                                                                                         |
|   | <b>Unicode</b>     | 00D0                                                                                                                                                                                                                                                             |
|   | <b>UTF-8</b>       | C390                                                                                                                                                                                                                                                             |
| ð | <b>Language(s)</b> | Icelandic, Faroese                                                                                                                                                                                                                                               |
|   | <b>Name</b>        | Latin small letter eth                                                                                                                                                                                                                                           |
|   | <b>Unicode</b>     | 00F0                                                                                                                                                                                                                                                             |
|   | <b>UTF-8</b>       | C3B0                                                                                                                                                                                                                                                             |
| Đ | <b>Language(s)</b> | Icelandic, Faroese                                                                                                                                                                                                                                               |
|   | <b>Name</b>        | Latin capital letter D with stroke                                                                                                                                                                                                                               |
|   | <b>Unicode</b>     | 0110                                                                                                                                                                                                                                                             |
|   | <b>UTF-8</b>       | C490                                                                                                                                                                                                                                                             |
| Đ | <b>Language(s)</b> | Croatian, Vietnamese, Sámi                                                                                                                                                                                                                                       |

|    |             |                                                                                      |
|----|-------------|--------------------------------------------------------------------------------------|
| đ  | Name        | Latin small letter d with stroke                                                     |
|    | Unicode     | 0111                                                                                 |
|    | UTF-8       | C491                                                                                 |
|    | Language(s) | Croatian, Vietnamese, Sámi                                                           |
| DZ | Name        | Latin capital letter DZ                                                              |
|    | Unicode     | 01F1                                                                                 |
|    | UTF-8       | C7B1                                                                                 |
|    | Language(s) | Hungarian                                                                            |
| Dz | Name        | Latin capital letter D with small letter z                                           |
|    | Unicode     | 01F2                                                                                 |
|    | UTF-8       | C7B2                                                                                 |
|    | Language(s) | Hungarian                                                                            |
| dz | Name        | Latin small letter dz                                                                |
|    | Unicode     | 01F3                                                                                 |
|    | UTF-8       | C7B3                                                                                 |
|    | Language(s) | Hungarian                                                                            |
| DŽ | Name        | Latin capital letter DZ with caron                                                   |
|    | Unicode     | 01C4                                                                                 |
|    | UTF-8       | C784                                                                                 |
|    | Language(s) | Croatian                                                                             |
| Dž | Name        | Latin capital letter D with small letter z with caron                                |
|    | Unicode     | 01C5                                                                                 |
|    | UTF-8       | C785                                                                                 |
|    | Language(s) | Croatian                                                                             |
| dž | Name        | Latin small letter dz with caron                                                     |
|    | Unicode     | 01C6                                                                                 |
|    | UTF-8       | C786                                                                                 |
|    | Language(s) | Croatian                                                                             |
| E  | Name        | Latin capital letter E                                                               |
|    | Unicode     | 0045                                                                                 |
|    | UTF-8       | 45                                                                                   |
|    | Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, |

Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish

e

**Name** Latin small letter e

**Unicode** 0065

**UTF-8** 65

**Language(s)** Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish

É

**Name** Latin capital letter E acute

**Unicode** 00C9

**UTF-8** C389

**Language(s)** Catalan, Czech, Dutch, French, Galician, Hungarian, Icelandic, Irish, Italian, Portuguese, Slovak, Spanish

é

**Name** Latin small letter e acute

**Unicode** 00E9

**UTF-8** C3A9

**Language(s)** Catalan, Czech, Dutch, French, Galician, Hungarian, Icelandic, Irish, Italian, Portuguese, Slovak, Spanish

È

**Name** Latin capital letter E grave

**Unicode** 00C8

**UTF-8** C388

**Language(s)** Catalan, Dutch, French, Italian, Maltese

è

**Name** Latin small letter e grave

**Unicode** 00E8

**UTF-8** C3A8

**Language(s)** Dutch, French, Italian, Maltese

Ê

**Name** Latin capital letter E circumflex

**Unicode** 00CA

**UTF-8** C38A

**Language(s)** Dutch, French, Portuguese

|   |             |                                  |
|---|-------------|----------------------------------|
| ê | Name        | Latin small letter e circumflex  |
|   | Unicode     | 00EA                             |
|   | UTF-8       | C3AA                             |
| Ë | Language(s) | Dutch, French, Portuguese        |
|   | Name        | Latin capital letter E diaeresis |
|   | Unicode     | 00CB                             |
| ë | UTF-8       | C38B                             |
|   | Language(s) | Dutch, French                    |
|   | Name        | Latin small letter e diaeresis   |
| Ě | Unicode     | 00EB                             |
|   | UTF-8       | C3AB                             |
|   | Language(s) | Dutch, French                    |
| Ě | Name        | Latin capital letter E hacek     |
|   | Unicode     | 011A                             |
|   | UTF-8       | C49A                             |
| ě | Language(s) | Czech                            |
|   | Name        | Latin small letter e hacek       |
|   | Unicode     | 011B                             |
| Ē | UTF-8       | C49B                             |
|   | Language(s) | Czech                            |
|   | Name        | Latin capital letter E macron    |
| ē | Unicode     | 0112                             |
|   | UTF-8       | C492                             |
|   | Language(s) | Latvian                          |
| ē | Name        | Latin small letter e macron      |
|   | Unicode     | 0113                             |
|   | UTF-8       | C493                             |
| Ę | Language(s) | Latvian                          |
|   | Name        | Latin capital letter E ogonek    |
|   | Unicode     | 0118                             |
| ę | UTF-8       | C498                             |
|   | Language(s) | Lithuanian, Polish               |

ė

|                    |                             |
|--------------------|-----------------------------|
| <b>Name</b>        | Latin small letter e ogonek |
| <b>Unicode</b>     | 0119                        |
| <b>UTF-8</b>       | C499                        |
| <b>Language(s)</b> | Lithuanian, Polish          |

Ė

|                    |                            |
|--------------------|----------------------------|
| <b>Name</b>        | Latin capital letter E dot |
| <b>Unicode</b>     | 0116                       |
| <b>UTF-8</b>       | C496                       |
| <b>Language(s)</b> | Lithuanian                 |

ė

|                    |                          |
|--------------------|--------------------------|
| <b>Name</b>        | Latin small letter e dot |
| <b>Unicode</b>     | 0117                     |
| <b>UTF-8</b>       | C497                     |
| <b>Language(s)</b> | Lithuanian               |

F

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter F                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 0046                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 46                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

f

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter f                                                                                                                                                                                                                                             |
| <b>Unicode</b>     | 0066                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 66                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

G

|                    |                                                                                                                                                                                       |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter G                                                                                                                                                                |
| <b>Unicode</b>     | 0047                                                                                                                                                                                  |
| <b>UTF-8</b>       | 47                                                                                                                                                                                    |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, |

|   |             |                                                                                                                                                                                                                                                                  |
|---|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| g |             | Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish                                                                                                                                                                                       |
|   | Name        | Latin small letter g                                                                                                                                                                                                                                             |
|   | Unicode     | 0067                                                                                                                                                                                                                                                             |
|   | UTF-8       | 67                                                                                                                                                                                                                                                               |
| Ġ | Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|   | Name        | Latin capital letter G breve                                                                                                                                                                                                                                     |
|   | Unicode     | 011E                                                                                                                                                                                                                                                             |
|   | UTF-8       | C49E                                                                                                                                                                                                                                                             |
| ğ | Language(s) | Turkish                                                                                                                                                                                                                                                          |
|   | Name        | Latin small letter g breve                                                                                                                                                                                                                                       |
|   | Unicode     | 011F                                                                                                                                                                                                                                                             |
|   | UTF-8       | C49F                                                                                                                                                                                                                                                             |
| Ġ | Language(s) | Turkish                                                                                                                                                                                                                                                          |
|   | Name        | Latin capital letter G dot                                                                                                                                                                                                                                       |
|   | Unicode     | 0120                                                                                                                                                                                                                                                             |
|   | UTF-8       | C4A0                                                                                                                                                                                                                                                             |
| ġ | Language(s) | Maltese                                                                                                                                                                                                                                                          |
|   | Name        | Latin small letter g dot                                                                                                                                                                                                                                         |
|   | Unicode     | 0121                                                                                                                                                                                                                                                             |
|   | UTF-8       | C4A1                                                                                                                                                                                                                                                             |
| Ģ | Language(s) | Maltese                                                                                                                                                                                                                                                          |
|   | Name        | Latin capital letter G cedilla                                                                                                                                                                                                                                   |
|   | Unicode     | 0122                                                                                                                                                                                                                                                             |
|   | UTF-8       | C4A2                                                                                                                                                                                                                                                             |
| ģ | Language(s) | Latvian                                                                                                                                                                                                                                                          |
|   | Name        | Latin small letter g cedilla                                                                                                                                                                                                                                     |
|   | Unicode     | 0123                                                                                                                                                                                                                                                             |
|   | UTF-8       | C4A3                                                                                                                                                                                                                                                             |
| ġ | Language(s) | Latvian                                                                                                                                                                                                                                                          |
|   | Name        | Latin capital letter G cedilla                                                                                                                                                                                                                                   |
|   | Unicode     | 0124                                                                                                                                                                                                                                                             |
|   | UTF-8       | C4A4                                                                                                                                                                                                                                                             |



H

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter H                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 0048                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 48                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

h

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter h                                                                                                                                                                                                                                             |
| <b>Unicode</b>     | 0068                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 68                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

Ĥ

|                    |                            |
|--------------------|----------------------------|
| <b>Name</b>        | Latin capital letter H bar |
| <b>Unicode</b>     | 0126                       |
| <b>UTF-8</b>       | C4A6                       |
| <b>Language(s)</b> | Maltese                    |

ħ

|                    |                          |
|--------------------|--------------------------|
| <b>Name</b>        | Latin small letter h bar |
| <b>Unicode</b>     | 0127                     |
| <b>UTF-8</b>       | C4A7                     |
| <b>Language(s)</b> | Maltese                  |

I

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter I                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 0049                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 49                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

i

|                |                      |
|----------------|----------------------|
| <b>Name</b>    | Latin small letter i |
| <b>Unicode</b> | 0069                 |
| <b>UTF-8</b>   | 69                   |

|   |             |                                                                                                                                                                                                                                                                  |
|---|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Í | Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|   | Name        | Latin capital letter I acute                                                                                                                                                                                                                                     |
|   | Unicode     | 00CD                                                                                                                                                                                                                                                             |
|   | UTF-8       | C38D                                                                                                                                                                                                                                                             |
| í | Language(s) | Catalan, Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Italian, Portuguese, Slovak, Spanish                                                                                                                                                               |
|   | Name        | Latin small letter i acute                                                                                                                                                                                                                                       |
|   | Unicode     | 00ED                                                                                                                                                                                                                                                             |
|   | UTF-8       | C3AD                                                                                                                                                                                                                                                             |
| Ì | Language(s) | Catalan, Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Italian, Portuguese, Slovak, Spanish                                                                                                                                                               |
|   | Name        | Latin capital letter I grave                                                                                                                                                                                                                                     |
|   | Unicode     | 00CC                                                                                                                                                                                                                                                             |
|   | UTF-8       | C38C                                                                                                                                                                                                                                                             |
| ì | Language(s) | Dutch, Italian, Maltese                                                                                                                                                                                                                                          |
|   | Name        | Latin small letter i grave                                                                                                                                                                                                                                       |
|   | Unicode     | 00EC                                                                                                                                                                                                                                                             |
|   | UTF-8       | C3AC                                                                                                                                                                                                                                                             |
| Î | Language(s) | Dutch, Italian, Maltese                                                                                                                                                                                                                                          |
|   | Name        | Latin capital letter I circumflex                                                                                                                                                                                                                                |
|   | Unicode     | 00CE                                                                                                                                                                                                                                                             |
|   | UTF-8       | C38E                                                                                                                                                                                                                                                             |
| î | Language(s) | Dutch, French, Romanian                                                                                                                                                                                                                                          |
|   | Name        | Latin small letter i circumflex                                                                                                                                                                                                                                  |
|   | Unicode     | 00EE                                                                                                                                                                                                                                                             |
|   | UTF-8       | C3AE                                                                                                                                                                                                                                                             |
| Ï | Language(s) | Dutch, French, Romanian                                                                                                                                                                                                                                          |
|   | Name        | Latin capital letter I diaeresis                                                                                                                                                                                                                                 |
| ï | Unicode     | 00CF                                                                                                                                                                                                                                                             |

