Source Han Sans Version 2.000

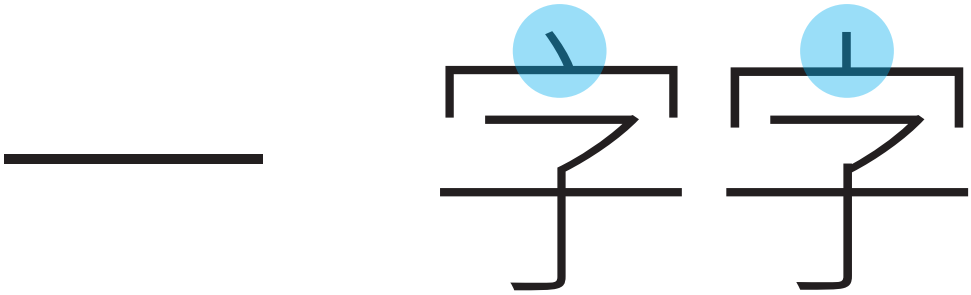
Built on November 2, 2018 • Released on November 19, 2018

Overview

Source Han Sans, designed by Ryoko Nishizuka (西塚涼子), is the companion sans serif–style Pan-CJK typeface family to Source Han Serif, and is offered in seven slightly different weights—ExtraLight, Light, Normal, Regular, Medium, Bold, and Heavy—and in several OpenType/CFF-based deployment configurations to accommodate various system requirements or, in some cases, limitations. Pan-CJK fonts, such as those provided by the *Source Han* typeface families, are intended to support and render the most important characters for Simplified Chinese, Traditional Chinese, Japanese, and Korean.

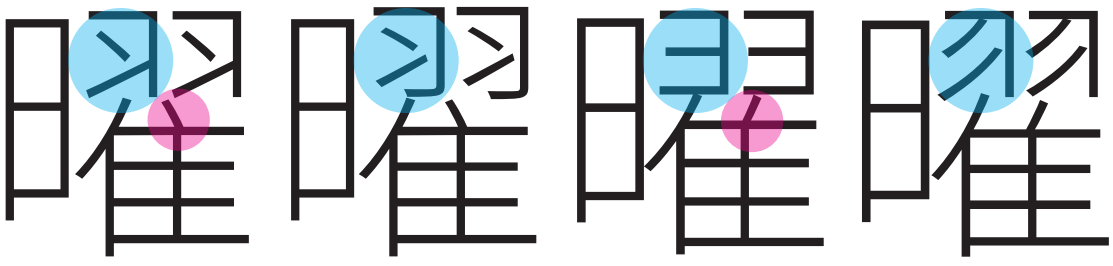
The samples on this page demonstrate that the differences for each language can be subtle or striking, depending on the ideograph, yet they all clearly share the same typeface style, design, weight, and other characteristics that are not necessarily tied to a particular language.

The first sample shows the completely shared form of U+4E00, along with the shared Simplified/Traditional Chinese and shared Japanese/Korean forms of U+5B57:



The second sample below shows, from left to right, the Simplified Chinese form of U+9AA8, its Traditional Chinese (Taiwan) form, and the form shared by Traditional Chinese (Hong Kong), Japanese, and Korean:

The third sample shows, also from left to right, the almost completely unshared Simplified Chinese, Traditional Chinese (Taiwan and Hong Kong), Japanese, and Korean forms of U+66DC:



The fourth and final example shows, again from left to right, the completely unshared Simplified Chinese, Traditional Chinese (Taiwan), Traditional Chinese (Hong Kong), Japanese, and Korean forms of U+8FD4:

The short passage shown below is Genesis 11:1 (创世记 11:1 in Simplified Chinese, 創世記 11:1 in Traditional Chinese, 創世記 11:1 in Japanese, and 창세기 11:1 in Korean) displayed in six languages and in three of the seven weights:

ExtraLight

Now the whole world had one language and a common speech.
那时，天下人的口音、言语都是一样。
那時，天下人的口音、言語都是一樣。
那時，天下人的口音、言語都是一樣。
全地は同じ発音、同じ言葉であった。
온 땅의 구음이 하나이요 언어가 하나이었더라.

Regular

Now the whole world had one language and a common speech.
那时，天下人的口音、言语都是一样。
那時，天下人的口音、言語都是一樣。
那時，天下人的口音、言語都是一樣。
全地は同じ発音、同じ言葉であった。
온 땅의 구음이 하나이요 언어가 하나이었더라.

Heavy

Now the whole world had one language and a common speech.
那时，天下人的口音、言语都是一样。
那時，天下人的口音、言語都是一樣。
那時，天下人的口音、言語都是一樣。
全地は同じ発音、同じ言葉であった。
온 땅의 구음이 하나이요 언어가 하나이었더라.

The pages that follow provide excruciating technical details about the font resources that are included in this open source project, and the information corresponds to Version 2.000.

Configurations

Source Han Sans is provided in four basic deployment configurations, each of which is described below, along with typical usage scenarios:

Language-specific OpenType/CFF (OTF)—45 font resources

This deployment configuration is available in five languages—Simplified Chinese, Traditional Chinese (Taiwan), Traditional Chinese (Hong Kong), Japanese, and Korean—and sets one language as the default (a default language is required due to the single 'cmap' table), and the 'locl' (*Localized Forms*) GSUB feature is expected to be used to access glyphs that are appropriate for the other four supported languages. (Note that the Regular and Bold weights include an additional font for all five language, which differ only in that the default glyphs for ASCII (U+0020 through U+007E), U+00A0 ( ), U+00A5 (¥), U+00AD (-), U+2011 (-), and U+20A9 (₩) are half-width instead of proportional, and the default glyph for U+2423 (␣) is half-width instead of full-width. These fonts include the additional “HW” identifier in their names.)

> *These fonts represent the most compact form that supports all languages and includes the complete set of glyphs, but this comes at the expense of requiring an application to properly support the 'locl' GSUB feature in order to display glyphs for languages other than the default one. In addition to using such an application, a good example of which is Adobe InDesign, the text—at the character, paragraph, or document level—must also be properly language-tagged.*

OpenType/CFF Collection (OTC)—7 font resources

This deployment configuration represents a “best of all possible worlds” in that there are separate font instances for each language, and while each font instance necessarily specifies a default language, the 'locl' GSUB feature can still be used to access the glyphs for the other languages. Like the language-specific OTFs, the Regular and Bold weights additionally include font instances for all five languages whose ASCII and small number of additional characters are half-width instead of the usual proportional.

> *These fonts offer the greater flexibility in that there is a single font resource that includes five or ten font instances, each with a different one of the five languages serving as the default. Users of these fonts simply choose the appropriate font in an application’s font menu, and the glyphs that are suitable for that language are displayed. However, OpenType/CFF Collections are not yet broadly supported. Windows 10 Anniversary Update (Version 1607, released on 2016-08-02), OS X Version 10.8, iOS Version 7.0, and Adobe CS6 applications represent the first environments that support this particular deployment format. Note that if you install the OTCs, you cannot install any of the corresponding language-specific OTFs, because they share the same names.*

Super OpenType/CFF Collection (Super OTC)—a single font resource

This deployment configuration packs all seven weights and all five languages, along with half-width variations of two of the seven weights, into a single font resource that includes a total of 45 font instances and 458,745 total glyphs. As a result of 'sfnt' table sharing, there are seven unique 'CFF ', 'hmtx', and 'vmtx' tables (one per weight), five unique 'GSUB' tables (one per language), and ten unique 'cmap' tables (one per language and proportional/half-width combination). These represent the largest tables, so greater sharing leads to a smaller overall footprint. This saves over 13MB compared to the seven separate OTCs. While each font

instance specifies a default language, the 'locl' GSUB feature can still be used to access the glyphs for the four other supported languages.

