

Noto Sans CJK TC (NotoSansCJKtc-Medium.otf dated 2021-04-30)

This CJK Sans font is intended for use with Traditional Chinese but also covers other CJK languages, those languages written with Latin-1 alphabets (not the Catalan precomposed L with middle dot, nor the Dutch ij digraphs, nor the uppercase Y with diaeresis added in Unicode), Vietnamese, and Russian. The Greek alphabet does not include accented letters, so is unusable for normal Greek text.

For variations of the Latin and Cyrillic alphabets I attempt to show the whole alphabet, followed by Article 1 of the UDHR (it is traditional!). For other scripts and languages I merely display the digits (if not common 0..9) and Article 1.

Sources:

http://unicode.org/udhr/assemblies/first_article_all.html

<http://www.omniglot.com/>

<https://en.wikipedia.org>

For Chinese and for Japanese Kanji (‘ Han’ ideograms) it is impossible to know exactly which codepoints you are likely to encounter. I have begun by assuming that older fonts are probably adequate for common usage and so I ignore CJK Extension B and later when trying to determine if a font is likely to be adequate.

I am primarily concerned with informal postings online, but I have tried to estimate which items might be found in more formal texts.

This font covers all my test Han codepoints.

CJK languages use an ellipsis codepoint to show omissions or tailing-off in speech. The codepoint is ideally placed at mid-level (most old CJK fonts do that if they do not cover non-English Latin languages), but in Unicode 14 the mid-level ellipsis (a mathematical symbol) was recommended for this purpose. Before that, Japanese fonts tended to ignore that mathematical symbol because it was not in the JIS tables of glyphs. This font places both ellipses on the baseline which is technically wrong.

This document was prepared using Xe_{La}TeX and fontspec. Missing glyphs will be replaced by an X in a box.

This font is available in many forms, either language-specific (TC) or all-CJK languages, in TTCs or individual files. The available weights cover from thin to black (not semi-bold nor extra-bold) and there are variable-weight versions. I am using the static files for Medium and Bold weights.

For the Latin and Cyrillic alphabets it uses normal-width characters.

This font seems to support Korean Hanja which can still be used in personal names and in signs for place names (i.e. it covered all the examples I could find), but some of the glyph shapes will look wrong to a Korean.

This font lacks the OpenType tag ‘cyril’ needed to use polyglossia for Cyrillic scripts.

1 Latin Alphabet (European variants)

1.1 Catalan

A a, B b, C c, Ç ç, D d, E e, É é, È è, F f, G g, H h, I i, J j, K k, L l, (L and l or L • L l • l are a different pronunciation to LL and ll), M m, N n, O o, Ó ó, Ò ò, P p, Q q, R r, S s, T t, U u, Ü ü, V v, W w, X x, Y y, Z z

Tots els éssers humans neixen lliures i iguals en dignitat i en drets. Són dotats de raó i de consciència, i han de comportar-se fraternalment els uns amb els altres.

1.2 Danish

A a, B b, C c, D d, E e, F f, G g, H h, I i, J j, K k, L l, M m, N n, O o, P p, Q q, R r, S s, T t, U u, V v, W w, X x, Y y, Z z, Æ æ, Ø ø, Å å

Alle mennesker er født frie og lige i værdighed og rettigheder. De er udstyret med fornuft og samvittighed, og de bør handle mod hverandre i en broderskabs ånd.

1.3 Dutch

As well as the standard 26 letters, the IJ digraph ( ) may be found.

Alle mensen worden vrij en gelijk in waardigheid en rechten geboren. Zij zijn begiftigd met verstand en geweten, en behoren zich jegens elkander in een geest van broederschap te gedragen.

1.4 French

A a, À à, Â â, Æ æ, Ä ä, B b, C c, Ç ç, D d, E e, É é, È è, Ê ê, Ë ë, F f, G g, I i, Î î, Ï ï, J j, K k, L l, M m, N n, (Ñ ñ may be seen on words of Spanish origin), O o, Ô ô, Œ œ, Ö ö, P p, Q q, R r, S s, T t, U u, Ù ù, Û û, Ü ü, V v, W w, X x, Y y,  y, Z z

Tous les êtres humains naissent libres et égaux en dignité et en droits. Ils sont doués de raison et de conscience et doivent agir les uns envers les autres dans un esprit de fraternité.