ï

|             |                                |
|-------------|--------------------------------|
| UTF-8       | C38F                           |
| Language(s) | Catalan, Dutch, French         |
| Name        | Latin small letter i diaeresis |

|         |      |
|---------|------|
| Unicode | 00EF |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C3AF |
|-------|------|

|             |                        |
|-------------|------------------------|
| Language(s) | Catalan, Dutch, French |
|-------------|------------------------|

|      |                               |
|------|-------------------------------|
| Name | Latin capital letter I macron |
|------|-------------------------------|

Ī

|         |      |
|---------|------|
| Unicode | 012A |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C4AA |
|-------|------|

|             |         |
|-------------|---------|
| Language(s) | Latvian |
|-------------|---------|

|      |                             |
|------|-----------------------------|
| Name | Latin small letter i macron |
|------|-----------------------------|

Ĭ

|         |      |
|---------|------|
| Unicode | 012B |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C4AB |
|-------|------|

|             |         |
|-------------|---------|
| Language(s) | Latvian |
|-------------|---------|

|      |                               |
|------|-------------------------------|
| Name | Latin capital letter I ogonek |
|------|-------------------------------|

Į

|         |      |
|---------|------|
| Unicode | 012E |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C4AE |
|-------|------|

|             |            |
|-------------|------------|
| Language(s) | Lithuanian |
|-------------|------------|

|      |                             |
|------|-----------------------------|
| Name | Latin small letter i ogonek |
|------|-----------------------------|

İ

|         |      |
|---------|------|
| Unicode | 012F |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C4AF |
|-------|------|

|             |            |
|-------------|------------|
| Language(s) | Lithuanian |
|-------------|------------|

|      |                            |
|------|----------------------------|
| Name | Latin capital letter I dot |
|------|----------------------------|

ı

|         |      |
|---------|------|
| Unicode | 0130 |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C4B0 |
|-------|------|

|             |         |
|-------------|---------|
| Language(s) | Turkish |
|-------------|---------|

|      |                              |
|------|------------------------------|
| Name | Latin small letter dotless i |
|------|------------------------------|

İ

|         |      |
|---------|------|
| Unicode | 0131 |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C4B1 |
|-------|------|

|             |         |
|-------------|---------|
| Language(s) | Turkish |
|-------------|---------|

|      |                         |
|------|-------------------------|
| Name | Latin capital letter IJ |
|------|-------------------------|

IJ

|         |      |
|---------|------|
| Unicode | 0132 |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C4B2 |
|-------|------|

ij

|                    |                       |
|--------------------|-----------------------|
| <b>Language(s)</b> | Dutch                 |
| <b>Name</b>        | Latin small letter ij |
| <b>Unicode</b>     | 0133                  |
| <b>UTF-8</b>       | C4B3                  |

J

|                    |                        |
|--------------------|------------------------|
| <b>Language(s)</b> | Dutch                  |
| <b>Name</b>        | Latin capital letter J |
| <b>Unicode</b>     | 004A                   |
| <b>UTF-8</b>       | 4A                     |

j

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
| <b>Name</b>        | Latin small letter j                                                                                                                                                                                                                                             |
| <b>Unicode</b>     | 006A                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 6A                                                                                                                                                                                                                                                               |

K

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
| <b>Name</b>        | Latin capital letter K                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 004B                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 4B                                                                                                                                                                                                                                                               |

k

|                    |                                                                                                                                                                                                         |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Language(s)</b> | Basque, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Slovak, Slovenian, Swedish, Turkish |
| <b>Name</b>        | Latin small letter k                                                                                                                                                                                    |
| <b>Unicode</b>     | 006B                                                                                                                                                                                                    |
| <b>UTF-8</b>       | 6B                                                                                                                                                                                                      |

K<sub>3</sub>

|             |                                |
|-------------|--------------------------------|
| Name        | Latin capital letter K cedilla |
| Unicode     | 0136                           |
| UTF-8       | C4B6                           |
| Language(s) | Latvian                        |

k<sub>3</sub>

|             |                              |
|-------------|------------------------------|
| Name        | Latin small letter k cedilla |
| Unicode     | 0137                         |
| UTF-8       | C4B7                         |
| Language(s) | Latvian                      |

L

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin capital letter L                                                                                                                                                                                                                                           |
| Unicode     | 004C                                                                                                                                                                                                                                                             |
| UTF-8       | 4C                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

l

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin small letter l                                                                                                                                                                                                                                             |
| Unicode     | 006C                                                                                                                                                                                                                                                             |
| UTF-8       | 6C                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

Ĺ

|             |                              |
|-------------|------------------------------|
| Name        | Latin capital letter L acute |
| Unicode     | 0139                         |
| UTF-8       | C4B9                         |
| Language(s) | Slovak                       |

ĺ

|             |                            |
|-------------|----------------------------|
| Name        | Latin small letter l acute |
| Unicode     | 013A                       |
| UTF-8       | C4BA                       |
| Language(s) | Slovak                     |

L'

|         |                              |
|---------|------------------------------|
| Name    | Latin capital letter L hacek |
| Unicode | 013D                         |
| UTF-8   | C4BD                         |

|    |             |                                   |
|----|-------------|-----------------------------------|
| I' | Language(s) | Slovak                            |
|    | Name        | Latin small letter I hacek        |
|    | Unicode     | 013E                              |
|    | UTF-8       | C4BE                              |
| Ḷ | Language(s) | Slovak                            |
|    | Name        | Latin capital letter L middle dot |
|    | Unicode     | 013F                              |
|    | UTF-8       | C4BF                              |
| ḷ | Language(s) | Catalan                           |
|    | Name        | Latin small letter l middle dot   |
|    | Unicode     | 0140                              |
|    | UTF-8       | C580                              |
| Ł  | Language(s) | Catalan                           |
|    | Name        | Latin capital letter L slash      |
|    | Unicode     | 0141                              |
|    | UTF-8       | C581                              |
| ł  | Language(s) | Polish                            |
|    | Name        | Latin small letter l slash        |
|    | Unicode     | 0142                              |
|    | UTF-8       | C582                              |
| Ļ  | Language(s) | Polish                            |
|    | Name        | Latin capital letter L cedilla    |
|    | Unicode     | 013B                              |
|    | UTF-8       | C4BB                              |
| ļ  | Language(s) | Latvian                           |
|    | Name        | Latin small letter l cedilla      |
|    | Unicode     | 013C                              |
|    | UTF-8       | C4BC                              |
| Ĺ  | Language(s) | Latvian                           |
|    | Name        | Latin capital letter LJ           |
|    | Unicode     | 01C7                              |
|    | UTF-8       | C787                              |
| Ľ  | Language(s) | Croatian                          |

Lj

|             |                                            |
|-------------|--------------------------------------------|
| Name        | Latin capital letter L with small letter j |
| Unicode     | 01C8                                       |
| UTF-8       | C788                                       |
| Language(s) | Croatian                                   |

lj

|             |                       |
|-------------|-----------------------|
| Name        | Latin small letter lj |
| Unicode     | 01C9                  |
| UTF-8       | C789                  |
| Language(s) | Croatian              |

M

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin capital letter M                                                                                                                                                                                                                                           |
| Unicode     | 004D                                                                                                                                                                                                                                                             |
| UTF-8       | 4D                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

m

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin small letter m                                                                                                                                                                                                                                             |
| Unicode     | 006D                                                                                                                                                                                                                                                             |
| UTF-8       | 6D                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

N

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin capital letter N                                                                                                                                                                                                                                           |
| Unicode     | 004E                                                                                                                                                                                                                                                             |
| UTF-8       | 4E                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

n

|         |                      |
|---------|----------------------|
| Name    | Latin small letter n |
| Unicode | 006E                 |
| UTF-8   | 6E                   |

|   |             |                                                                                                                                                                                                                                                                  |
|---|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Ń | Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|   | Name        | Latin capital letter N acute                                                                                                                                                                                                                                     |
|   | Unicode     | 0143                                                                                                                                                                                                                                                             |
|   | UTF-8       | C583                                                                                                                                                                                                                                                             |
| ń | Language(s) | Polish                                                                                                                                                                                                                                                           |
|   | Name        | Latin small letter n acute                                                                                                                                                                                                                                       |
|   | Unicode     | 0144                                                                                                                                                                                                                                                             |
|   | UTF-8       | C584                                                                                                                                                                                                                                                             |
| Ñ | Language(s) | Polish                                                                                                                                                                                                                                                           |
|   | Name        | Latin capital letter N tilde                                                                                                                                                                                                                                     |
|   | Unicode     | 00D1                                                                                                                                                                                                                                                             |
|   | UTF-8       | C391                                                                                                                                                                                                                                                             |
| ñ | Language(s) | Basque, Galician, Spanish                                                                                                                                                                                                                                        |
|   | Name        | Latin small letter n tilde                                                                                                                                                                                                                                       |
|   | Unicode     | 00F1                                                                                                                                                                                                                                                             |
|   | UTF-8       | C3B1                                                                                                                                                                                                                                                             |
| Ņ | Language(s) | Basque, Galician, Spanish                                                                                                                                                                                                                                        |
|   | Name        | Latin capital letter N cedilla                                                                                                                                                                                                                                   |
|   | Unicode     | 0145                                                                                                                                                                                                                                                             |
|   | UTF-8       | C585                                                                                                                                                                                                                                                             |
| ņ | Language(s) | Latvian                                                                                                                                                                                                                                                          |
|   | Name        | Latin small letter n cedilla                                                                                                                                                                                                                                     |
|   | Unicode     | 0146                                                                                                                                                                                                                                                             |
|   | UTF-8       | C587                                                                                                                                                                                                                                                             |
| ň | Language(s) | Latvian                                                                                                                                                                                                                                                          |
|   | Name        | Latin capital letter N hacek                                                                                                                                                                                                                                     |
|   | Unicode     | 0147                                                                                                                                                                                                                                                             |
|   | UTF-8       | C587                                                                                                                                                                                                                                                             |
| Ň | Language(s) | Czech, Slovak                                                                                                                                                                                                                                                    |
|   | Name        | Latin capital letter N cedilla                                                                                                                                                                                                                                   |
|   | Unicode     | 0148                                                                                                                                                                                                                                                             |
|   | UTF-8       | C588                                                                                                                                                                                                                                                             |



|    |             |                                                                                                                                                                                                                                                                  |
|----|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ň  | Name        | Latin small letter n hacek                                                                                                                                                                                                                                       |
|    | Unicode     | 0148                                                                                                                                                                                                                                                             |
|    | UTF-8       | C588                                                                                                                                                                                                                                                             |
| Ň  | Language(s) | Czech, Slovak                                                                                                                                                                                                                                                    |
|    | Name        | Latin capital letter ENG                                                                                                                                                                                                                                         |
|    | Unicode     | 014A                                                                                                                                                                                                                                                             |
| ŋ  | UTF-8       | C58A                                                                                                                                                                                                                                                             |
|    | Language(s) | Sámi, Mende, IPA                                                                                                                                                                                                                                                 |
|    | Name        | Latin small letter eng                                                                                                                                                                                                                                           |
| ŋ  | Unicode     | 014B                                                                                                                                                                                                                                                             |
|    | UTF-8       | C58B                                                                                                                                                                                                                                                             |
|    | Language(s) | Sámi, Mende, IPA                                                                                                                                                                                                                                                 |
| NJ | Name        | Latin capital letter NJ                                                                                                                                                                                                                                          |
|    | Unicode     | 01CA                                                                                                                                                                                                                                                             |
|    | UTF-8       | C78A                                                                                                                                                                                                                                                             |
| Nj | Language(s) | Croatian                                                                                                                                                                                                                                                         |
|    | Name        | Latin capital letter N with small letter j                                                                                                                                                                                                                       |
|    | Unicode     | 01CB                                                                                                                                                                                                                                                             |
| nj | UTF-8       | C78B                                                                                                                                                                                                                                                             |
|    | Language(s) | Croatian                                                                                                                                                                                                                                                         |
|    | Name        | Latin small letter nj                                                                                                                                                                                                                                            |
| nj | Unicode     | 01CC                                                                                                                                                                                                                                                             |
|    | UTF-8       | C78C                                                                                                                                                                                                                                                             |
|    | Language(s) | Croatian                                                                                                                                                                                                                                                         |
| O  | Name        | Latin capital letter O                                                                                                                                                                                                                                           |
|    | Unicode     | 004F                                                                                                                                                                                                                                                             |
|    | UTF-8       | 4F                                                                                                                                                                                                                                                               |
| O  | Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|    | Name        | Latin small letter o                                                                                                                                                                                                                                             |
|    | Unicode     | 006F                                                                                                                                                                                                                                                             |