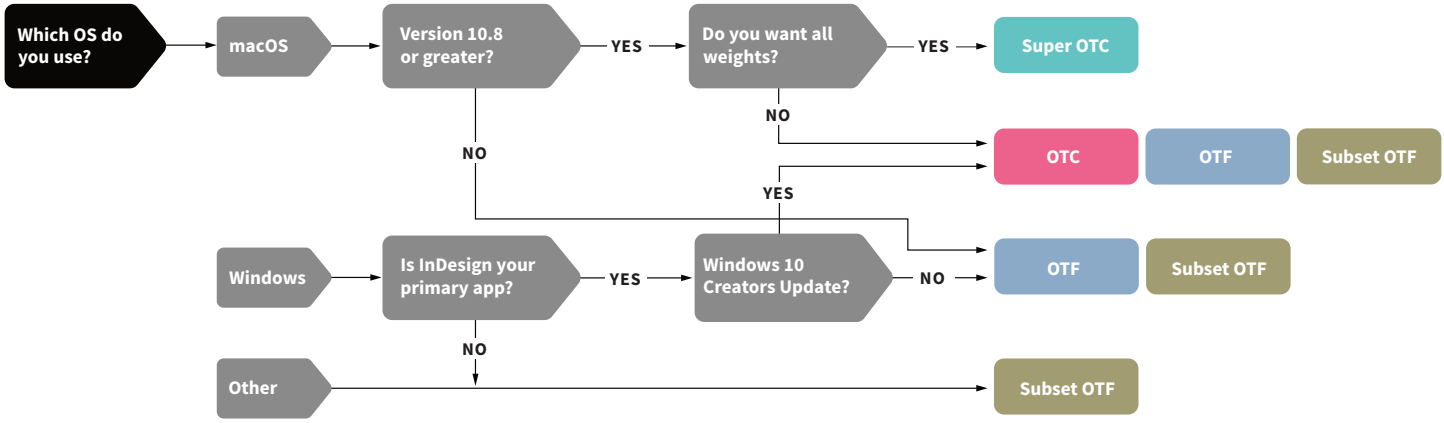
This font offers the greatest flexibility in that there is a single font resource that includes all 45 font instances, one for each of the seven weights and five languages, along with half-width versions of the Regular and Bold weights. Users of this deployment format simply choose the appropriate font in an application's font menu, and the glyphs that are suitable for that language are displayed. The Super OTC is subject to the same caveats and limitations as the weight-specific OTCs. Please be aware that while Windows 10 Anniversary Update (Version 1607) is the first version of Windows OS to support OTCs, Windows 10 Creators Update (Version 1703, released on 2017-04-05) is necessary to support the Super OTC, due to its large number of font instances.

Region-specific Subset OpenType/CFF (Subset OTF)—35 font resources

This deployment configuration includes five different subsets, and each subset includes only the glyphs that are necessary for Simplified Chinese, Traditional Chinese (Taiwan), Traditional Chinese (Hong Kong), Japanese, or Korean.

These fonts are considered the most broadly usable because the 'locl' GSUB feature is not required to access the region-specific glyphs. Instead, only the glyphs that are necessary for each region are included. This deployment configuration is recommended for users who need only the glyphs for a specific region, and also desire the smallest possible footprint. These fonts are expected to behave the same as conventional Simplified Chinese, Traditional Chinese, Japanese, or Korean fonts.

The flowchart below may be helpful in deciding which deployment configurations are usable in your current or preferred working environment:



If the flowchart suggests that you are able to make use of more than one deployment configuration, please re-examine the descriptions above to determine which one would better satisfy your usage needs.

Our Recommendation

If the flowchart indicates that you are able to make use of more than one deployment configuration, we generally recommend one or more of the region-specific subset OTFs, mainly due to the fact that they are usable in a broader range of environments. If your environment supports the 45-font Super OTC, its biggest advantage—other than being a single font resource that it easy to manage—is that it is the most space-efficient deployment configuration that includes all 458,745 glyphs. If you choose to download and install the language-specific OTFs, OTCs, or Super OTC, please be mindful of their requirements for accessing region-specific glyphs and various OS limitations.

Font Resources

The table below lists all 88 font resources that are included in this release, organized by format and language, and providing their file and PostScript names (for the OTCs and Super OTC, the order of the fonts in the fourth column represents the actual order of the fonts in those font resources):

Format	Language	File Name	PostScript Name/Names
OTF	Simplified Chinese	*SourceHanSansSC-ExtraLight.otf*	SourceHanSansSC-ExtraLight
		SourceHanSansSC-Light.otf	SourceHanSansSC-Light
		SourceHanSansSC-Normal.otf	SourceHanSansSC-Normal
		SourceHanSansSC-Regular.otf	SourceHanSansSC-Regular
		SourceHanSansHWSC-Regular.otf	SourceHanSansHWSC-Regular
		SourceHanSansSC-Medium.otf	SourceHanSansSC-Medium
		SourceHanSansSC-Bold.otf	SourceHanSansSC-Bold
		SourceHanSansHWSC-Bold.otf	SourceHanSansHWSC-Bold
		SourceHanSansSC-Heavy.otf	SourceHanSansSC-Heavy
	Traditional Chinese—Taiwan	*SourceHanSansTC-ExtraLight.otf*	SourceHanSansTC-ExtraLight
		SourceHanSansTC-Light.otf	SourceHanSansTC-Light
		SourceHanSansTC-Normal.otf	SourceHanSansTC-Normal
		SourceHanSansTC-Regular.otf	SourceHanSansTC-Regular
		SourceHanSansHWTC-Regular.otf	SourceHanSansHWTC-Regular
		SourceHanSansTC-Medium.otf	SourceHanSansTC-Medium
		SourceHanSansTC-Bold.otf	SourceHanSansTC-Bold
		SourceHanSansHWTC-Bold.otf	SourceHanSansHWTC-Bold
		SourceHanSansTC-Heavy.otf	SourceHanSansTC-Heavy
	Traditional Chinese—Hong Kong	*SourceHanSansHC-ExtraLight.otf*	SourceHanSansHC-ExtraLight
		SourceHanSansHC-Light.otf	SourceHanSansHC-Light
		SourceHanSansHC-Normal.otf	SourceHanSansHC-Normal
		SourceHanSansHC-Regular.otf	SourceHanSansHC-Regular
		SourceHanSansHWHC-Regular.otf	SourceHanSansHWHC-Regular
		SourceHanSansHC-Medium.otf	SourceHanSansHC-Medium
		SourceHanSansHC-Bold.otf	SourceHanSansHC-Bold
		SourceHanSansHWHC-Bold.otf	SourceHanSansHWHC-Bold
		SourceHanSansHC-Heavy.otf	SourceHanSansHC-Heavy
	Japanese	*SourceHanSans-ExtraLight.otf*	SourceHanSans-ExtraLight
		SourceHanSans-Light.otf	SourceHanSans-Light
		SourceHanSans-Normal.otf	SourceHanSans-Normal
		SourceHanSans-Regular.otf	SourceHanSans-Regular
		SourceHanSansHW-Regular.otf	SourceHanSansHW-Regular
		SourceHanSans-Medium.otf	SourceHanSans-Medium
		SourceHanSans-Bold.otf	SourceHanSans-Bold
		SourceHanSansHW-Bold.otf	SourceHanSansHW-Bold
		SourceHanSans-Heavy.otf	SourceHanSans-Heavy
	Korean	*SourceHanSansK-ExtraLight.otf*	SourceHanSansK-ExtraLight
		SourceHanSansK-Light.otf	SourceHanSansK-Light
		SourceHanSansK-Normal.otf	SourceHanSansK-Normal
		SourceHanSansK-Regular.otf	SourceHanSansK-Regular
		SourceHanSansHWK-Regular.otf	SourceHanSansHWK-Regular
		SourceHanSansK-Medium.otf	SourceHanSansK-Medium
		SourceHanSansK-Bold.otf	SourceHanSansK-Bold
		SourceHanSansHWK-Bold.otf	SourceHanSansHWK-Bold
		SourceHanSansK-Heavy.otf	SourceHanSansK-Heavy
OTC	All	*SourceHanSans-ExtraLight.ttc*	SourceHanSans-ExtraLight, SourceHanSansK-ExtraLight, SourceHanSansSC-ExtraLight, SourceHanSansTC-ExtraLight, SourceHanSansHC-ExtraLight