1.5 German

A a, Ä ä, B b, C c, D d, E e, F f, G g, H h, I i, J j, K k, L l, M m, N n, O o, Ö ö, P p, Q q, R r, S s, T t, U u, Ü ü, V v, W w, X x, Y y, Z z and ß (but ss in Switzerland)

Alle Menschen sind frei und gleich an Würde und Rechten geboren. Sie sind mit Vernunft und Gewissen begabt und sollen einander im Geist der Brüderlichkeit begegnen.

1.6 Icelandic

A a, Á á, B b, D d, Ð ð, (eth is not used for the first letter of any word) E e, É é, F f, G g, H h, I i, Í í, J j, K k, L l, M m, N n, O o, P p, R r, S s, T t, U u, Ú ú, V v, X x, Y y, Ý ý, Þ þ, Æ æ, Ö ö

Hver maður er borinn frjálss og jafn öðrum að virðingu og réttindum. Menn eru gæddir vitsmunum og samvitzku, og ber þeim að breyta bróðurlega hverjum við annan.

1.7 Italian

A a, À à, B b, C c, D d, E e, É é, È è, F f, G g, H h, I i, Í í, Ì ì, Î î, L l, M m, N n, O o, Ó ó, Ò ò, P p, Q q, R r, S s, T t, U u, Ú ú, Û ù, V v, Z z

Tutti gli esseri umani nascono liberi ed eguali in dignità e diritti. Essi sono dotati di ragione e di coscienza e devono agire gli uni verso gli altri in spirito di fratellanza.

1.8 Portuguese

A a, Á á, Â â, Ã ã, À à, B b, C c, Ç ç, D d, E e, É é, Ê ê, F f, G g, H h, I i, Í í, J j, K k, L l, M m, N n, O o, Ó ó, Ô ô, Õ õ, P p, Q q, R r, S s, T t, U u, Ú ú, V v, W w, X x, Y y, Z z

Todos os seres humanos nascem livres e iguais em dignidade e em direitos. Dotados de razão e de consciência, devem agir uns para com os outros em espírito de fraternidade.

1.9 Spanish

A a, Á á, B b, C c, D d, E e, É é, F f, G g, H h, I i, J j, K k, L l, M m, N n, Ñ ñ, O o, Ó ó, P p, Q q, R r, S s, T t, U u, Ú ú, Ü ü, V v, W w, X x, Y y, Z z

Todos los seres humanos nacen libres e iguales en dignidad y derechos y, dotados como están de razón y conciencia, deben comportarse fraternalmente los unos con los otros.

2 Cyrillic Alphabet

2.1 Russian

А а, Б б, В в, Г г, Д д, Е е, Ё ё, Ж ж, З з, И и, Й й, К к, Л л, М м, Н н, О о, П п, Р р, С с, Т т, У у, Ф ф, Х х, Ц ц, Ч ч, Ш ш, Щ щ, Ъ ъ, Ы ы, Ь ь, Э э. Ю ю, Я я

Все люди рождаются свободными и равными в своем достоинстве и правах. Они наделены разумом и совестью и должны поступать в отношении друг друга в духе братства.

3 CJK scripts

For Chinese and Japanese these examples might use the styles appropriate to the selected variant (TC) of this Noto CJK font

3.1 Cantonese

人人生出嚟就係自由嘅，喺尊嚴同權利上一律平等。佢哋具有理性同良心，而且應該用兄弟間嘅關係嚟互相對待。

3.2 Japanese

すべての人間は、生まれながらにして自由であり、かつ、尊厳と権利とについて平等である。人間は、理性と良心とを授けられており、互いに同胞の精神をもって行動しなければならない。

3.3 Korean

모든 인간은 태어날 때부터 자유로우며 그 존엄과 권리에 있어 동등하다. 인간은 천부적으로 이성과 양심을 부여받았으며 서로 형제애의 정신으로 행동하여야 한다.