|   |             |                                                                                                                                                                                                                                                                  |
|---|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Ó | UTF-8       | 6F                                                                                                                                                                                                                                                               |
|   | Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|   | Name        | Latin capital letter O acute                                                                                                                                                                                                                                     |
|   | Unicode     | 00D3                                                                                                                                                                                                                                                             |
| ó | UTF-8       | C393                                                                                                                                                                                                                                                             |
|   | Language(s) | Catalan, Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Italian, Polish, Portuguese, Slovak, Spanish                                                                                                                                                       |
|   | Name        | Latin small letter o acute                                                                                                                                                                                                                                       |
|   | Unicode     | 00F3                                                                                                                                                                                                                                                             |
| Ò | UTF-8       | C3B3                                                                                                                                                                                                                                                             |
|   | Language(s) | Catalan, Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Italian, Polish, Portuguese, Slovak, Spanish                                                                                                                                                       |
|   | Name        | Latin capital letter O grave                                                                                                                                                                                                                                     |
|   | Unicode     | 00D2                                                                                                                                                                                                                                                             |
| ò | UTF-8       | C392                                                                                                                                                                                                                                                             |
|   | Language(s) | Catalan, Dutch, Italian, Maltese                                                                                                                                                                                                                                 |
|   | Name        | Latin small letter o grave                                                                                                                                                                                                                                       |
|   | Unicode     | 00F2                                                                                                                                                                                                                                                             |
| Ô | UTF-8       | C3B2                                                                                                                                                                                                                                                             |
|   | Language(s) | Catalan, Dutch, Italian, Maltese                                                                                                                                                                                                                                 |
|   | Name        | Latin capital letter O circumflex                                                                                                                                                                                                                                |
|   | Unicode     | 00D4                                                                                                                                                                                                                                                             |
| ô | UTF-8       | C394                                                                                                                                                                                                                                                             |
|   | Language(s) | Dutch, French, Portuguese, Slovak                                                                                                                                                                                                                                |
|   | Name        | Latin small letter o circumflex                                                                                                                                                                                                                                  |
|   | Unicode     | 00F4                                                                                                                                                                                                                                                             |
| Õ | UTF-8       | C3B4                                                                                                                                                                                                                                                             |
|   | Language(s) | Dutch, French, Portuguese, Slovak                                                                                                                                                                                                                                |
|   | Name        | Latin capital letter O tilde                                                                                                                                                                                                                                     |

Õ

|             |                              |
|-------------|------------------------------|
| Name        | Latin capital letter O tilde |
| Unicode     | 00D5                         |
| UTF-8       | C395                         |
| Language(s) | Estonian, Portuguese         |

õ

|             |                            |
|-------------|----------------------------|
| Name        | Latin small letter o tilde |
| Unicode     | 00F5                       |
| UTF-8       | C3B5                       |
| Language(s) | Estonian, Portuguese       |

Ö

|             |                                                                          |
|-------------|--------------------------------------------------------------------------|
| Name        | Latin capital letter O diaeresis                                         |
| Unicode     | 00D6                                                                     |
| UTF-8       | C396                                                                     |
| Language(s) | Dutch, Estonian, Finnish, German, Hungarian, Icelandic, Swedish, Turkish |

ö

|             |                                                                          |
|-------------|--------------------------------------------------------------------------|
| Name        | Latin small letter o diaeresis                                           |
| Unicode     | 00F6                                                                     |
| UTF-8       | C3B6                                                                     |
| Language(s) | Dutch, Estonian, Finnish, German, Hungarian, Icelandic, Swedish, Turkish |

Ő

|             |                                     |
|-------------|-------------------------------------|
| Name        | Latin capital letter O double acute |
| Unicode     | 0150                                |
| UTF-8       | C590                                |
| Language(s) | Hungarian                           |

ő

|             |                                   |
|-------------|-----------------------------------|
| Name        | Latin small letter o double acute |
| Unicode     | 0151                              |
| UTF-8       | C591                              |
| Language(s) | Hungarian                         |

Ø

|             |                              |
|-------------|------------------------------|
| Name        | Latin capital letter O slash |
| Unicode     | 00D8                         |
| UTF-8       | C398                         |
| Language(s) | Danish, Norwegian            |

ø

|             |                            |
|-------------|----------------------------|
| Name        | Latin small letter o slash |
| Unicode     | 00F8                       |
| UTF-8       | C3B8                       |
| Language(s) | Danish, Norwegian          |

Ō

|             |                               |
|-------------|-------------------------------|
| Name        | Latin capital letter O macron |
| Unicode     | 014C                          |
| UTF-8       | C58C                          |
| Language(s) | Latvian                       |

ō

|             |                             |
|-------------|-----------------------------|
| Name        | Latin small letter o macron |
| Unicode     | 014D                        |
| UTF-8       | C58D                        |
| Language(s) | Latvian                     |

Œ

|             |                         |
|-------------|-------------------------|
| Name        | Latin capital letter OE |
| Unicode     | 0152                    |
| UTF-8       | C592                    |
| Language(s) | French                  |

œ

|             |                       |
|-------------|-----------------------|
| Name        | Latin small letter oe |
| Unicode     | 0153                  |
| UTF-8       | C593                  |
| Language(s) | French                |

P

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin capital letter P                                                                                                                                                                                                                                           |
| Unicode     | 0050                                                                                                                                                                                                                                                             |
| UTF-8       | 50                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

p

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin small letter p                                                                                                                                                                                                                                             |
| Unicode     | 0070                                                                                                                                                                                                                                                             |
| UTF-8       | 70                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

Q

|         |                        |
|---------|------------------------|
| Name    | Latin capital letter Q |
| Unicode | 0051                   |
| UTF-8   | 51                     |

q

**Language(s)** Basque, Catalan, Danish, Dutch, English, Finish, French, Galician, German, Hungarian, Irish, Italian, Maltese, Norwegian, Portuguese, Spanish

**Name** Latin small letter q

**Unicode** 0071

**UTF-8** 71

**Language(s)** Basque, Catalan, Danish, Dutch, English, Finish, French, Galician, German, Hungarian, Irish, Italian, Maltese, Norwegian, Portuguese, Spanish

**Name** Latin capital letter R

**Unicode** 0052

**UTF-8** 52

R

**Language(s)** Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish

**Name** Latin small letter r

**Unicode** 0072

**UTF-8** 72

r

**Language(s)** Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish

**Name** Latin capital letter R acute

**Unicode** 0154

**UTF-8** C594

**Language(s)** Slovak

Ř

**Name** Latin small letter r acute

**Unicode** 0155

**UTF-8** C595

**Language(s)** Slovak

ř

Ř

|             |                              |
|-------------|------------------------------|
| Name        | Latin capital letter R hacek |
| Unicode     | 0158                         |
| UTF-8       | C598                         |
| Language(s) | Czech                        |

ř

|             |                            |
|-------------|----------------------------|
| Name        | Latin small letter r hacek |
| Unicode     | 0159                       |
| UTF-8       | C599                       |
| Language(s) | Czech                      |

Ŕ

|             |                                |
|-------------|--------------------------------|
| Name        | Latin capital letter R cedilla |
| Unicode     | 0156                           |
| UTF-8       | C596                           |
| Language(s) | Latvian                        |

ŗ

|             |                              |
|-------------|------------------------------|
| Name        | Latin small letter r cedilla |
| Unicode     | 0157                         |
| UTF-8       | C597                         |
| Language(s) | Latvian                      |

Š

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin capital letter S                                                                                                                                                                                                                                           |
| Unicode     | 0053                                                                                                                                                                                                                                                             |
| UTF-8       | 53                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

š

|             |                                                                                                                                                                                                                                                                  |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Name        | Latin small letter s                                                                                                                                                                                                                                             |
| Unicode     | 0073                                                                                                                                                                                                                                                             |
| UTF-8       | 73                                                                                                                                                                                                                                                               |
| Language(s) | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

Ś

|         |                              |
|---------|------------------------------|
| Name    | Latin capital letter S acute |
| Unicode | 015A                         |
| UTF-8   | 059A                         |

|   |             |                                                                              |
|---|-------------|------------------------------------------------------------------------------|
| ś | Language(s) | Polish                                                                       |
|   | Name        | Latin small letter s acute                                                   |
|   | Unicode     | 015B                                                                         |
|   | UTF-8       | 059B                                                                         |
| Œ | Language(s) | Polish                                                                       |
|   | Name        | Latin capital letter S cedilla                                               |
|   | Unicode     | 015E                                                                         |
|   | UTF-8       | C59E                                                                         |
| Ÿ | Language(s) | (Romanian [should not be used]) Turkish                                      |
|   | Name        | Latin small letter s cedilla                                                 |
|   | Unicode     | 015F                                                                         |
|   | UTF-8       | C59F                                                                         |
| Ș | Language(s) | (Romanian [should not be used]) Turkish                                      |
|   | Name        | Latin capital letter S with comma below                                      |
|   | Unicode     | 0218                                                                         |
|   | UTF-8       | C898                                                                         |
| ș | Language(s) | Romanian                                                                     |
|   | Name        | Latin small letter s with comma below                                        |
|   | Unicode     | 0219                                                                         |
|   | UTF-8       | C899                                                                         |
| Ș | Language(s) | Romanian                                                                     |
|   | Name        | Latin capital letter S hacek                                                 |
|   | Unicode     | 0160                                                                         |
|   | UTF-8       | C5A0                                                                         |
| š | Language(s) | Croatian, Czech, Estonian, [Finnish], Latvian, Lithuanian, Slovak, Slovenian |
|   | Name        | Latin small letter s hacek                                                   |
|   | Unicode     | 0161                                                                         |
|   | UTF-8       | C5A1                                                                         |
| ß | Language(s) | Croatian, Czech, Estonian, [Finnish], Latvian, Lithuanian, Slovak, Slovenian |
|   | Name        | Latin small letter sharp s                                                   |
|   | Unicode     | 00DF                                                                         |
|   | UTF-8       | C39F                                                                         |

T

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Language(s)</b> | German                                                                                                                                                                                                                                                           |
| <b>Name</b>        | Latin capital letter T                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 0054                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 54                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

t

|                |                      |
|----------------|----------------------|
| <b>Name</b>    | Latin small letter t |
| <b>Unicode</b> | 0074                 |
| <b>UTF-8</b>   | 74                   |

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Ţ

|                    |                                 |
|--------------------|---------------------------------|
| <b>Name</b>        | Latin capital letter T cedilla  |
| <b>Unicode</b>     | 0162                            |
| <b>UTF-8</b>       | C5A2                            |
| <b>Language(s)</b> | (Romanian [should not be used]) |

ţ

|                |                              |
|----------------|------------------------------|
| <b>Name</b>    | Latin small letter t cedilla |
| <b>Unicode</b> | 0163                         |
| <b>UTF-8</b>   | C5A3                         |

|                    |                                 |
|--------------------|---------------------------------|
| <b>Language(s)</b> | (Romanian [should not be used]) |
|--------------------|---------------------------------|

Ț

|                |                                         |
|----------------|-----------------------------------------|
| <b>Name</b>    | Latin capital letter T with comma below |
| <b>Unicode</b> | 021A                                    |
| <b>UTF-8</b>   | C89A                                    |

|                    |          |
|--------------------|----------|
| <b>Language(s)</b> | Romanian |
|--------------------|----------|

ț

|                |                                       |
|----------------|---------------------------------------|
| <b>Name</b>    | Latin small letter t with comma below |
| <b>Unicode</b> | 021B                                  |
| <b>UTF-8</b>   | C89B                                  |

|                    |          |
|--------------------|----------|
| <b>Language(s)</b> | Romanian |
|--------------------|----------|