Format	Language	File Name	PostScript Name/Names
OTC (cont'd)	ALL	*SourceHanSans-Light.ttc*	SourceHanSans-Light, SourceHanSansK-Light, SourceHanSansSC-Light, SourceHanSansTC-Light, SourceHanSansHC-Light
		SourceHanSans-Normal.ttc	SourceHanSans-Normal, SourceHanSansK-Normal, SourceHanSansSC-Normal, SourceHanSansTC-Normal, SourceHanSansHC-Normal
		SourceHanSans-Regular.ttc	SourceHanSans-Regular, SourceHanSansK-Regular, SourceHanSansSC-Regular, SourceHanSansTC-Regular, SourceHanSansHC-Regular, SourceHanSansHW-Regular, SourceHanSansHWK-Regular, SourceHanSansHWSC-Regular, SourceHanSansHWTC-Regular, SourceHanSansHWHC-Regular
		SourceHanSans-Medium.ttc	SourceHanSans-Medium, SourceHanSansK-Medium, SourceHanSansSC-Medium, SourceHanSansTC-Medium, SourceHanSansHC-Medium
		SourceHanSans-Bold.ttc	SourceHanSans-Bold, SourceHanSansK-Bold, SourceHanSansSC-Bold, SourceHanSansTC-Bold, SourceHanSansHC-Bold, SourceHanSansHW-Bold, SourceHanSansHWK-Bold, SourceHanSansHWSC-Bold, SourceHanSansHWTC-Bold, SourceHanSansHWHC-Bold
		SourceHanSans-Heavy.ttc	SourceHanSans-Heavy, SourceHanSansK-Heavy, SourceHanSansSC-Heavy, SourceHanSansTC-Heavy, SourceHanSansHC-Heavy
Super OTC	All	*SourceHanSans.ttc*	SourceHanSans-ExtraLight, SourceHanSansK-ExtraLight, SourceHanSansSC-ExtraLight, SourceHanSansTC-ExtraLight, SourceHanSansHC-ExtraLight, SourceHanSans-Light, SourceHanSansK-Light, SourceHanSansSC-Light, SourceHanSansTC-Light, SourceHanSansHC-Light, SourceHanSans-Normal, SourceHanSansK-Normal, SourceHanSansSC-Normal, SourceHanSansTC-Normal, SourceHanSansHC-Normal, SourceHanSans-Regular, SourceHanSansK-Regular, SourceHanSansSC-Regular, SourceHanSansTC-Regular, SourceHanSansHC-Regular, SourceHanSansHW-Regular, SourceHanSansHWK-Regular, SourceHanSansHWSC-Regular, SourceHanSansHWTC-Regular, SourceHanSansHWHC-Regular, SourceHanSans-Medium, SourceHanSansK-Medium, SourceHanSansSC-Medium, SourceHanSansTC-Medium, SourceHanSansHC-Medium, SourceHanSans-Bold, SourceHanSansK-Bold, SourceHanSansSC-Bold, SourceHanSansTC-Bold, SourceHanSansHC-Bold, SourceHanSansHW-Bold, SourceHanSansHWK-Bold, SourceHanSansHWSC-Bold, SourceHanSansHWTC-Bold, SourceHanSansHWHC-Bold, SourceHanSans-Heavy, SourceHanSansK-Heavy, SourceHanSansSC-Heavy, SourceHanSansTC-Heavy, SourceHanSansHC-Heavy
Subset OTF	CN	*SourceHanSansCN-ExtraLight.otf*	SourceHanSansCN-ExtraLight
		SourceHanSansCN-Light.otf	SourceHanSansCN-Light
		SourceHanSansCN-Normal.otf	SourceHanSansCN-Normal
		SourceHanSansCN-Regular.otf	SourceHanSansCN-Regular
		SourceHanSansCN-Medium.otf	SourceHanSansCN-Medium
		SourceHanSansCN-Bold.otf	SourceHanSansCN-Bold
		SourceHanSansCN-Heavy.otf	SourceHanSansCN-Heavy
	TW	*SourceHanSansTW-ExtraLight.otf*	SourceHanSansTW-ExtraLight
		SourceHanSansTW-Light.otf	SourceHanSansTW-Light
		SourceHanSansTW-Normal.otf	SourceHanSansTW-Normal
		SourceHanSansTW-Regular.otf	SourceHanSansTW-Regular
		SourceHanSansTW-Medium.otf	SourceHanSansTW-Medium
		SourceHanSansTW-Bold.otf	SourceHanSansTW-Bold
		SourceHanSansTW-Heavy.otf	SourceHanSansTW-Heavy

Format	Language	File Name	PostScript Name/Names
Subset OTF (cont'd)	HK	*SourceHanSansHK-ExtraLight.otf*	SourceHanSansHK-ExtraLight
		SourceHanSansHK-Light.otf	SourceHanSansHK-Light
		SourceHanSansHK-Normal.otf	SourceHanSansHK-Normal
		SourceHanSansHK-Regular.otf	SourceHanSansHK-Regular
		SourceHanSansHK-Medium.otf	SourceHanSansHK-Medium
		SourceHanSansHK-Bold.otf	SourceHanSansHK-Bold
		SourceHanSansHK-Heavy.otf	SourceHanSansHK-Heavy
	JP	*SourceHanSansJP-ExtraLight.otf*	SourceHanSansJP-ExtraLight
		SourceHanSansJP-Light.otf	SourceHanSansJP-Light
		SourceHanSansJP-Normal.otf	SourceHanSansJP-Normal
		SourceHanSansJP-Regular.otf	SourceHanSansJP-Regular
		SourceHanSansJP-Medium.otf	SourceHanSansJP-Medium
		SourceHanSansJP-Bold.otf	SourceHanSansJP-Bold
		SourceHanSansJP-Heavy.otf	SourceHanSansJP-Heavy
	KR	*SourceHanSansKR-ExtraLight.otf*	SourceHanSansKR-ExtraLight
		SourceHanSansKR-Light.otf	SourceHanSansKR-Light
		SourceHanSansKR-Normal.otf	SourceHanSansKR-Normal
		SourceHanSansKR-Regular.otf	SourceHanSansKR-Regular
		SourceHanSansKR-Medium.otf	SourceHanSansKR-Medium
		SourceHanSansKR-Bold.otf	SourceHanSansKR-Bold
		SourceHanSansKR-Heavy.otf	SourceHanSansKR-Heavy

Glyph Set Particulars

Glyph Set & Region-specific Subsets

The number of glyphs in each font resource—except for the region-specific subset OTFs—is 65,535 (CIDs 0 through 65534), which is at the architectural limit for CID-keyed fonts (65,535 glyphs).

The table below indicates the number of glyphs that are included in the region-specific subset OTFs, whose figures include a common set of 3,240 glyphs that correspond to various characters, symbols, and punctuation. Also provided are the names of the subset definition files that can be found in the Resources folder of the release branch of this open source project.

Language	Glyphs	Subset Definition File	Supported Standards
Simplified Chinese	31,032	*AI0-SourceHanSans.CN*	All GB 18030 hanzi, all 8,105 hanzi of *Tōngyòng Guīfàn Hànzìbiǎo* (通用规范汉字表), 199 of which are outside of GB 18030
Traditional Chinese—Taiwan	20,945	*AI0-SourceHanSans.TW*	All Big Five hanzi (aka CNS 11643 Planes 1 and 2), all HKSCS-2016 hanzi (the glyphs may or may not adhere to the Taiwan MOE glyph standard), seven ETen hanzi, 15 additional hanzi
Traditional Chinese—Hong Kong	20,936	*AI0-SourceHanSans.HK*	All Big Five hanzi (aka CNS 11643 Planes 1 and 2), all HKSCS-2016 hanzi, seven ETen hanzi, 15 additional hanzi
Japanese	17,933	*AI0-SourceHanSans.JP*	All Adobe-Japan1-6 kanji (a superset of those in JIS X 0208, JIS X 0213 & JIS X 0212)
Korean	24,960	*AI0-SourceHanSans.KR*	All modern (11,172) and 500 high-frequency archaic hangul syllables, conjoining hangul jamo (with full archaic hangul support), all KS X 1001 and KS X 1002 hanja (7,476), 722 additional hanja

Of course, the font resources that include the full set of 65,535 glyphs support all of the standards that are listed in the above table, and employ some method of accessing the glyphs for different languages when they occupy the same Unicode code point and require a different shape.

The ordering file, *Al0-SourceHanSans*, which is provided in the Resources folder of the release branch of this open source project, lists all 65,535 CIDs in the first column, and shows the FDArray and row font structure in the second and third columns, respectively, along with the Unicode-based working glyph names in the fourth column. All 65,535 working glyph names are unique, and all—with the exception of the ones for CID+0 (the *.notdef* glyph) and CIDs 65485 through 65534—use a "uni" (BMP) or "u" (outside BMP) prefix followed by uppercase hexadecimal digits. Glyphs that are represented by (or can be considered) sequences are made up of concatenations of the appropriate Unicode-based glyph names. Identifiers for regions and other purposes are also used.