3.4 Simplified Chinese

人人生而自由,在尊严和权利上一律平等。他们赋有理性和良心,并应以兄弟关系的精神相对待。

3.5 Traditional Chinese

In some of my files I use ' TW' as a shorthand for Traditional Chinese used in Taiwan. This is the Mandarin version of Article 1:

人人生而自由，在尊嚴和權利上一律平等。他們賦有理性和良心，並應以兄弟關係的精神互相對待。

However, the common dialect in Taiwan, Taiwanese Hokkien (臺語) a.k.a. Minnan, is different and was what the abandoned UKai and UMin fonts were hoping to support. This is its Article 1, and this font fits one less ideogram on the first line for this text when using XeLaTeX.

人皆生而自由；在尊嚴及權利上均各平等。人各賦有理性良知，誠應和睦相處，情同手足。

4 Other variants of the Latin alphabet

4.1 Vietnamese

The interesting features of Vietnamese are the use of horns on ‘o’ and ‘u’ to provide additional vowels, and the use of diacritics to show the tone. The letters themselves are pre-composed, but displaying some of the individual tone markings requires combining accents and this font does not provide those.

A a, Ă ă, Â â, B b, C c, D d, Đ đ, E e, Ê ê, G g, H h, I i, K k, L l, M m, N n, O o, Ô ô, Ó ó, P p, Q q, R r, S s, T t, U u, Ú ú, V v, X x, Y y

The additional markings for tone are as follows.

If two descriptions, they are northern / southern :

A a, Ă ă, Â â, E e, Ê ê, I i, O o, Ô ô, Ơ ơ, U u, Ứ ứ, Y y mid level
 ˘ À à, Ằ ằ, Ẳ ẳ, È è, Ẻ ẻ, Ì ì, Ò ò, Ỗ ỗ, Ỡ ỡ, Ù ù, Ừ ừ, ò ò low falling

☒ Ǻ ǻ, Ǻ ǻ, Ǻ ǻ, Ẹ ẹ, Ẹ ẹ, ỉ ỉ, Ỏ ỏ, Ỏ ỏ, Ỡ ỡ, Ỡ ỡ, Ừ ừ, Ừ ừ, Ỡ ỡ, Ỡ ỡ mid falling / dipping

☒ Āā, ǺǺ, ǺǺ, Ēē, Ēē, Īī, Ōō, Ōō, Ōō, Ūū, Ūū, Ỹỹ glottalized rising / dipping

Á á, Ă ă, Ã ã, É é, Ê ê, Í í, Ó ó, Ô ô, Ǫ ǫ, Ú ú, Ǻ ǻ, Ý ý high rising

☒ A a, Ă ă, Â â, E e, Ê ê, I i, O o, Ô ô, Ơ ơ, U u, Ư ư, Y y glottalized falling / low rising

Tất cả mọi người sinh ra đều được tự do và bình đẳng về nhân phẩm và quyền lợi. Mọi con người đều được tạo hóa ban cho lý trí và lương tâm và cần phải đối xử với nhau trong tình anh em.

5 Quotation-marks, etc.

In real-world usage, particularly when looking at news sites, the presence or absence of the correct quote characters will make a difference. As a coder, I use ‘...’ or ”...” but in good text there are variations and this font supports:

left and right single quotes ‘...’	angle brackets <...>
left and right double quotes “...”	double angle brackets «...»
single turned comma, with right single quote ‘...’	corner brackets 「...」
double low 9 and left double quotes „...“	white corner brackets 『...』
guillemets «...»	
angle quotes <...>	

In Spanish languages, the inverted exclamation-mark ¡ and the inverted question-mark ¿ are necessary.

6 Symbols

6.1 Currency Symbols

Many countries have symbols for their currencies. In alphabetic order:

⌔ Cedi (Ghana)	⌔ Manat (Azerbaijan)
¢ Cent (many places)	⌔ Naira (Nigeria)
⌔ Colon (Costa Rica, El Salvador)	£ Pound Sterling
⌔ Cruzeiro (Brazil)	⌔ Rouble (Russia)
\$ Dollar (many places)	⌔ Tenge (Kazakhstan)
₫ Dong (Vietnam)	⌔ Turkish Lira
€ Euro	₩ Won (Korea)
⌔ Hryvnia (Ukraine)	¥ Yen (Japan)

6.2 Other symbols in text

© Copyright	¶ Pilcrow
† Dagger	® Registered
‡ Double Dagger	§ Section
‰ Per mille	™ Trademark