Ť

|                    |                              |
|--------------------|------------------------------|
| <b>Name</b>        | Latin capital letter T hacek |
| <b>Unicode</b>     | 0164                         |
| <b>UTF-8</b>       | C5A4                         |
| <b>Language(s)</b> | Czech, Slovak                |

ť

|                    |                            |
|--------------------|----------------------------|
| <b>Name</b>        | Latin small letter t hacek |
| <b>Unicode</b>     | 0165                       |
| <b>UTF-8</b>       | C5A5                       |
| <b>Language(s)</b> | Czech, Slovak              |

Þ

|                    |                            |
|--------------------|----------------------------|
| <b>Name</b>        | Latin capital letter THORN |
| <b>Unicode</b>     | 00DE                       |
| <b>UTF-8</b>       | C39E                       |
| <b>Language(s)</b> | Icelandic                  |

þ

|                    |                          |
|--------------------|--------------------------|
| <b>Name</b>        | Latin small letter thorn |
| <b>Unicode</b>     | 00FE                     |
| <b>UTF-8</b>       | C3BE                     |
| <b>Language(s)</b> | Icelandic                |

U

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter U                                                                                                                                                                                                                                           |
| <b>Unicode</b>     | 0055                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 55                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

u

|                    |                                                                                                                                                                                                                                                                  |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter u                                                                                                                                                                                                                                             |
| <b>Unicode</b>     | 0075                                                                                                                                                                                                                                                             |
| <b>UTF-8</b>       | 75                                                                                                                                                                                                                                                               |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

Ú

|                |                              |
|----------------|------------------------------|
| <b>Name</b>    | Latin capital letter U acute |
| <b>Unicode</b> | 00DA                         |
| <b>UTF-8</b>   | C39A                         |

ú

|                    |                                                                                                    |
|--------------------|----------------------------------------------------------------------------------------------------|
| <b>Language(s)</b> | Catalan, Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Italian, Portuguese, Slovak, Spanish |
| <b>Name</b>        | Latin small letter u acute                                                                         |
| <b>Unicode</b>     | 00FA                                                                                               |
| <b>UTF-8</b>       | C3BA                                                                                               |
| <b>Language(s)</b> | Catalan, Czech, Dutch, Galician, Hungarian, Icelandic, Irish, Italian, Portuguese, Slovak, Spanish |

Ù

|                |                              |
|----------------|------------------------------|
| <b>Name</b>    | Latin capital letter U grave |
| <b>Unicode</b> | 00D9                         |
| <b>UTF-8</b>   | C399                         |

ù

|                    |                                 |
|--------------------|---------------------------------|
| <b>Language(s)</b> | Dutch, French, Italian, Maltese |
| <b>Name</b>        | Latin small letter U grave      |
| <b>Unicode</b>     | 00F9                            |
| <b>UTF-8</b>       | C3B9                            |
| <b>Language(s)</b> | Dutch, French, Italian, Maltese |

Û

|                    |                                   |
|--------------------|-----------------------------------|
| <b>Name</b>        | Latin capital letter U circumflex |
| <b>Unicode</b>     | 00DB                              |
| <b>UTF-8</b>       | C39B                              |
| <b>Language(s)</b> | Dutch, French, Turkish            |

û

|                    |                                 |
|--------------------|---------------------------------|
| <b>Name</b>        | Latin small letter U circumflex |
| <b>Unicode</b>     | 00FB                            |
| <b>UTF-8</b>       | C3BB                            |
| <b>Language(s)</b> | Dutch, French, Turkish          |

Ü

|                    |                                                                       |
|--------------------|-----------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter U diaeresis                                      |
| <b>Unicode</b>     | 00DC                                                                  |
| <b>UTF-8</b>       | C39C                                                                  |
| <b>Language(s)</b> | Catalan, Dutch, Estonian, French, German, Hungarian, Spanish, Turkish |

ü

|                    |                                                                       |
|--------------------|-----------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter u diaeresis                                        |
| <b>Unicode</b>     | 00FC                                                                  |
| <b>UTF-8</b>       | C3BC                                                                  |
| <b>Language(s)</b> | Catalan, Dutch, Estonian, French, German, Hungarian, Spanish, Turkish |

Ú

|             |                                     |
|-------------|-------------------------------------|
| Name        | Latin capital letter U double acute |
| Unicode     | 0170                                |
| UTF-8       | C5B0                                |
| Language(s) | Hungarian                           |

ú

|             |                                   |
|-------------|-----------------------------------|
| Name        | Latin small letter u double acute |
| Unicode     | 0171                              |
| UTF-8       | C5B1                              |
| Language(s) | Hungarian                         |

Ů

|             |                             |
|-------------|-----------------------------|
| Name        | Latin capital letter U ring |
| Unicode     | 016E                        |
| UTF-8       | C5AE                        |
| Language(s) | Czech                       |

ů

|             |                           |
|-------------|---------------------------|
| Name        | Latin small letter u ring |
| Unicode     | 016F                      |
| UTF-8       | C5AF                      |
| Language(s) | Czech                     |

Ū

|             |                               |
|-------------|-------------------------------|
| Name        | Latin capital letter U macron |
| Unicode     | 016A                          |
| UTF-8       | C5AA                          |
| Language(s) | Latvian, Lithuanian           |

ū

|             |                             |
|-------------|-----------------------------|
| Name        | Latin small letter u macron |
| Unicode     | 016B                        |
| UTF-8       | C5AB                        |
| Language(s) | Latvian, Lithuanian         |

Ų

|             |                               |
|-------------|-------------------------------|
| Name        | Latin capital letter U ogonek |
| Unicode     | 0172                          |
| UTF-8       | C5B2                          |
| Language(s) | Lithuanian                    |

ų

|             |                             |
|-------------|-----------------------------|
| Name        | Latin small letter u ogonek |
| Unicode     | 0173                        |
| UTF-8       | C5B3                        |
| Language(s) | Lithuanian                  |

V

|                    |                                                                                                                                                                                                                                                          |
|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter V                                                                                                                                                                                                                                   |
| <b>Unicode</b>     | 0056                                                                                                                                                                                                                                                     |
| <b>UTF-8</b>       | 56                                                                                                                                                                                                                                                       |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

v

|                    |                                                                                                                                                                                                                                                          |
|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter v                                                                                                                                                                                                                                     |
| <b>Unicode</b>     | 0076                                                                                                                                                                                                                                                     |
| <b>UTF-8</b>       | 76                                                                                                                                                                                                                                                       |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

W

|                    |                                                                                                                                          |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter W                                                                                                                   |
| <b>Unicode</b>     | 0057                                                                                                                                     |
| <b>UTF-8</b>       | 57                                                                                                                                       |
| <b>Language(s)</b> | Basque, Croatian, Danish, Dutch, English, Finish, French, German, Hungarian, Irish, Italian, Maltese, Norwegian, Polish, Slovak, Swedish |

w

|                    |                                                                                                                                          |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter w                                                                                                                     |
| <b>Unicode</b>     | 0077                                                                                                                                     |
| <b>UTF-8</b>       | 77                                                                                                                                       |
| <b>Language(s)</b> | Basque, Croatian, Danish, Dutch, English, Finish, French, German, Hungarian, Irish, Italian, Maltese, Norwegian, Polish, Slovak, Swedish |

X

|                    |                                                                                                                                                                                                |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter X                                                                                                                                                                         |
| <b>Unicode</b>     | 0058                                                                                                                                                                                           |
| <b>UTF-8</b>       | 58                                                                                                                                                                                             |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Danish, Dutch, English, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Maltese, Norwegian, Portuguese, Romanian, Slovak, Swedish, Spanish |

X

|                    |                                                                                                                                                                                                |
|--------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter x                                                                                                                                                                           |
| <b>Unicode</b>     | 0078                                                                                                                                                                                           |
| <b>UTF-8</b>       | 78                                                                                                                                                                                             |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Danish, Dutch, English, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Italian, Maltese, Norwegian, Portuguese, Romanian, Slovak, Swedish, Spanish |

Y

|                    |                                                                                                                                                                                                               |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter Y                                                                                                                                                                                        |
| <b>Unicode</b>     | 0059                                                                                                                                                                                                          |
| <b>UTF-8</b>       | 59                                                                                                                                                                                                            |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Latvian, Lithuanian, Norwegian, Polish, Slovak, Spanish, Swedish, Turkish |

y

|                    |                                                                                                                                                                                                               |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter y                                                                                                                                                                                          |
| <b>Unicode</b>     | 0079                                                                                                                                                                                                          |
| <b>UTF-8</b>       | 79                                                                                                                                                                                                            |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Icelandic, Irish, Latvian, Lithuanian, Norwegian, Polish, Slovak, Spanish, Swedish, Turkish |

Ý

|                    |                              |
|--------------------|------------------------------|
| <b>Name</b>        | Latin capital letter Y acute |
| <b>Unicode</b>     | 00DD                         |
| <b>UTF-8</b>       | C39D                         |
| <b>Language(s)</b> | Czech, Icelandic, Slovak     |

ý

|                    |                            |
|--------------------|----------------------------|
| <b>Name</b>        | Latin small letter y acute |
| <b>Unicode</b>     | 00FD                       |
| <b>UTF-8</b>       | C3BD                       |
| <b>Language(s)</b> | Czech, Icelandic, Slovak   |

ÿ

|                    |                                  |
|--------------------|----------------------------------|
| <b>Name</b>        | Latin capital letter Y diaeresis |
| <b>Unicode</b>     | 0178                             |
| <b>UTF-8</b>       | C3BF                             |
| <b>Language(s)</b> | French                           |

ÿ

|                    |                                |
|--------------------|--------------------------------|
| <b>Name</b>        | Latin small letter y diaeresis |
| <b>Unicode</b>     | 00FF                           |
| <b>UTF-8</b>       | C3BF                           |
| <b>Language(s)</b> | French                         |

Z

|                    |                                                                                                                                                                                                                                           |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter Z                                                                                                                                                                                                                    |
| <b>Unicode</b>     | 005A                                                                                                                                                                                                                                      |
| <b>UTF-8</b>       | 5A                                                                                                                                                                                                                                        |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Irish, Italian, Latvian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

z

|                    |                                                                                                                                                                                                                                           |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Name</b>        | Latin small letter z                                                                                                                                                                                                                      |
| <b>Unicode</b>     | 007A                                                                                                                                                                                                                                      |
| <b>UTF-8</b>       | 7A                                                                                                                                                                                                                                        |
| <b>Language(s)</b> | Basque, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Hungarian, Irish, Italian, Latvian, Maltese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, Turkish |

Ź

|                    |                              |
|--------------------|------------------------------|
| <b>Name</b>        | Latin capital letter Z acute |
| <b>Unicode</b>     | 0179                         |
| <b>UTF-8</b>       | C5B9                         |
| <b>Language(s)</b> | Polish                       |

ź

|                    |                            |
|--------------------|----------------------------|
| <b>Name</b>        | Latin small letter z acute |
| <b>Unicode</b>     | 017A                       |
| <b>UTF-8</b>       | C5BA                       |
| <b>Language(s)</b> | Polish                     |

Ž

|                    |                                                                              |
|--------------------|------------------------------------------------------------------------------|
| <b>Name</b>        | Latin capital letter Z hacek                                                 |
| <b>Unicode</b>     | 017D                                                                         |
| <b>UTF-8</b>       | C5BD                                                                         |
| <b>Language(s)</b> | Croatian, Czech, Estonian, [Finnish], Latvian, Lithuanian, Slovak, Slovenian |

ž

|                |                            |
|----------------|----------------------------|
| <b>Name</b>    | Latin small letter z hacek |
| <b>Unicode</b> | 017E                       |

Ž

|             |                                                                              |
|-------------|------------------------------------------------------------------------------|
| UTF-8       | C5BE                                                                         |
| Language(s) | Croatian, Czech, Estonian, [Finnish], Latvian, Lithuanian, Slovak, Slovenian |

|      |                            |
|------|----------------------------|
| Name | Latin capital letter Z dot |
|------|----------------------------|

|         |      |
|---------|------|
| Unicode | 017B |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C5BB |
|-------|------|

|             |                 |
|-------------|-----------------|
| Language(s) | Maltese, Polish |
|-------------|-----------------|

|      |                          |
|------|--------------------------|
| Name | Latin small letter z dot |
|------|--------------------------|

|         |      |
|---------|------|
| Unicode | 017C |
|---------|------|

|       |      |
|-------|------|
| UTF-8 | C5BC |
|-------|------|

|             |                 |
|-------------|-----------------|
| Language(s) | Maltese, Polish |
|-------------|-----------------|