Weights

The table below shows sample glyphs in each of the seven weights, ranging from ExtraLight to Heavy. The ExtraLight and Heavy weights represent the master designs, and the five intermediate weights are the result of multiple master interpolation (the interpolation ratios are provided):

ExtraLight—0	Light—160	Normal—320	Regular—420	Medium—560	Bold—780	Heavy—1000
汉漢 漢한	汉漢 漢한	汉漢 漢한	汉漢 漢한	汉漢 漢한	汉漢 漢한	汉漢 漢한

Glyph Complement PDFs

Included in the GlyphComplements folder of the release branch of this open source project are seven Unicode-based glyph complement PDFs, one for each weight, that provide a visual synopsis of the UTF-32 'cmap' tables for each of the five languages: Japanese, Korean, Simplified Chinese, Traditional Chinese (Taiwan), and Traditional Chinese (Hong Kong). For each code point that maps to a glyph, there are three types of annotations, described as follows according to their position relative to the code-point box:

Upper-Left—Glyph width: **F** = Full-width, **H** = Half-width, **M** = Monospaced (hangul letters and syllables), **P** = Proportional, **Q** = Quarter-width, **T** = Tall (U+3031, U+3032, and the vertical forms of U+2E3A and U+2E3B), **W** = Wide (U+2E3A and U+2E3B), **Z** = Zero (non-spacing)

Upper-Right—Language (Region): **C** = Simplified Chinese (China), **H** = Traditional Chinese (Hong Kong), **J** = Japanese, **K** = Korean, **T** = Traditional Chinese (Taiwan)

Bottom—The CID of the glyph

Each glyph complement PDF contains five bookmarked 389-page sections, one for each language, meaning 1,945 pages in total. (There are actually 1,958 pages, because pp 779 through 791 represent the 500 pre-composed high-frequency archaic hangul syllables, which are described in the next paragraph.) Glyphs that are tall (T), wide (W), or non-spacing (Z) may exceed or appear outside the code-point box, which includes those for U+20DD, U+20DE, U+2E3A, U+2E3B, U+302A through U+302D, U+3031, and U+3032.

Also included in the glyph complement PDFs on pp 779 through 791 (13 pages) are the 500 pre-composed high-frequency archaic hangul syllables, ordered by their two- or three-character combining sequences, and bookmarked under the "Korean" bookmark.

Unencoded Glyphs

Not shown in the Unicode-based glyph complement PDFs are glyphs that are unencoded.

Ignoring code points that share different Simplified Chinese, Traditional Chinese, Japanese, Korean, and proportional/half-width glyphs, there are 3,167 unencoded glyphs in each 65,535-glyph font resource. The region-specific subset OTFs include considerably fewer unencoded glyphs.

Approximately one-third of the unencoded glyphs are Japanese ideographs (kanji), all of which represent kanji included in Adobe-Japan1-6. Some of these have been explicitly identified as JIS90 (JIS X 0208-1990) glyphs according to their source glyph names (CIDs 61896 through 62064; 169 glyphs) and are reflected in the 'jp90' GSUB feature that is specific to Japanese fonts and font instances, and the remainder have been identified according to their registered IVSes (CIDs 62065 through 63107; 1,043 glyphs) in the *Adobe-Japan1* IVD (Ideographic Variation Database) Collection, and are reflected in the Format 14 'cmap' subtable in the same fonts and font instances.

The bulk of the remaining unencoded glyphs are the 500 high-frequency archaic hangul syllables, the glyphs for combing jamo, vertical forms, and a small number of other variants.

Latin, Greek & Cyrillic Glyphs

Included in all font resources is a rich set of Latin glyphs that support not only ASCII and ISO/IEC 8859-1 (aka ISO Latin 1), but also the characters that are necessary for broadly-used CJK transliteration and transcription systems, along with those that are necessary for Latin-based Vietnamese. A basic set of glyphs for Greek and Cyrillic, with proportional metrics, is also included.

Source Han Sans Versus Source Sans Pro & Source Code Pro

The Latin, Latin-like, Greek, and Cyrillic glyphs in *Source Han Sans* are derived from—but not identical to—Source Sans Pro. The same is true for the half-width glyphs in terms of their relationship with Source Code Pro. The Latin and Latin-like glyphs in a typical CJK font represent a minority, and when it comes to harmonizing glyphs of different scripts, it is better to modify the minority to harmonize with the majority, and not vice versa. In addition, half-width glyphs in typical CJK fonts are also expected to be precisely half-width.

There are two primary differences between the glyphs that are common in *Source Han Sans* and *Source Sans Pro*:

- The interpolation ratios for the weights are different. *Source Han Sans* is available in seven weights: ExtraLight, Light, Normal, Regular, Medium, Bold, and Heavy. *Source Sans Pro* is available in six: ExtraLight, Light, Regular, Semibold, Bold, and Black. While some of the weight names are the same, one should not assume that the interpolation ratios are the same. They will be relatively close, but not precisely the same.
- The glyphs in *Source Han Sans* that are derived from *Source Sans Pro* have been adapted for use in *Source Han Sans*, which mainly involves scaling. In the case of the ExtraLight, Regular, and Heavy/Black weights, the *Source Sans Pro* glyphs were scaled to 110%, 113%, and 115%, respectively. Thus, the *Source Han Sans* glyphs appear to be slightly larger than those in *Source Sans Pro*, particularly in the heavier weights.

The half-width Latin glyphs in *Source Han Sans*, which are the default for the half-width (“HW”) OTFs and OTC font instances in only the Regular and Bold weights, and which are also exposed via the 'hwid' GSUB feature in the other OTFs and OTC font instances, are different from the glyphs in *Source Code Pro* as follows:

- Like *Source Sans Pro*, the interpolation ratios are different for all weights.
- The half-width Latin glyphs in *Source Han Sans* are precisely half-width, meaning half an em or 500-unit horizontal advances. The glyphs in *Source Code Pro* are monospaced, using 600-unit horizontal advances, meaning that they are not precisely half-width.

- The glyphs themselves are also different, particularly the one for zero (0) whose glyph in *Source Han Sans* lacks a center dot to more easily distinguish it from uppercase O, which is important when using a font to edit or display source code.

The table below compares *Source Han Sans* with *Source Sans Pro* and *Source Code Pro* for three weights, ExtraLight, Regular, and Heavy/Black:

Weight	Source Han Sans & Source Han Sans HW / 'hwid'	Source Sans Pro & Source Code Pro
ExtraLight	Unicode Version 11.0	↔ Unicode Version 11.0
	Unicode Version 11.0	↔ Unicode Version 11.0
Regular	Unicode Version 11.0	↔ Unicode Version 11.0
	Unicode Version 11.0	↔ Unicode Version 11.0
Heavy/ Black	**Unicode Version 11.0**	↔ **Unicode Version 11.0**
	Unicode Version 11.0	↔ **Unicode Version 11.0**

Vertical Glyphs

The usual and expected set of vertical glyphs is included, some of which are region- or language-specific. In addition, all glyphs for kana, meaning not only those for small kana, include a vertical glyph variant. A small number of vertical glyphs happen to be encoded for compatibility reasons, most of which can be found in the U+FE*xx* range, but they are still accessible via the 'vert' GSUB feature as vertical variants of the horizontal forms that are encoded elsewhere.

The pre-rotated non–full-width glyphs that are typically accessible via the effectively-deprecated 'vrt2' GSUB feature have been intentionally excluded from the glyph set.