ž

## Arrows

|   |         |                  |
|---|---------|------------------|
| ← | Name    | left arrow       |
|   | Unicode | 2190             |
|   | UTF-8   | E28690           |
| ↑ | Name    | up arrow         |
|   | Unicode | 2191             |
|   | UTF-8   | E28691           |
| → | Name    | right arrow      |
|   | Unicode | 2192             |
|   | UTF-8   | E28692           |
| ↓ | Name    | down arrow       |
|   | Unicode | 2193             |
|   | UTF-8   | E28693           |
| ↔ | Name    | left right arrow |
|   | Unicode | 2194             |
|   | UTF-8   | E28694           |
| ↕ | Name    | up down arrow    |
|   | Unicode | 2195             |
|   |         |                  |

|   |                |                         |
|---|----------------|-------------------------|
|   | <b>UTF-8</b>   | E28695                  |
| ↖ | <b>Name</b>    | upper left arrow        |
|   | <b>Unicode</b> | 2196                    |
|   | <b>UTF-8</b>   | E28696                  |
| ↗ | <b>Name</b>    | upper right arrow       |
|   | <b>Unicode</b> | 2197                    |
|   | <b>UTF-8</b>   | E28697                  |
| ↘ | <b>Name</b>    | lower right arrow       |
|   | <b>Unicode</b> | 2198                    |
|   | <b>UTF-8</b>   | E28698                  |
| ↙ | <b>Name</b>    | lower left arrow        |
|   | <b>Unicode</b> | 2199                    |
|   | <b>UTF-8</b>   | Es8699                  |
| ⇐ | <b>Name</b>    | left double arrow       |
|   | <b>Unicode</b> | 21D0                    |
|   | <b>UTF-8</b>   | E28790                  |
| ⇑ | <b>Name</b>    | up double arrow         |
|   | <b>Unicode</b> | 21D1                    |
|   | <b>UTF-8</b>   | E28791                  |
| ⇒ | <b>Name</b>    | right double arrow      |
|   | <b>Unicode</b> | 21D2                    |
|   | <b>UTF-8</b>   | E28792                  |
| ⇓ | <b>Name</b>    | down double arrow       |
|   | <b>Unicode</b> | 21D3                    |
|   | <b>UTF-8</b>   | E28793                  |
| ↔ | <b>Name</b>    | left right double arrow |
|   | <b>Unicode</b> | 21D4                    |
|   | <b>UTF-8</b>   | E28794                  |



|                                                                                     |                |                          |
|-------------------------------------------------------------------------------------|----------------|--------------------------|
|    | <b>Name</b>    | up down double arrow     |
|                                                                                     | <b>Unicode</b> | 21D5                     |
|                                                                                     | <b>UTF-8</b>   | E28795                   |
|    | <b>Name</b>    | upper left double arrow  |
|                                                                                     | <b>Unicode</b> | 21D6                     |
|                                                                                     | <b>UTF-8</b>   | E28796                   |
|    | <b>Name</b>    | upper right double arrow |
|                                                                                     | <b>Unicode</b> | 21D7                     |
|                                                                                     | <b>UTF-8</b>   | E28797                   |
|    | <b>Name</b>    | lower right double arrow |
|                                                                                     | <b>Unicode</b> | 21D8                     |
|                                                                                     | <b>UTF-8</b>   | E28798                   |
|  | <b>Name</b>    | lower left double arrow  |
|                                                                                     | <b>Unicode</b> | 21D9                     |
|                                                                                     | <b>UTF-8</b>   | E28799                   |
|  | <b>Name</b>    | white left arrow         |
|                                                                                     | <b>Unicode</b> | 21E6                     |
|                                                                                     | <b>UTF-8</b>   | E287A6                   |
|  | <b>Name</b>    | white up arrow           |
|                                                                                     | <b>Unicode</b> | 21E7                     |
|                                                                                     | <b>UTF-8</b>   | E287A7                   |
|  | <b>Name</b>    | white right arrow        |
|                                                                                     | <b>Unicode</b> | 21E8                     |
|                                                                                     | <b>UTF-8</b>   | E287A8                   |
|  | <b>Name</b>    | white down arrow         |
|                                                                                     | <b>Unicode</b> | 21E9                     |
|                                                                                     | <b>UTF-8</b>   | E287A9                   |

|                                                                                   |                |                         |
|-----------------------------------------------------------------------------------|----------------|-------------------------|
|  | <b>Name</b>    | delete to the right key |
|                                                                                   | <b>Unicode</b> | 2326                    |
|                                                                                   | <b>UTF-8</b>   | E28CA6                  |
|  | <b>Name</b>    | delete to the left key  |
|                                                                                   | <b>Unicode</b> | 232B                    |
|                                                                                   | <b>UTF-8</b>   | E28CAB                  |

## Processing instructions

### *General use*

In general, the use of processing instructions is forbidden. The only exceptions are those mentioned in this document. So if any provider uses processing instructions for technical or other reasons which do not comply with the specifications in this document, they have to be removed from the instances before transferring them to OPOCE.

All attributes which might be declared within processing instructions must have their values between double quotes.

### *The XML declaration*

Although the XML declaration is not a processing instruction in the strict sense of the term, it has a similar format. Therefore we handle it in this chapter.

Each instance which is transferred to OPOCE has to start with the XML declaration. It specifies the version of the XML specification and the character encoding used for the document. These items are to be provided in form of so-called pseudo-attributes.

The declaration is case dependent. It always starts with the keyword "xml".

The current version of the XML specification is specified as an attribute named "version". The only value accepted is "1.0". If in the course of time W3C decides to adopt a new version of XML, it is up to the OPOCE to decide when and how this version will be used in the context of Formex.

The character encoding is specified by means of the attribute "encoding". The only value accepted is "utf-8". Any other value may only be used with written confirmation by OPOCE.

The XML declaration looks like this (attribute values must always be quoted using the double quote character ("")):

```
<?xml version="1.0" encoding="utf-8"?>
```

*NB:* Other attributes which are foreseen in the XML specification are not accepted.

### *Page breaks*

The processing instruction PAGE is used to mark up the page breaks in an instance. Therefore, it may appear anywhere. The processing instruction has three attributes, "NO", "ID.VOL" and "PAGE.SIZE". The first one is mandatory for each occurrence of the processing instruction.

The first page break must be inserted at the bottom of the first page of a given document. The value given to the NO attribute is the number of the following page. The processing instruction must be repeated at the bottom of every page, with the exception of the last page.

In the case where a PAGE processing instruction has to be added at the end of a structural element, as far as possible it has to be placed outside all elements which logically belong to the same page.

If the processing instruction PAGE appears within a text block between two words, it has to be followed by a space. No spaces may be added, if it appears outside a text block. The objective of this rule is to get a valid document by simply deleting the processing instructions.

*Attention:* The previous rule is different from Formex v3.

When the page break falls within a divided word, the processing instruction must be placed after the word from which the caesura has been removed.

Page breaks in inclusions (INCL.ELEMENT) must also be marked up using the PAGE processing instruction. When a page break occurs at the end of the included document, the PAGE processing instruction will be inserted not in the included document, but just after the inclusion element in the calling document. The value of the NO attribute is the number of the page following the last page of the inclusion.

Specific processing is required for tables which cover a double page (format A3), for example tables relating to Taric. In these cases, the processing instruction is only used every two pages, i.e. at the bottom of right-hand pages. In addition, the attribute PAGE.SIZE has to be included with the value "A3", the value "A4" being the default.

For documents published over two or more volumes the ID.VOL attribute is used to indicate the identifier of the volume containing the page.

The processing instruction must not be used within a page which collects footnotes at the end of the document.

## **Column page break**

The COL.PAGE processing instruction is used to mark up the page break in a column of a table when the contents of the cell cover two pages.

It is also present in other models which are similar to table or column structures.

The COL.PAGE processing instruction is used when the contents of the cell cover two pages. It is usually inserted after the tag indicating the end of the line.

The use of COL.PAGE does not under any circumstances replace the PAGE processing instruction.

```
<ROW>
  <CELL COL="1">HITACHI</CELL>
  <CELL COL="2">
    <P>SK-H5</P>
    <P>SK-H501</P>
    <P>DK-7700</P>
    <!-- ... -->
    <P>FP-C10 (R) </P><?COL.PAGE?>
    <P>FP-C10 (S) </P>
    <P>FP-C10 (V) </P>
    <P>FP-C10 (F) </P>
    <!-- ... -->
    <P>Z-ONE-D (C) </P>
  </CELL>
  <CELL COL="3">
    <P>GM-5 (A) </P>
    <P>GM-5-R2 (A) </P>
    <P>GM-5-R2 </P>
    <P>GM-50 </P>
  </CELL>
  <CELL COL="4">
    <P>RU-C1 (B) </P>
    <P>RU-C1 (D) </P>
    <P>RU-C1 </P>
    <!-- ... -->
    <P>RC-Z21 </P>
```

```

</CELL>
<CELL COL="7">
  <P>CA-Z1</P>
  <P>CA-Z2</P>
  <P>CA-Z1SJ</P>
  <P>CA-Z1SP</P>
  <P>CA-Z1M</P>
  <P>CA-Z1M2</P>
  <P>CA-Z1HB</P>
  <P>CA-C10</P><?COL.PAGE?>
  <P>CA-C10SP</P>
  <P>CA-C10SJA</P>
  <P>CA-C10M</P>
  <P>CA-C10B</P>
</CELL>
</ROW>

```

## CLG.MDFO

The CLG.MDFO processing instruction is used:

- to mark up the parts of the consolidated text which do not come from the basic act but from another document such as a modifying act or a corrigendum, and
- to save the source of these modifications.

This processing instruction is used to indicate the start of a modification. It is always accompanied by a CLG.MDFC processing instruction which indicates the end of the modification.

Various attributes describe the specialties of each modification:

The value of the attribute ID is the unique identifier of the occurrence of the modification in the record. The format of the value is of the type

OxxxxnnnMddddeeee

where:

- O: is a fixed value (Open);
- xxx: this value corresponds to the production sequence during which the modifications were introduced for the first time;
- nnn: this value corresponds to the version of the production sequence during which the modifications were introduced for the first time;
- M: is a fixed value;
- dddd: this value corresponds to the order of the modification in the phase (the record);
- eeee: this value is used to define a serial number when there are several occurrences of the same modification; for example, in the case of a global replacement. When there is only a single occurrence, the value is '0000'. When there are several occurrences, the value of the first occurrence is '0001'.

The values 'xxx' and 'nnn' can be deduced from the value assigned to the attribute PROD.SEQ of the element INFO.CONSLEG for the production sequence during which the modifications were introduced for the first time.

The different values must always be represented by three digits. If necessary, the leading positions must be filled with zeros.

For each value assigned to the attribute ID, there must be one and only one occurrence of the processing instruction CLG.MDFC of which the attribute IDREF contains this value.

The value of the attribute IDREF refers to the occurrence of the CLG.MDFC processing instruction, which indicates the end of the modification in question.

For each value assigned to the attribute IDREF, there must be one and only one occurrence of the processing instruction CLG.MDFC of which the attribute ID contains this value.

The ACTION attribute is used to define the type of the modification. The following values are permitted:

- INSERTED: a straight forward insertion,
- DELETED: a straight forward deletion,
- REPEALED: the document is repealed,
- REPLACED: an insertion following a replacement,
- RENUMBERED: an insertion following a renumbering.

For a straight forward deletion, the deleted text must be saved. In this case the value "DELETED" is used. The old text and (where applicable) its structure are stored in the record. The beginning of the deleted text is marked up with the markup CLG.MDFO and the end of the deleted text is marked up with the markup CLG.MDFC.

In the case of the "REPLACED" and "RENUMBERED" values, however, the old text and (where applicable) its structure cannot be stored in the record.

The attribute LEVEL is used to define the nature of the modification.

The following values are permitted:

- STRUCTURE: a modification relating to a structural element,
- TEXT: a modification relating to a portion of text.

The distinction between the two values is not rigid. In some cases the context is indispensable for determining the value. These values are assigned during preparation.

The COMMAND attribute is used to define the way in which the modification has been described.

The following values are permitted:

- EXPLICIT: the modification is described explicitly in a modifying act or in a corrigendum (default value). Modifications described globally, such as global replacements, are also considered as explicit modifications.
- IMPLICIT: the modification is not described explicitly in a modifying act or in a corrigendum. Generally, these are modifications which are implied by an explicit modification; for example, grammatical modifications.
- GENERAL: the modification is described in general documents. Unlike documents containing implicit and explicit modifications, these documents do not necessarily appear on the list of modifying acts.

The ACTIVE.DOC attribute is used to specify the Celex reference of the active document. This is the document that describes the modification to be carried out.

The ACTIVE.LOC attribute is used to identify in more detail the location of the modification command within the active document.

The description of the location may consist of the following elements:

- specific abbreviations;
- numerical values representing the occurrence;
- the colon symbol: this symbol separates the abbreviation from the numerical value;
- the semicolon symbol: if the location consists of several abbreviations indicating several levels in the location, each abbreviation is separated by a semicolon;
- the symbol '.AND.': if there are several references to different locations, the references are separated by a logical connector in the form of '.AND.'.

Spaces are not allowed, except for those placed before and after a logical connector such as '.AND.'.

For abbreviations, the following codes may be used:

- AD: address
- AL: paragraph
- AP: appendix
- AX: annex
- AR: article
- CE: cell (of a table)
- CH: chapter
- CL: column

- CO: considering (preamble)
- DA: dash (EM-DASH)
- DC: declaration
- DI: disposition
- EL: element of list
- FA: final act
- FG: figure
- FI: final part
- FN: footnote
- LI: list
- LT: letter (in the context of an agreement in the form of an exchange of letters)
- PA: paragraph
- PB: preamble
- PC: protocol
- PP: physical paragraph
- PR: part
- PT: item
- RO: row (of a table)
- SC: subchapter
- SD: subdivision
- SE: section
- SN: phrase
- ST: subtitle
- SU: contents
- TA: table
- TI: title
- TT: title as subdivision
- TO: all of the act
- VI: approved

In some cases, additional information may be provided by the following abbreviations:

- (AB): at the beginning
- (AF): after
- (BE): before
- (EN): at the end
- (GL): global
- (IL): insertion in sorted list
- (NN): new numbering
- (ON): old numbering

In the context of paragraphs and articles, special abbreviations may be used:

- .A: 2nd
- .B: 3rd
- .C: 4th
- .D: 5th
- .E: 6th
- .F: 7th
- .G: 8th
- .H: 9th
- .I: 10th
- etc.

If there are more than 26 occurrences, the following abbreviations are used: .AA, .AB, .AC etc.

Some examples:

The value 'AR:1.A;PT:3;PT:c .AND. AX:1' refers to the locations 'Article 1A, item 3c' and 'Annex 1'.

If there is only one annex, the location is indicated as 'AX:0'.

The attribute MOD.LEVEL is used to identify the level of the modification. The default value is '1'. When there is a second-generation modification, which means a modification of the modifying act, the value assigned to MOD.LEVEL is '2', and so on.

Modifications of a higher level must always be imbedded in a lower-level modification. For example, a modification of level '2' must always be integrated into a modification of level '1'.

## **CLG.MDFC**

This processing instruction CLG.MDFC is used to indicate the end of a modification. It is always accompanied by a CLG.MDFO processing instruction which indicates the start of the modification.

The value of the attribute ID is the unique identifier of the occurrence of the modification in the record.

The format of the value is of the type

CxxxxnnnMdddeee

where:

- C: is a fixed value (Close);
- xxx: this value corresponds to the production sequence;
- nnn: this value corresponds to the version;
- M: is a fixed value;
- dddd: this value corresponds to the order of the modification in the phase;
- eeee: this value is used to define a serial number when the same modification is concerned.

The values 'xxx' and 'nnn' must correspond to the values assigned to the attribute PROD.SEQ of the element INFO.CONSLEG.

The different values must always be represented by three digits. If necessary, the leading positions must be filled with zeros.

For each value assigned to the attribute ID, there must be one and only one occurrence of the processing instruction CLG.MDFO of which the attribute IDREF contains this value.

The value of the attribute IDREF refers to the occurrence of the processing instruction CLG.MDFC, which indicates the end of the modification in question.

For each value assigned to the attribute IDREF, there must be one and only one occurrence of the processing instruction CLG.MDFO of which the attribute ID contains this value.

---

## **Composition of filenames**

### ***General rules related to OJ instances***

|                |     |                                                                                                                                                                                              |
|----------------|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [ 1] filename  | ::= | file.OJ   file.ECR   file.Schema   file.PDF  <br>file.CL   file.Sec.Leg                                                                                                                      |
| [ 2] file.OJ   | ::= | series subseries year ojno language '.'<br>(volume page sequence '.')? (container '.')?<br>type                                                                                              |
| [ 3] series    | ::= | 'A'   'C'   'L'   'P'   'S'   'X'                                                                                                                                                            |
| [ 4] subseries | ::= | '_'   'A'   'E'   'I'   'M'                                                                                                                                                                  |
| [ 5] year      | ::= | d{4}                                                                                                                                                                                         |
| [ 6] ojno      | ::= | d{3}                                                                                                                                                                                         |
| [ 7] language  | ::= | 'BG'   'CS'   'DA'   'DE'   'EL'   'EN'   'ES'<br>  'ET'   'FI'   'FR'   'GA'   'HR'   'HU'  <br>'IS'   'IT'   'LT'   'LV'   'MT'   'NL'   'PL'<br>  'PT'   'RO'   'SK'   'SL'   'SV'   'TR' |
| [ 8] volume    | ::= | d{2}                                                                                                                                                                                         |
| [ 9] page      | ::= | d{4}   ('C'   'E'   'S') d{3}                                                                                                                                                                |

|                   |     |                                                                                    |
|-------------------|-----|------------------------------------------------------------------------------------|
| [10] sequence     | ::= | d{2}                                                                               |
| [11] container    | ::= | 'doc'   'frag'   'toc'                                                             |
| [12] type         | ::= | 'frg'   'xml'   'tif'   'tgz'   'zip'                                              |
| [13] file.Sec.Leg | ::= | 'DD' year '_' chapter tome language '.' (page sequence '.')? (container '.')? type |
| [14] chapter      | ::= | d{2}                                                                               |
| [15] tome         | ::= | d{3}                                                                               |
| [16] d            | ::= | '0'   '1'   '2'   '3'   '4'   '5'   '6'   '7'<br>  '8'   '9'                       |

## Vocabulary

|              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|--------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| chapter      | chapter within the context of secondary legislation publications                                                                                                                                                                                                                                                                                                                                                                                                                            |
| container    | indicator for instance describing the composition of a document or a publication:<br><br>no container: XML instance, document component<br><br>doc: description of a document<br><br>frag: encapsulation of fragments<br><br>toc: table of contents of a volume (not a table of contents within a document!)                                                                                                                                                                                |
| d            | symbolic placeholder for digits                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| file.OJ      | file within a OJ publication                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| file.ECR     | file within European Court Reports                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| file.Schema  | file containing a Formex schema version                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| file.PDF     | file within a publication in PDF                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| file.CL      | file containing a consolidated act                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| file.Sec.Leg | table of contents or specific file within a secondary legislation publication                                                                                                                                                                                                                                                                                                                                                                                                               |
| language     | language of the document or the publication                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| ojno         | number of the <i>Official Journal</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| page         | starting page of the document; in some cases the number is preceded by a letter, which indicates that the numbering does not follow the normal pagination; the letter 'S' indicates supplement pages which are used for the cover pages of a publication of for inserted forms in the case of competitions, 'E' indicates any other inserted object, 'C' precedes the number of cover pages of a consolidated document; in any case the length of the pagination never exceeds 4 characters |
| sequence     | sequence of document on the starting page                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| series       | series of the <i>Official Journal</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                       |



|           |                                                                                                                                                                                                                                                                                                                                                                                                                |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| subseries | indication of an <i>Official Journal</i> annexed to another one (A), only existing in an electronic format (E), being an isolated publication (I) or a special edition in Maltese (M)                                                                                                                                                                                                                          |
| tome      | tome within a chapter of the secondary legislation                                                                                                                                                                                                                                                                                                                                                             |
| type      | indication of the contents format:<br><br>frg: The file contains fragments of a bigger document; the content is supposed to be coded in XML (Formex), but it must not necessarily be valid nor well-formed.<br><br>xml: The file contains a valid XML (Formex) instance. This is why a container of type 'frag' should always have this file extension.<br><br>tif: the file contains an image in TIFF format. |
| volume    | indication of the volume of a publication, in most cases this value is 01                                                                                                                                                                                                                                                                                                                                      |
| year      | year of the publication of the <i>Official Journal</i>                                                                                                                                                                                                                                                                                                                                                         |

### Examples:

|                              |                                                                                                                                                                                                           |
|------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L_2002309EN.01000101.xml     | English version of instance of the document starting on page 1 of volume 1, sequence position 1, in <i>Official Journal</i> series L, no subseries, number 309 of the year 2002                           |
| L_2002309EN.01000101.doc.xml | English version of the description of the composition of the document starting on page 1 of volume 1, sequence position 1, in <i>Official Journal</i> series L, no subseries, number 309 of the year 2002 |
| L_2002309EN.toc.xml          | description of the publication of an <i>Official Journal</i> , English version, number 309 of 2002, series L                                                                                              |
| L_2002309EN.xml              | table of contents accompanying the PDF package (SUMMARY.PDF instance)                                                                                                                                     |
| L_2002309EN.xml.tgz          | compressed archive of the delivery of the Formex version of an OJ                                                                                                                                         |

### General rules for ECR instances

These definitions may only be used in the context of the production of European Court Reports.

|               |     |                                                                                                                                                                                     |
|---------------|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [17] file.ECR | ::= | "ECR" court year lg part subpart? "." fascicle ((start.page page.sequence)? ("case"   "notice"))? ("xml"   "tif")                                                                   |
| [18] court    | ::= | "CJ"   "T1"   "F1"   "F2"                                                                                                                                                           |
| [19] year     | ::= | d{4}                                                                                                                                                                                |
| [20] lg       | ::= | "BG"   "CS"   "DA"   "DE"   "EL"   "EN"   "ES"   "ET"   "FI"   "FR"   "GA"   "HR"   "HU"   "IS"   "IT"   "LT"   "LV"   "MT"   "NL"   "PL"   "PT"   "RO"   "SK"   "SL"   "SV"   "XM" |
| [21] part     | ::= | "A"   "B"   "C"                                                                                                                                                                     |

|      |                                |     |                   |
|------|--------------------------------|-----|-------------------|
| [22] | subpart (only for CST reports) | ::= | "1"   "2"         |
| [23] | fascicle                       | ::= | d{2}              |
| [24] | start.page                     | ::= | ([aprs]d{5}) d{6} |
| [25] | page.sequence                  | ::= | d{2}              |

## Examples

|                                  |                           |
|----------------------------------|---------------------------|
| ECRCJ2005FRA.01.xml              | description of a fascicle |
| ECRCJ2005FRA.0100043901.case.xml | description of a case     |
| ECRCJ2005FRA.0100044101.xml      | contents                  |

## The schema file

The file containing the XML schema has a special filename. It is composed as follows:

|      |             |     |                                                           |
|------|-------------|-----|-----------------------------------------------------------|
| [27] | file.Schema | ::= | 'formex-' release '-' date '.xd'                          |
| [28] | release     | ::= | d{2} "." d{2} "cj"                                        |
| [29] | date        | ::= | d{8}                                                      |
| [30] | d           | ::= | '0'   '1'   '2'   '3'   '4'   '5'   '6'   '7'   '8'   '9' |

The schema file which currently is in force can be found under the following address:

— <http://formex.publications.europa.eu/>

## Example

|                          |                                                            |
|--------------------------|------------------------------------------------------------|
| formex-01.00-20040501.xd | Schema file release 01.00, entering into force May 1, 2004 |
|--------------------------|------------------------------------------------------------|

## The PDF files

All the PDF filenames should follow the definition below:

|      |           |     |                                                                          |
|------|-----------|-----|--------------------------------------------------------------------------|
| [31] | file.PDF  | ::= | series subseries OJno date language (startpage endpage)? quality? '.pdf' |
| [32] | date      | ::= | year month day                                                           |
| [33] | month     | ::= | d{2}                                                                     |
| [34] | day       | ::= | d{2}                                                                     |
| [35] | startpage | ::= | (d{4})   ([cser]d{3})                                                    |
| [36] | endpage   | ::= | (d{4})   ([cser]d{3})                                                    |
| [36] | quality   | ::= | '.x'                                                                     |

PDF filenames must be in lower-case.