CIDFont Resource & CFF Particulars

CIDFont Resource Structure

The font resources that include 65,535 glyphs began their life as an *Adobe-Identity-0* ROS CIDFont resource that includes 18 FDArray elements, each of which specifies its own hinting parameters. The table below shows the names of each of the 18 FDArray elements, its index, the CIDs and CID ranges that are included, and the total number of glyphs:

FDArray Name	Index	CIDs & CID Ranges	Glyphs
Alphabetic	0	59079–59104, 59111–59136	52
AlphabeticDigits	1	958–977, 59062–59071, 59218–59228	41
Bopomofo	2	1649–1691, 1801–1828, 65342	72
Dingbats	3	102, 111, 116, 149, 181, 245–253, 255–256, 720, 723–724, 731–732, 734–736, 741, 743, 747, 752–759, 762, 764–766, 768–769, 794–830, 832–957, 978–1077, 1238–1285, 1287–1292, 1294–1328, 1330–1354, 1384–1459, 1546–1549, 1553, 1644, 1785–1800, 1881–2440, 58979–59061, 59072–59078, 59105–59110, 59137–59141, 59205–59211, 59229–59318, 59321–59476, 63139–63152, 63279–63282, 65140–65156, 65243–65244, 65248, 65359–65460	1,570
DingbatsDigits	4	770–793	24

FDArray Name	Index	CIDs & CID Ranges	Glyphs
Generic	5	0, 1078–1237, 1286, 1293, 65485–65534	213
HDingbats	6	59212–59217	6
HKana	7	59142–59204	63
HWidth	8	63164–63179, 63190–63261	88
HWidthCJK	9	63262–63278	17
HWidthDigits	10	63180–63189	10
Hangul	11	372–627, 1692–1784, 47582–58854, 63283–65139	13,479
Ideographs	12	1357–1383, 1829–1864, 2441–47581, 58855–58973, 59477–61863, 61896–63107	48,922
Kana	13	1460–1545, 1550–1552, 1554–1643, 1645–1648, 1865–1880, 61872–61895	223
Proportional	14	1–101, 103–110, 112–115, 117–148, 150–180, 182–244, 254, 257–371, 628–719, 721–722, 725–730, 733, 737–740, 742, 744–746, 748–751, 760–761, 763, 767, 831, 1329, 1355–1356, 58974–58978, 59319–59320, 61864–61871, 63163	492
ProportionalCJK	15	63108–63138	31
ProportionalDigits	16	63153–63162	10
VKana	17	65157–65242, 65245–65247, 65249–65341, 65343–65358, 65461–65484	222

CFF Subroutinization

All 'CFF ' tables have been subroutinized. The size savings ranges anywhere from 1.5 to 3MB for the 65,535-glyph OTFs and OTCs. The ExtraLight weight exhibits the greatest size savings.

The AFDKO *tx* tool and its "-cff +S -no_futile" command-line options were used to convert the CIDFont resources into CFFs and to subroutinize them. The latest subroutinizer in *tx* is almost three orders of magnitude faster than the one that is currently in *makeotf*, meaning that the process takes a small number of minutes instead of several hours. The resulting subroutinized CFFs were subsequently spliced into the 'sfnt' font resources (aka OpenType/CFF fonts) using the *sfntedit* tool.

Unicode Particulars

Unicode Mappings

The Format 12 (UTF-32) 'cmap' subtable of each language-specific OTF and OTC specifies 44,806 meaningful mappings, and the region-specific subset OTFs obviously include less. Note that some glyphs map from multiple code points, such as the range U+2F00 through U+2FD5, along with a large chunk of the CJK Compatibility Ideographs. When the ten UTF-32 CMap resources are combined, a total of 62,368 glyphs are covered, which leaves 3,167 glyphs as being not directly unencoded. One of the mappings, U+31BB, is expected to be included in Unicode Version 13.0 (2020), but has been deemed stable enough to implement.

In addition to the ideographs for which there are obviously a large number of language-specific glyphs, the following code points also exhibit language-specific variation:

Unicode	Simplified Chinese	Traditional Chinese	Japanese	Korean
U+2018	‘	‘	‘	‘
U+2019	’	’	’	’
U+201C	“	“	“	“

Unicode	Simplified Chinese	Traditional Chinese	Japanese	Korean
U+201D	”	”	”	”
U+2264	≤	≤	≤	≤
U+2265	≥	≥	≥	≥
U+226E	≮	≮	≮	≮
U+226F	≯	≯	≯	≯
U+3001	、	、	、	、
U+3002	。	。	。	。
U+FF01	！	！	！	！
U+FF0C	，	，	，	，
U+FF0E	．	．	．	．
U+FF1A	：	：	：	：
U+FF1B	；	；	；	；
U+FF1F	？	？	？	？

Included in the Resources folder of the release branch of this open source project are the raw (aka human-readable, with one mapping per line) UTF-32 mapping files—named *utf32-cn.map*, *utf32-tw.map*, *utf32-hk.map*, *utf32-jp.map*, and *utf32-kr.map*—that are used to compile the UTF-32 CMap resources—named *UniSourceHanSansCN-UTF32-H*, *UniSourceHanSansTW-UTF32-H*, *UniSourceHanSansHK-UTF32-H*, *UniSourceHanSansJP-UTF32-H*, and *UniSourceHanSansKR-UTF32-H, respectively*—that the AFDKO *makeotf* tool uses to generate the Format 12 (UTF-32) 'cmap' subtables. Also included in this project are the raw UTF-32 mapping files—named *utf32hw-cn.map*, *utf32hw-tw.map*, *utf32hw-hk.map*, *utf32hw-jp.map*, and *utf32hw-kr.map*—that are used to compile the UTF-32 CMap resources that map ASCII (U+0020 through U+007E), U+00A0 ( ), U+00A5 (¥), U+00AD (-), U+2011 (-), U+20A9 (₩), and U+2423 (␣) to half-width forms—named *UniSourceHanSansHWCN-UTF32-H*, *UniSourceHanSansHWTW-UTF32-H*, *UniSourceHanSansHWHK-UTF32-H*, *UniSourceHanSansHWJP-UTF32-H*, and *UniSourceHanSansHWKR-UTF32-H*.

Matching UTF-16 CMap resources, which should not be used to build the OpenType/CFF fonts, are provided in the Resources folder of the release branch of this open source project for good measure.

Unicode Coverage

In addition to complete URO (*Unified Repertoire & Ordering*; up through U+9FEF for Unicode Version 11.0), Extension A, and modern hangul syllable coverage, the 65,535-glyph font resources completely cover the following 256-character Unicode blocks: U+00*xx*, U+11*xx*, U+2F*xx* through U+33*xx* (except for U+332C), U+D7*xx*, U+FF*xx*, and U+1F2*xx* (except for U+1F260 through U+1F265)

CJK Unified Ideographs Extension G

Although CJK Unified Ideographs Extension G (aka IRG Working Set 2015) is not yet encoded, and therefore its tentative Plane 3 code points are not yet stable, the language-specific OTFs include glyphs for four of its ideographs. (The region-specific subset OTFs do not include these glyphs.)

The table below lists these four ideographs by providing their IRG Working Set 2015 serial numbers (their source references are in parentheses), their language-specific forms, and the IDSes (*Ideographic Description Sequences*) that can be used to access their glyphs via the 'ccmp' GSUB feature that is generally enabled in most applications:

IRG	CN	TW	HK	JP	KR	IDS
02063 (UTC-01200)	KR	KR	KR	KR	㤅	⿰氵恩
04318 (UTC-01312)	邐	CN	CN	CN	CN	⿺辶⿱穴⿰月⿰幺⿱幺长⿱言马⿱幺长 刂心
04319 (UTC-00791)	邐	HK	邐	邐	邐	⿺辶⿱穴⿰月⿰幺⿱幺長⿱言馬⿱幺長 刂心
04752 (UK-02960)	龗雲龍雲龍龍	龗雲龍雲龍龍	龗雲龍雲龍龍	龗雲龍雲龍龍	JP	⿱雲⿰雲龍雲⿰龍龍

For those who wish to repurpose the IDSes in the last column of the above table, please be aware that there is a U+200A HAIR SPACE character immediately after the first IDC (*Ideographic Description Character*), which was necessary to prevent the 'ccmp' GSUB from being invoked, and as a result, the IDS is visible. Instead, simply select and copy the desired ideograph from the CN, TW, HK, JP, or KR column in the above table, which will copy the underlying three-, nine-, or 19-character IDS, then repurpose in other applications.

Unicode Variation Sequences

All font resources include 16 SVSes (*Standardized Variation Sequences*) that correspond to eight full-width CJK punctuation characters, and can be used to explicitly invoke them, thereby overriding the default glyphs. These SVSes will be included in Unicode Version 12.0 (2019), but have been deemed stable enough to implement. The table below lists these SVSes and the glyphs that correspond to them:

Unicode	VS1 (U+FE00)—Corner-Justified Form	VS2 (U+FE01)—Centered Form
U+3001	あ汉、汉あ	永、永
U+3002	あ汉。汉あ	永。永

Unicode	VS1 (U+FE00)—Corner-Justified Form	VS2 (U+FE01)—Centered Form
U+FF01	汉！汉	あ永！永あ
U+FF0C	あ汉，汉あ	永，永
U+FF0E	あ汉．汉あ	永・永
U+FF1A	汉：汉	あ永：永あ
U+FF1B	汉；汉	あ永；永あ
U+FF1F	汉？汉	あ永？永あ

The use of these SVSes is meant as a “plain text” alternative to language-tagging in order to achieve the same results, which means that subsequent language-tagging of such text, explicit or otherwise, will override the results in potentially unpredictable ways. In other words, these SVSes should be considered a last-resort means of displaying particular forms of these full-width CJK punctuation characters. Also see the table that starts on page 25 to learn how these eight full-width CJK punctuation characters are expected to behave in vertical writing mode.

Each Simplified Chinese font and font instance additionally includes nine SVSes that correspond to nine of the 1,002 Standardized Variants that were introduced in Unicode Version 6.3. Six of these SVSes are default (directly encoded). 14 of its 25 total UVSes are default, and the remaining 11 are non-default. The provided *SourceHanSans_CN_sequences.txt* file specifies the UVSes.

Each Traditional Chinese (Taiwan) font and font instance additionally includes two SVSes that correspond to two of the 1,002 Standardized Variants that were introduced in Unicode Version 6.3. Both of these SVSes are default (directly encoded). 10 of its 18 total UVSes are default, and the remaining eight are non-default. The provided *SourceHanSans_TW_sequences.txt* file specifies the UVSes.

Each Traditional Chinese (Hong Kong) font and font instance additionally includes 14 SVSes that correspond to 14 of the 1,002 Standardized Variants that were introduced in Unicode Version 6.3. 10 of its 30 total UVSes are default, and the remaining 20 are non-default. The provided *SourceHanSans_HK_sequences.txt* file specifies the UVSes.

All IVSes from the registered Adobe-Japan1 IVD Collection—except for <6CE8 E0102> (Adobe-Japan1-6 CID+12869), which is excluded because it is outside the scope of the *Source Han Sans* glyph set—are specified in the Format 14 'cmap' subtable of each Japanese font and font instance, along with 89 of the 1,002 Standardized Variants that were introduced in Unicode Version 6.3. This means that 14,682 *Adobe-Japan1* IVSes and 105 SVSes are included. 13,319 of these UVSes are default, meaning that the glyph is directly encoded, and the remaining 1,468 are non-default (unencoded or encoded in a CJK Compatibility Ideograph block, at least for Japanese fonts and font instances). The provided *SourceHanSans_JP_sequences.txt* file specifies the UVSes.

Each Korean font and font instance additionally includes 270 SVSes that correspond to 270 of the 1,002 Standardized Variants that were introduced in Unicode Version 6.3. All of these SVSes are default (directly encoded). The 36 IVSes from the registered KRName IVD Collection are also supported. 293 of its 322 total UVSes are default, and the remaining 29 are non-default. The provided *SourceHanSans_KR_sequences.txt* file specifies the UVSes.

Glyph Sharing Statistics

One of the defining characteristics of Pan-CJK typeface designs is the significant sharing of glyphs across its supported languages. However, in order to honor regional conventions, some code points, in particular those for ideographs, may require more than one glyph per code point. Of course, the URO exhibits the greatest variation, in terms of including a large number of code points that require multiple language-specific glyphs. In general, as one progresses through the CJK Unified Ideograph extensions, from Extension A through Extension F, the number of code points that require multiple language-specific glyphs diminishes.

The table below lists several code point categories, and shows how many glyphs are used to represent the 44,806 code points:

			CJK Unified Ideograph Extensions						
		URO	A	B	C	D	E	F	Other
Glyphs	1	8,783	6,126	2,080	47	34	111	5	14,703
	2	7,641	444	28					192
	3	3,747	12						47
	4	739							1
	5	66							

Of particular interest should be the 66 highlighted URO code points that have five unique glyphs, one per language. The table below shows these 66 ideographs for each of the five supported languages:

Simplified Chinese	傑僭割廱匾叟喝塌姿贏幰 廋扇扉摩溲潛瀛瘦瞎磨窖 竇箭篠簉糙綢蠯贏翁鶼翩 盲贏艘花裯褐謁譖豁贏轄 返迷途造週遍遭選遼鄰釁 閟雕靠靡颼飯驎鬣魔麗麟

Traditional Chinese—Taiwan	傑僭割廱匾叟喝塌姿嬴幰 廋扇扉摩溲潛瀛瘦瞎磨窖 竇箭篠簉糙綢鼕羸翁翦翩 肓羸艘花裯褐謁譖豁贏轄 返迷途造週遍遭選遼鄰釁 閟雕靠靡颼飯驎鬣魔麗麟
Traditional Chinese—Hong Kong	傑僭割廱匾叟喝塌姿嬴幰 廋扇扉摩溲潛瀛瘦瞎磨窖 竇箭篠簉糙綢鼕羸翁翦翩 肓羸艘花裯褐謁譖豁贏轄 返迷途造週遍遭選遼鄰釁 閟雕靠靡颼飯驎鬣魔麗麟
Japanese	傑僭割廱匾叟喝塌姿嬴幰 廋扇扉摩溲潛瀛瘦瞎磨窖 竇箭篠簉糙綢鼕羸翁翦翩 肓羸艘花裯褐謁譖豁贏轄 返迷途造週遍遭選遼鄰釁 閟雕靠靡颼飯驎鬣魔麗麟

Korean	傑僭割廱匾叟喝塌姿嬴幰 廋扇扉摩溲潛瀛瘦瞎磨窖 竇箭篠簉糙綢蠶羸翁翦翩 肓羸艘花裯褐謁譖豁羸轄 返迷途造週遍遭選遼鄰釁 閥雕靠靡颼飯驎鬣魔麗麟

The Resources folder in the release branch of this open source project includes a two-sheet and six-column spreadsheet, named *region-map.xlsx*, which shows the glyphs that are used for each code point. The first sheet covers the URO (U+4E00 through U+9FEF), and the second one covers Extension A (U+3400 through U+4DB5). In order to make clear whether a glyph is shared, and with what language or languages, the primary language of the glyph is shown instead of the glyph itself. For only the KR, TW, and HK columns, if a code point is outside the scope of the KS standards, Big Five, or Big Five plus HKSCS-2016, respectively, a "#" is shown after the region code to indicate that the code point is outside the scope of those particular standards. The same data is also provided in "plain text" (UTF-8) format as the file named *region-map-utf8.txt*.

UAX #50 Compliance

Source Han Sans is one of the first font implementations that is compliant with UAX #50 (*Unicode Vertical Text Layout*). Only the substitutions in the 'vert' GSUB feature are expected to be used, and the 'vrt2' GSUB feature, which is a subset of the 'vert' GSUB feature, is included only because some environments, such as Windows and some Microsoft applications, require it to be present. In particular, pre-rotated non–full-width glyphs have been excluded from the 'vrt2' GSUB feature, and substitutions for arrows and arrow-like characters have also been excluded from both GSUB features.

Language Particulars

Simplified Chinese: GB 18030 & Tōngyòng Guīfàn Hànzìbiǎo

In addition to supporting GB 18030, which primarily amounts to Simplified Chinese glyphs for all URO and Extension A code points plus six Extension B code points, China's latest list of 8,105 hanzi (通用规范汉字表 *Tōngyòng Guīfàn Hànzìbiǎo*), which includes 196 additional Extension B through E code points, along with three that were appended to the URO for 199 in total, is also supported. Among these 199 hanzi, 36 map to Extension B, 44 map to Extension C, eight map to Extension D, 108 map to Extension E, and three have been appended to the URO (U+9FCD through U+9FCF).

For those who wish to develop versions of the fonts that support the 24 PUA (*Private Use Area*) code points required by GB 18030, please reference the *utf32-gb18030pua24.map* file that is provided in the Resources folder of the release branch of this open source project.

Traditional Chinese—Taiwan: Big Five + CNS 11643 Planes 1 & 2

Beginning with this release, there are separate Traditional Chinese fonts and font instances for Taiwan and Hong Kong. The scope of Traditional Chinese for Taiwan is limited to Big Five (equivalent to CNS 11643 Planes 1 and 2), and the glyphs mostly adhere to the Taiwan MOE (Ministry of Education) glyph standard. Any CJK Unified Ideograph code point that is outside the scope of Big Five is not likely to display appropriately for Traditional Chinese (Taiwan) use.

Traditional Chinese—Hong Kong: Big Five & HKSCS-2016

The Traditional Chinese (Hong Kong) fonts and font instances include as their scope Big Five and HKSCS-2016, with the glyphs for both mostly adhering to Hong Kong conventions.

Japanese: Adobe-Japan1-6 Correspondence Table & JIS Coverage

The *aj16-kanji.txt* mapping file that is provided in the Resources folder of the release branch of this open source project shows how all Adobe-Japan1-6 kanji map to working glyph names as specified in the fourth field of the included *AI0-SourceHanSans* ordering file. In order to support the *Adobe-Japan1* IVD Collection, glyphs for all Adobe-Japan1-6 kanji—except for <6CE8 E0102> (Adobe-Japan1-6 CID+12869), which is excluded because it is outside the scope of the *Source Han Sans* glyph set—are included.