The table of contents should have as start and end page the number '0000'. The archive should contain also the table of contents of the delivered Official Journal in XML format (should be named following the example below).

## Vocabulary

|           |                                                                                                         |
|-----------|---------------------------------------------------------------------------------------------------------|
| date      | The date of the publication of the OJ.                                                                  |
| endpage   | Number of the last page in the document.                                                                |
| startpage | Number of the first page in the document.                                                               |
| quality   | Indication of the PDF quality: 'x' refers to PDF/X, absence of quality (default value) refers to PDF/A. |

## Examples

|                             |                                  |
|-----------------------------|----------------------------------|
| 1_30920020503en00000000.pdf | PDF document (table of contents) |
| 1_30920020503en00010013.pdf | PDF document (pages 1-13)        |
| 1_30920020503en00140029.pdf | PDF document (pages 14-29)       |
| 1_30920020503en.xml         | XML document (Table of Contents) |

## General rules for Consolidation

Formex files delivered in the context of consolidation (contents: clfilename, archive: clarchive) have to be named as follows:

|                   |     |                                                                                                                                                                                              |
|-------------------|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [39] file.CL      | ::= | series year celexcode lg prodseq "." incseq<br>"." (container ".")? type                                                                                                                     |
| [40] clarchive    | ::= | series year celexcode lg prodseq "."<br>contenttype "." intermediate? archivetype                                                                                                            |
| [41] series       | ::= | "CL"                                                                                                                                                                                         |
| [42] year         | ::= | d{4}                                                                                                                                                                                         |
| [43] celexcode    | ::= | c{4}                                                                                                                                                                                         |
| [44] lg           | ::= | "BG"   "CS"   "DA"   "DE"   "EL"   "EN"   "ES"<br>  "ET"   "FI"   "FR"   "GA"   "HR"   "HU"  <br>'IS'   "IT"   "LT"   "LV"   "MT"   "NL"   "PL"<br>  "PT"   "RO"   "SK"   "SL"   "SV"   "XM" |
| [45] prodseq      | ::= | d{7}                                                                                                                                                                                         |
| [46] incseq       | ::= | d{4}                                                                                                                                                                                         |
| [47] container    | ::= | "doc"   "sls"                                                                                                                                                                                |
| [48] type         | ::= | "xml"   "tif"   "pdf"                                                                                                                                                                        |
| [49] contenttype  | ::= | "xml"   "pdf"                                                                                                                                                                                |
| [50] intermediate | ::= | "int."                                                                                                                                                                                       |
| [51] archivetype  | ::= | "zip"   "tgz"                                                                                                                                                                                |
| [49] c            | ::= | 'A'   'B'   'C'   'D'   'E'   'F'   'G'   'H'  <br>'I'   'J'   'K'  <br><br>'L'   'M'   'N'   'O'   'P'   'Q'   'R'   'S'<br>  'T'   'U'   'V'  <br><br>'W'   'X'   'Y'   'Z'                |
| [50] d            | ::= | '0'   '1'   '2'   '3'   '4'   '5'   '6'   '7'<br>  '8'   '9'                                                                                                                                 |

## Vocabulary

|             |                            |
|-------------|----------------------------|
| archivetype | type of compression method |
|-------------|----------------------------|

|              |                                                                                             |
|--------------|---------------------------------------------------------------------------------------------|
| c            | symbolic placeholder for a character                                                        |
| celexcode    | CELEX code (5 positions, capital letter followed by 4 digits)                               |
| container    | flag for contents of instance: document description or consolidation description            |
| contenttype  | type of format in archive, XML or PDF                                                       |
| d            | symbolic placeholder for digits                                                             |
| incseq       | incremental sequence number starting with 0001, incremented by 1                            |
| intermediate | flag for intermediate, not definitive delivery                                              |
| lg           | symbolic placeholder for the language of the document or archive                            |
| prodseq      | production sequence number as defined by the attribute PROD.SEQ of the INFO.CONSLEG element |
| series       | fixed value: CL                                                                             |
| type         | file format                                                                                 |
| year         | year of the basic document of the consolidated family                                       |

## ***Examples***

|                                   |                                                      |
|-----------------------------------|------------------------------------------------------|
| CL1953D0030FR0040010.xml.zip      | archive with XML content                             |
| CL1953D0030FR0040010.xml.int.zip  | intermediate delivery of an archive with XML content |
| CL1953D0030FR0040010.0001.doc.xml | document description                                 |
| CL1953D0030FR0040010.0001.xml     | consolidated XML document                            |
| CL1953D0030FR0040010.0001.sls.xml | consolidation history description                    |

## ***General rules for secondary legislation***

The naming of the documents which are translations of documents published in the OJ will follow the existing rules. For the documents that are specific to the special edition, the following convention should be adopted:

|                   |     |                                                                                                                                                                                                        |
|-------------------|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [51] file.Sec.Leg | ::= | 'DD' year '_' chapter.se volume.se language<br>'.' (volume page sequence '.')? (container<br>'.')? type                                                                                                |
| [52] year         | ::= | d{4}                                                                                                                                                                                                   |
| [53] chapter.se   | ::= | d{2}                                                                                                                                                                                                   |
| [54] volume.se    | ::= | d{3}                                                                                                                                                                                                   |
| [55] language     | ::= | 'BG'   'CS'   'DA'   'DE'   'EL'   'EN'   'ES'<br>  'ET'   'FI'   'FR'   'GA'   'HR'   'HU'  <br>'IS'   'IT'   'LT'   'LV'   'MK'   'MT'   'NL'<br>  'PL'   'PT'   'RO'   'SK'   'SL'   'SV'  <br>'TR' |
| [56] page         | ::= | d{4}   ('C' 'E' 'S') d{3}                                                                                                                                                                              |

|                |                                                                  |
|----------------|------------------------------------------------------------------|
| [57] sequence  | ::= d{2}                                                         |
| [58] container | ::= 'doc'   'frag'   'toc'                                       |
| [59] type      | ::= 'frg'   'xml'   'tif'   'tgz'   'zip'                        |
| [60] d         | ::= '0'   '1'   '2'   '3'   '4'   '5'   '6'   '7'<br>  '8'   '9' |

## ***Vocabulary***

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| chapter.se | chapter number of the special edition                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| container  | <p>indicator for instance describing the composition of a document or a publication:</p> <p>no container: XML instance, document component</p> <p>doc: description of a document</p> <p>frag: encapsulation of fragments</p> <p>toc: table of contents of a volume (not a table of contents within a document!)</p>                                                                                                                                                                                   |
| d          | symbolic placeholder for digits                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| language   | language of the document or the publication (a list of valid values is available within the schema)                                                                                                                                                                                                                                                                                                                                                                                                   |
| page       | starting page of the document; in some cases the number is preceded by a letter, which indicates that the numbering does not follow the normal pagination; the letter 'S' indicates supplement pages which are used for the cover pages of a publication or for inserted forms in the case of competitions, 'E' indicates any other inserted object, 'C' precedes the number of the cover pages of a consolidated document; in any case the length of the pagination never exceeds 4 characters       |
| sequence   | sequence of document on the starting page                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| type       | <p>indication of the contents format:</p> <p>frg: The file contains fragments of a bigger document; the content is supposed to be coded in XML (Formex), but it must not necessarily be valid nor well-formed</p> <p>xml: The file contains a valid XML (Formex) instance. This is why a container of type 'frag' should always have this file extension</p> <p>xsl: The file contains a layout description in form of a stylesheet in XSL</p> <p>tif: the file contains an image in TIFF format.</p> |
| volume.se  | volume number of the special edition                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| year       | year of the publication of the Official Journal                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

## ***Examples:***

|                               |                                                                                                                                                                                      |
|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DD2004_02001CS.000101.xml     | Czech version of instance of the first document starting on page 1, in Special Edition chapter 2 volume 1 of the year 2004                                                           |
| DD2004_02001CS.000101.doc.xml | Czech version of the description of the composition of the first document starting on page 1, in Special Edition chapter 2 volume 1 of the year 2004                                 |
| L_1975100CS.01000101.xml      | Czech version of instance of the document that corresponds to the translation of one published in OJL, year 1975, number 100, volume 1, page 1, sequence 1                           |
| L_1975100CS.01000101.doc.xml  | Czech version of the description of the composition of the document that corresponds to the translation of one published in OJL, year 1975, number 100, volume 1, page 1, sequence 1 |
| DD2004_02001CS.toc.xml        | description of the publication of a Special Edition, Czech version, chapter 2 volume 1 of the year 2004                                                                              |
| DD2004_02001CS.xml            | table of contents accompanying the PDF package (SUMMARY.PDF instance)                                                                                                                |
| DD2004_02001CS.xml.tgz        | compressed archive of the delivery of the Formex version of a special edition                                                                                                        |
| DD2004_02001CS.xml            | instance accompanying the PDF delivery for the Special Edition, Czech version, chapter 2 volume 1 of the year 2004                                                                   |

## Fragmentation

### 1. Explanations

Large documents may be divided into fragments which are referenced from within the main document or sub-document instance (also see the element "INCL.ELEMENT"). A fragment can contain any portion of the document which must perfectly suit into the superior instance when the content of the fragment just replaces the reference. A merge process will validate the correct use.

As a fragment can contain different elements on the entry level or root level, it is neither always a valid nor a well-formed document. In order to maintain the possibility to validate a fragment (i.e. a file with the extension ".frg"), it has to be accompanied by a file with the same name, but with the suffix ".frag.xml" instead. This instance supplies a single root element (always FRAGMENT) which only contains an entity reference to the contents of the fragment (= file ".frg"). The entity with the unique name "frag" is defined in the subset of the document type statement in the ".frag.xml" instance.

**N.B.:** The delivery of the ".frag.xml" instance is mandatory for each fragment (".frg").

### 2. Example

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE FRAGMENT [
  <!ENTITY frag SYSTEM "CA2002104EN.01000601.frg">
]>
<FRAGMENT xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="Formex4-01.01.20040501.xd">
  &frag;
</FRAGMENT>
```

## Exchange of data (1)

## **Introduction**

This document will describe the exchange of Formex data between the producer of the data and the *Office for Official Publications of the European Communities*.

## **Exchange mode**

The exchange of data between the contractor and the Publications Office (or its representative) is normally executed by using the network. The contractor has to take into account all costs which he considers necessary, while the Publications Office offers the entry point. From a technical point of view, the Publications Office uses a native TCP/IP network which is connected to the public networks by means of ISDN. The transfer protocol for files is FTP (*File Transfer Protocol*), the transfer mode is binary.

If the contractor chooses the network for the transfer of files, a confirmation message can be sent by the Publications Office. These messages will normally be transmitted by e-mail, but other ways are possible as well.

The initiator of the exchange has to ask the Publications Office, Control and Security Section (INFRA/CS), for an adequate access authorisation. He also has to inform about the network address which will be used as well as about the users who will be implied in these procedures (address, telephone number, telefax, e-mail, etc.).

The *Publications Office* uses a firewall system to protect its installations. If the contractor considers it necessary and useful, he may install a similar protection procedure on his own costs. The Publications Office will not be responsible for any attack via IT installations that might be registered.

If the request for access authorisation sent by the contractor is accepted, he will receive, by confidential mail, a user account, a password and the name of the directory where the files have to be transferred to.

## **Transmission**

The contractor transfers the data concerning the language version of each *Official Journal* issue in a single transmission. The name of the transmission is constructed as follows:

SAYYYYNNNLL.OJS.ext

|      |     |                                                                                                                                                            |
|------|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| S    | ::= | series of the <i>Official Journal</i> : C or L                                                                                                             |
| A    | ::= | extension of the series of the <i>Official Journal</i> :<br>A for annexed issue, E for electronic issue (OJ<br>CE only), _ (underscore) for normal issues. |
| YYYY | ::= | the year of the publication                                                                                                                                |
| NNN  | ::= | current number of the <i>Official Journal</i> , always<br>three digits                                                                                     |
| LL   | ::= | language version: one of the following codes has<br>to be used: CS DA DE EL EN ES ET FI FR GA HU IT<br>LT LV MT NL PL PT SK SL SV                          |
| OJS  | ::= | composition of the string OJ and the contents of<br>the S field                                                                                            |
| ext  | ::= | extension according to mode used for the<br>creation: <i>zip</i> or <i>tar.gz</i> .                                                                        |

## **Structure of transmission**

The transmission consists of only one complete language version of an *Official Journal* issue. The archive has to respect the following tree structure:

./SAYYYYNNN/LL/

For the single components see above.

## ***Format for exchange on CDROM***

The standard for the exchange on CDROM is based on ISO 9660, level 2, with a maximum length of filenames up to 31 characters.

The production of the CDROM has to respect the following extensions to the standard ISO 9660 in dependence on the platform for the generation:

1. the option 'Rock Ridge' for UNIX based systems (long file names, conservation of permissions, ...)
  2. the option 'Joliet' for Windows™ based systems (long file names, depth of directories, ...).
- 
- (1) The specifications for the exchange of data may be modified according to the production environment of the *Publications Office*. In cooperation with its partners the Publications Office may adapt the procedure to technological evolution.

## **The image formats**

### ***Preliminary remarks***

This part of the physical specifications clarifies the use of TIFF as format for images which are included in Formex instances (INCL.ELEMENT).

The specifications are based on the international standard TIFF (*Tagged Image File Format*), version 6.0, described in the document

### ***The TIFF standard***

The TIFF 6.0 specifications are composed of two parts. The first one describes the kernel of the standard. It has to be fulfilled by any tool which is said to be TIFF compatible. The second part contains extensions which have been added for special purposes. Therefore they may not be available in all tools.

The kernel describes four image types:

- bilevel encoding, black and white: the two formats TIF\_G3 and TIF\_G4 only allow the encoding of black and white images;
- encoding of different grey scales (16 or 256 levels);
- encoding of colours with a palette included in the TIFF file itself; in that case the pixels reference the entries of the palette;
- encoding of colours according to the RGB methodology (*Red - Green - Blue*).

All tools which are compliant to TIFF 6.0 have to handle these four formats.

The specifications of compression algorithms within the kernel of the standard are limited. For images in black and white, two methods are allowed: TIF\_G3 and PackBits. The other formats only support the method PackBits.

The format TIF\_G4 was introduced as recommendation by CCITT for the encoding of fax transmissions. Its only interest consists in the higher rates of compression. As this format is not part of the kernel, it is not supported by all tools on the market.

### ***Images within Formex deliveries***

Within the Formex deliveries the Publications Office only accepts TIFF images which can be displayed by Quick View and Imaging and be inserted into Word 97 documents.

For these reasons, only the following formats will be accepted by the Publications Office:

- black and white images with TIF\_G3 or PackBits compression; the photometric interpretation has to be set to 'min-is-white' (in general, this is the default value of this tag);
- images with different grey scales, without any compression;
- images with RGB colour encoding, without any compression.



The resolution has to be set to 400 dpi.

For compatibility and legal reasons, all extensions of the standard are forbidden.

## ***Control of the image formats***

The TIFF images have to be controlled by means of the library LIB TIFF V3.4 which is available from the Internet site

<http://www.libtiff.org/>.

The images have to be recognized by the tool TIFFINFO. They have to follow one of the above-mentioned formats as specified by the kernel.

---

## **DOI (Digital Object Identifier)**

### ***Introduction***

- a DOI for each number within the OJ series and for each linguistic version, the DOI suffix is based either on logical references to the publication or on the catalogue number; it is not excluded to have two identifiers according to the dissemination context;
- a DOI for an identified document the title of which appears in the table of contents for each linguistic version.

The DOI suffix has to be based on information which is available at the moment of publication and which is not dependent from any kind of deep analysis.

### ***Formalised DOI-OJ Grammar***

```
doi_doc          ::=      prefix "/" suffix_doc
doi_oj           ::=      prefix "/" suffix_oj
a                ::=      "A" | "B" | "C" | "D" | "E" | "F" | "G" | "H" |
                          "I" | "J" | "K" |
                          "L" | "M" | "N" | "O" | "P" | "Q" | "R" | "S" |
                          "T" | "U" | "V" |
                          "W" | "X" | "Y" | "Z"
d                ::=      "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" |
                          "8" | "9"
issn             ::=      d{4} "-" d{3} (d | "X")
lang.ver         ::=      «ISO 639-2»
log.id           ::=      series sub-series year "." num.cur
log.id.ext       ::=      log.id "." tome page sequence
num.cur          ::=      d{3}
page             ::=      (d | "s" | "e" | "c") d{3}
prefix           ::=      "10.3000" | "10.3001"
sequence         ::=      d{2}
series           ::=      "C" | "L"
sub-series       ::=      "_" | "A" | "E" | "G" | "M"
suffix_doc       ::=      log.id.ext "." lang.ver
suffix_oj        ::=      issn "." (log.id "." lang.ver)
tome             ::=      d{2}
year            ::=      d{4}
```

## ***Vocabulary***

|            |                                                         |
|------------|---------------------------------------------------------|
| doi_doc    | document identifier                                     |
| doi_oj     | identifier for OJ (single number)                       |
| a          | collection of capital letters allowed in the identifier |
| d          | digit                                                   |
| issn       | formatted ISSN                                          |
| lang.ver   | language code according to ISO 639-2                    |
| log.id     | basic components of logical identifier                  |
| log.id.ext | extended logical identifier                             |
| num.cur    | current number of OJ                                    |
| page       | page number (if necessary with prefix)                  |
| prefix     | DOI prefix                                              |
| sequence   | sequence of document on page                            |
| series     | main series of OJ                                       |
| sub-series | sub-series                                              |
| suffix_doc | DOI suffix for a document                               |
| suffix_oj  | DOI suffix for a number of OJ                           |
| tome       | number of tome                                          |
| year       | publication year                                        |

### ***Examples:***

|                                          |                                       |
|------------------------------------------|---------------------------------------|
| doi:10.nnnn/1725-<br>2563.L_2008.116.fre | OJ L series in FR (logical reference) |
| doi:10.nnnn/L_2008.116.01000301.fre      | Second document in the OJ             |

## **Catalog of schema versions**

### ***1. Explanations***

An XML instance in a file with the unique name "catalog.xml". It only contains a list of the different versions which have been used for the production of the electronic deliveries of Official Journal publications. The root element "catalog" contains one or more elements named "entry". Each element is supplied with up to four attributes.

The attribute "file" contains the filename of the catalog. It is mandatory.

The attribute "version" indicates the version of the Formex schema. It is also mandatory.

The attribute "date" specifies the date when the version entered into force. It is mandatory as well.

The optional attribute "comment" may contain information on the use of the concerned version.

## **2. Example**

```
<?xml version="1.0" encoding="UTF-8"?>
<catalog xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="catalog.xd">
  <entry file="formex4-0.99-20040301.xd" version="0.99" date="20040301" comment="for
test reasons only"/>
</catalog>
```

## **Physical specifications - Differences**

### ***Preliminary remark***

The compiled differences only concern the physical specifications.

### ***Release 0.98***

- correction of filename conventions

### ***Release 0.99***

- name for schema file (filenames) [not yet applied for this release];  
news section;  
catalog of schema versions;  
manual for use of fragmentation;  
precision on rules for the use of elements DATE, TITLE, PRINTER.

### ***Release 1.00***

- clarifications on processing instructions
- format for data exchange on CDROM

### ***Release 1.01***

- suppression of presentation instructions

### ***Release 1.02 and 2.00***

- integration of several characters in order to become compatible with EUAlbertina  
redefinition of character ranges  
suppression of data exchange via magnetic tapes  
text entities allowed in the context of fragment handling  
filename convention for special editions in Maltese

### ***Release 3.00***

- correction of some errors in the documentation; clarification of ellipsis
- numbering of notes continued over several documents: NNC attribute added to "root" elements
- IN.SUMMARY attribute added to ITEM.REF
- specific element (SE) for secondary legislation
- values BG/RO added to CRIT language attribute
- optional attribute ITEM.ID added to ITEM element (only for internal use, see Manual)
- TOC, ANNOTATION, and GR.ANNOTATION added to CONS.DOC and CONS.DOC.GEN
- suppression of element INCLUSIONS from BIB.INSTANCE.CONC
- attribute values ROMAN/roman added to NOTE numbering scheme
- elements NOTE, QUOT.END, QUOT.START added to t\_btx.form
- modification of IDENTIFIER attribute for ARTICLE and PARAG elements
- clarification of use of HT element on complete paragraphs
- Bulgarian and Romanian quotation marks added to physical specifications
- Update of table of characters for Basque, Catalan and Galician

### ***Release 4.0***

- addition of special elements which are only in the context of the Recueil of the Court of Justice; these elements do not concern the production of the Official Journal;
- revision of the type t\_sectiontype to fit with the new structure of the OJ L and C;
- revision of the element SECTION.SUMMARY.

### ***Release 4.cl***

- addition of filename convention for European Court Reports;
- addition of 70 types and elements for the instances produced as European Court Reports.

### ***Release 5.00***

- addition of new type definitions: t\_doi.doc, t\_doi.oj, t\_doi.pub, t\_prod.id, t\_fin.id;
- addition of new elements: DOI.OJ, DOI.PUB, PROD.ID, FIN.ID;
- modification of existing elements: ACT, AGR, ANNEX, BIB.DOC, BIB.INSTANCE, BIB.VOLUME, CJT, COMPETITION, CORR, DOC.REF.ECR, GENERAL, GR.SEQ, ITEM.REF, LETTER, DOC.MAIN.PUB;
- modification in the ranges of accepted characters.

### ***Release 5.20***

- adaptations due to EU extension: HR, IS.
- adaptations to ECR elements

### ***Release 5.56***

- protection of personal rights
-

introduction of element ANONYMOUS, introduction of attribute REPUBLICATION on toc level

### ***Release 5.57 and 5.58***

- New release of Vademecum of CJ
- modification of ECR models, modification of element INDEX and LIST

### ***Release 5.59***

- new attribute value "0022" added to simpleType "t\_quotation.end" and "t\_quotation.start", for new character of quotation mark used in Dutch

### ***Release 6.00, 6.01, 6.02, 6.02.1, 6.02.2, 6.02.3, 6.02.4***

- Add support of OJ AbA
- new element NO.ELI added as optional in DOCUMENT.REF.CONNS
- added two new unicode characters "0406" and "0456" to the list allowed characters
- new element LINK created to be used for any uri link
- new element REF.DOC.SE created to be used for references to special editions
- creation of type t\_chapter, needed for element REF.DOC.SE
- new pattern of type t\_volume, needed for element REF.DOC.SE
- added new elements LINK and REF.DOC.SE inside type t\_btx
- added new elements LINK and REF.DOC.SE inside type t\_btx.seq
- added new elements LINK and REF.DOC.SE inside type t\_btx.seq
- new character added in allowed character ranges: unicode 25A1 for white square
- Introduction of new type of document, SUMMARY and RESUME
- new element TITLE added as optional in INDEX
- element TITLE in SUMMARY made as optional
- RESUME added as value REF.PDF.ECR element
- DATE.PUB no longer mandatory for REF.DOC.SE element
- DATE.PUB no longer mandatory for REF.DOC.SE element
- Introduction of new element DEC.TITLE added as optional in CASE
- new character added in allowed character ranges: unicode 202F for thin non-breaking space
- NO.CASE element can be used in BIB.INSTANCE element
- New element ANSM that can be used in BIB.DOC, BIB.INSTANCE and BIB.INSTANCE.COS elements.

### ***Nota bene***

Detailed information on the changes in the schema can also be found in the beginning of the schema itself.

---