Due to the JIS standard coverage of Adobe-Japan1-6 that is inherited by *Source Han Sans*, all JIS X 0208, JIS X 0213, and JIS X 0212 kanji are therefore supported. JIS2004 (aka JIS X 0213:2004) glyphs are the default for the relevant code points. A small number of characters in the JIS standards, such as those for IPA, along with additional Latin, Greek, and Cyrillic that were not deemed necessary, have been intentionally excluded.

Although *Source Han Sans* includes the same kanji as Adobe-Japan1-6, including a large number of kanji variants, compatibility shouldn't be expected for documents that were authored using applications that specify glyphs by CID. The only Adobe-Japan1-6 compatibility that should be expected is at the Unicode level, which includes the *Adobe-Japan1* IVSes that are specified in the Format 14 'cmap' subtable.

Korean: Hangul Glyphs & Hanja Coverage

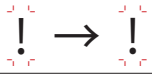
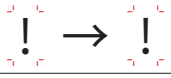
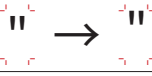
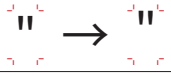
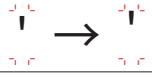
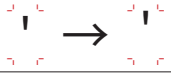
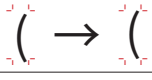
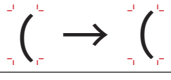
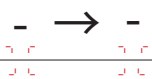
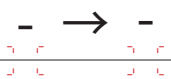
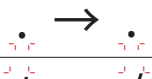
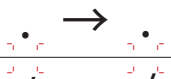
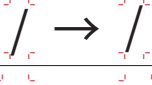
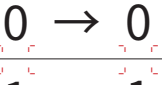
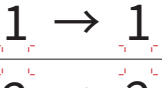
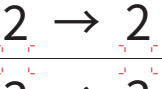
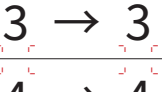
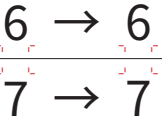
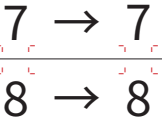
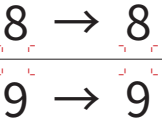
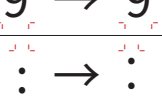
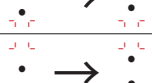
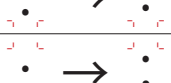
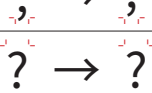
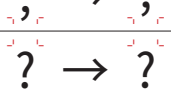
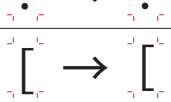
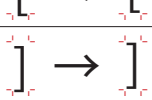
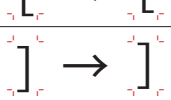
Glyphs for all modern Korean hangul symbols, letters (including compatibility versions), and syllables are included, along with the additional glyphs necessary to compose archaic hangul via the 'ljmo', 'vjmo', and 'tjmo' GSUB features. Also included are glyphs for 500 high-frequency archaic hangul syllables in pre-composed form, which are made accessible via the 'ccmp' GSUB feature.

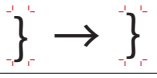
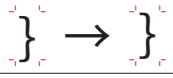
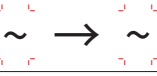
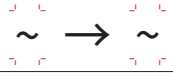
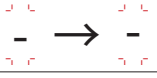
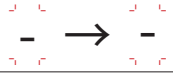
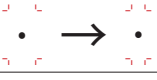
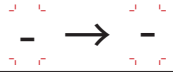
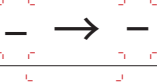
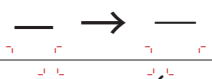
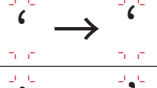
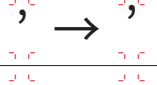
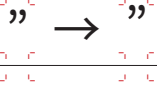
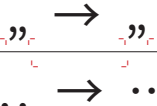
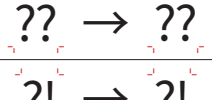
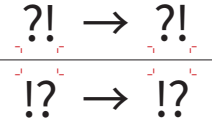
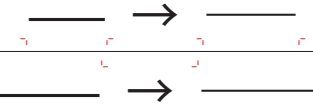
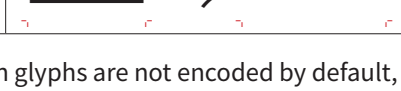
While the horizontal advances of the glyphs for Korean hangul symbols (in the U+32*xx* block) are full-width (1000 units), those for Korean hangul letters and syllables are monospaced at 920 units.

The *ks-hanja.txt* mapping file that is provided in the Resources folder of the release branch of this open source project shows how the hanja in the KS X 1001 (4,620) and KS X 1002 (2,856) standards map to working glyph names as specified in the fourth field of the included *AI0-SourceHanSans* ordering file.

Proportional & Half-Width CJK Punctuation

Included in these fonts are special forms of the proportional and half-width punctuation shown in the table below, which have been tailored for CJK use in that they are aligned to the em-box, not to Latin features, and which are accessible via the 'locl' GSUB feature:

Unicode	Western versus CJK		Chinese	Japanese	Korean
	Proportional	Half-Width[1]			
U+0020					Yes
U+0021					Yes
U+0022			Yes	Yes	Yes
U+0027			Yes	Yes	Yes
U+0028					Yes
U+0029					Yes
U+002C					Yes
U+002D					Yes
U+002E					Yes
U+002F					Yes
U+0030[2]			Yes	Yes	Yes
U+0031[2]			Yes	Yes	Yes
U+0032[2]			Yes	Yes	Yes
U+0033[2]			Yes	Yes	Yes
U+0034[2]			Yes	Yes	Yes
U+0035[2]			Yes	Yes	Yes
U+0036[2]			Yes	Yes	Yes
U+0037[2]			Yes	Yes	Yes
U+0038[2]			Yes	Yes	Yes
U+0039[2]			Yes	Yes	Yes
U+003A					Yes
U+003B					Yes
U+003F					Yes
U+005B					Yes
U+005D					Yes

Unicode	Western versus CJK		Chinese	Japanese	Korean
	Proportional	Half-Width[1]			
U+007B					Yes
U+007D					Yes
U+007E					Yes
U+00AD					Yes
U+00B7					Yes
U+2011					Yes
U+2013					Yes
U+2014			Yes	Yes	Yes
U+2018			Yes[2]	Yes	Yes
U+2019			Yes[2]	Yes	Yes
U+201A			Yes	Yes	Yes
U+201C			Yes[2]	Yes	Yes
U+201D			Yes[2]	Yes	Yes
U+201E			Yes	Yes	Yes
U+2026[2]			Yes	Yes	Yes
U+203C			Yes	Yes	Yes
U+2047			Yes	Yes	Yes
U+2048			Yes	Yes	Yes
U+2049			Yes	Yes	Yes
U+2E3A			Yes	Yes	Yes
U+2E3B			Yes	Yes	Yes

1 The half-width glyphs are not encoded by default, except for the Regular and Bold OTFs and OTC font instances that include the “HW” identifier in their names, and are accessible via the 'hwid' GSUB feature in all OTFs and OTC font instances .

2 These characters are unique in that the CJK forms are encoded by default, and the Western forms are accessible when the text is language-tagged for English.

3 The default glyph for this code point is full-width, not proportional, so the 'pwid' GSUB feature must first be invoked to access the proportional glyph that is tailored for CJK use.

OpenType Particulars

Menu Names

The table below shows the English and localized Family names for each font and font instance:

Configuration	Family Name—English	Family Name—Localized
Simplified Chinese OTF & OTC	*Source Han Sans SC*	思源黑体
	Source Han Sans HW SC	思源黑体 HW
Traditional Chinese OTF & OTC—Taiwan	*Source Han Sans TC*	思源黑體
	Source Han Sans HW TC	思源黑體 HW
Traditional Chinese OTF & OTC—Hong Kong	*Source Han Sans HC*	思源黑體 香港
	Source Han Sans HW HC	思源黑體 香港 HW
Japanese OTF & OTC	*Source Han Sans*	源ノ角ゴシック
	Source Han Sans HW	源ノ角ゴシック HW
Korean OTF & OTC	*Source Han Sans K*	본고딕
	Source Han Sans HW K	본고딕 HW
Simplified Chinese OTF (subset)	*Source Han Sans CN*	思源黑体 CN
Traditional Chinese OTF (subset)—Taiwan	*Source Han Sans TW*	思源黑體 TW
Traditional Chinese OTF (subset)—Hong Kong	*Source Han Sans HK*	思源黑體 HK
Japanese OTF (subset)	*Source Han Sans JP*	源ノ角ゴシック JP
Korean OTF (subset)	*Source Han Sans KR*	본고딕 KR

For the region-specific subset OTFs, the English and localized menu names also include two-letter region identifiers. Their PostScript names, as shown in the table in the “Font Resources” section on page 5, use the same two-letter region identifiers. With the exception of Japanese, the language-specific OTFs and OTCs include a one- or two-letter language identifier only for their English menu names.

The Regular weight in all fonts and font instances is style-linked to the Bold weight. For applications that support style-linking, the Regular weight becomes the Bold weight if the “Bold” style is selected, and the Bold weight may not appear in the font menu.

Because the OTCs and the corresponding language-specific OTFs specify identical PostScript and Family names, they cannot be installed in the same environment.

OpenType Tables

All font resources, with the exception of the OTCs, include the following 17 OpenType tables: 'BASE', 'CFF ', 'DSIG', 'GDEF', 'GPOS', 'GSUB', 'OS/2', 'VORG', 'cmap', 'head', 'hhea', 'hmtx', 'maxp', 'name', 'post', 'vhea', and 'vmtx'. The OTCs do not include a 'DSIG' table.

The five or ten font instances in each of the seven OTCs share the following ten OpenType tables: 'BASE', 'CFF ', 'GDEF', 'VORG', 'hhea', 'hmtx', 'maxp', 'post', 'vhea', and 'vmtx'. The following six OpenType tables are not completely shared by the five or ten font instances in each OTC: 'GPOS', 'GSUB', 'OS/2', 'cmap', 'head', and 'name'. The Super OTC shares OpenType tables more efficiently.

OpenType Table Overrides

Several values in particular OpenType tables have been overridden from their otherwise default values. The subsections below detail some of the more important table-specific overrides that have been applied.

The 'OS/2' Table

The *OS/2.sTypoLineGap* value has been set to 0 (zero) units, and is also reflected in the *hhea.LineGap* and *vhea.lineGap* values. The *OS/2.usWinAscent* and *OS/2.usWinDescent* values have been calculated by removing excessively tall and other vertical-only glyphs—for U+2E3A, U+2E3B, U+302A through U+302D, U+3031, and U+3032—from the equation, and have been harmonized across all seven weights. These same harmonized settings are also reflected in the *hhea.Ascender* and *hhea.Descender* values. This is for the benefit of applications that use these values for determining default leading. These and other 'OS/2' table settings are intended to provide consistent cross-platform line spacing (aka vertical metrics).

The 'name' Table

Unlike mainstream OpenType/CFF CJK fonts, a *name.ID=20* string is not specified because there are no legacy (non-Unicode) encodings that meaningfully correspond to these fonts. In addition, the 'name' table does not include any Macintosh (*PlatformID=1*) strings, which was accomplished by invoking the AFDKO *makeotf* tool's "-omitMacNames" command-line option. This means that the 'name' table includes only Unicode strings.

The 'vmtx' Table

In addition to specifying alternate vertical origins for full-width Latin and Latin-like glyphs that rest on the Latin baseline, proper vertical origins and vertical advances are also specified for the glyphs that correspond to U+3031 and U+3032, and to the vertical forms of U+02D9, U+2E3A, U+2E3B, U+302E, and U+302F.

OpenType GSUB Features

All fonts and font instances include the OpenType GSUB features (see the OpenType Feature Registry for additional information) as detailed in the table below:

	OTF & OTC					Subset OTF				
GSUB Feature	SC	TC	HC	J	K	CN	TW	HK	JP	KR
aalt	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
calt	Yes	Yes	Yes	Yes	Yes					Yes
ccmp	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
dlig	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
fwid[1]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
hist	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
hwid[1]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
jp78[1]				Yes					Yes	
jp83[1]				Yes					Yes	
jp90[1]				Yes					Yes	
liga	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ljmo	Yes	Yes	Yes	Yes	Yes					Yes
locl	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
nlck[1]				Yes					Yes	
pwid[1]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ruby[1]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	OTF & OTC					Subset OTF				
GSUB Feature	SC	TC	HC	J	K	CN	TW	HK	JP	KR
tjmo	Yes	Yes	Yes	Yes	Yes					Yes
vert[1]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
vjmo	Yes	Yes	Yes	Yes	Yes					Yes
vrt2[2]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

1 This GSUB feature is inherited by the 'aalt' GSUB feature.
2 This GSUB feature is a subset—not superset, as is usually the case—of the 'vert' GSUB feature.

All fonts and font instances that include hangul letters or syllables include a Korean-specific glyph for the *space* character (U+0020), whose width is set to 280 units for all weights. The width of the proportional glyph for the *space* character ranges from 220 units (in ExtraLight) to 229 units (in Heavy) in all fonts and font instances, except for the half-width fonts and font instances that include the “HW” identifier in their names and whose *space* glyph is 500 units. A contextual substitution in the 'calt' GSUB feature substitutes the proportional *space* glyph with the Korean-specific version only when it is surrounded by a hangul syllable (modern or archaic), regardless of whether it is pre-composed or combining.

The 'ccmp' GSUB Feature

The 'ccmp' GSUB feature is used to form the appropriate glyphs that correspond to the sequences needed to support a small number of kana, many of which are included in JIS X 0213 but are intentionally unencoded, along with the 500 high-frequency pre-composed archaic hangul syllables. PDF specimens for the latter can be found in the GlyphComplements folder of the release branch of this open source project. A complete listing of the former is provided in the table below:

Unicode Sequence	Glyph	Present in JIS X 0213
<004D 0300>	M̀	
<004D 0304>	M̄	
<006D 0300>	m̀	
<006D 0304>	m̄	
<00CA 0304>	Ê̄	
<00CA 030C>	Ê̌	
<00EA 0304>	ê̄	
<00EA 030C>	ê̌	
<3042 3099>	あ゙	
<3044 3099>	い゙	
<3048 3099>	え゙	
<304A 3099>	お゙	
<304B 309A>	か゚	Yes

Unicode Sequence	Glyph	Present in JIS X 0213
<304D 309A>	ぎ	Yes
<304F 309A>	ぐ	Yes
<3051 309A>	げ	Yes
<3053 309A>	ご	Yes
<3093 3099>	ん	
<30A2 3099>	ア	
<30A4 3099>	イ	
<30A8 3099>	エ	
<30AA 3099>	オ	
<30AB 309A>	カ	Yes
<30AD 309A>	キ	Yes
<30AF 309A>	ク	Yes
<30B1 309A>	ケ	Yes
<30B3 309A>	コ	Yes
<30BB 309A>	セ	Yes
<30C4 309A>	ツ	Yes
<30C8 309A>	ト	Yes
<30F3 3099>	ン	
<31F7 309A>	ㇷ	Yes

This GSUB feature is also used to support the two- and three-character sequences shown in the table below:

Unicode Sequence	Unicode	Western Glyph	CJK Glyph
<2014 2014 2014>	U+2E3B		
<2014 2014>	U+2E3A		
<2015 2015 2015>	U+2E3B		
<2015 2015>	U+2E3A		

Unicode Sequence	Unicode	Western Glyph	CJK Glyph
<3033 3035>	U+3031		
<3034 3035>	U+3032		

The glyphs that result from the first four sequences can be overridden, in terms of Western versus CJK glyph style, by applying the 'locl' GSUB feature, which entails using an application that supports this OpenType feature and properly language-tagging the text. These sequences merely serve as a convenience mechanism for environments that do not support language tagging.

See the table in the “CJK Unified Ideographs Extension G” section on page 13 for details about four additional characters that are handled via the 'ccmp' GSUB feature.

The 'locl' GSUB Feature

The 'locl' GSUB feature plays a critical role in the language-specific OTFs in that it represents the *only* mechanism within the font resource for accessing the glyphs for the non-default languages. If the 'locl' GSUB feature is not supported or not properly used, the default glyphs are used. Each non-default language is handled via a separate lookup that is associated with the appropriate language and script, and one of its purposes is to mimic the 'cmap' table of the target language.

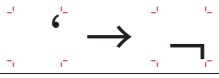
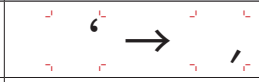
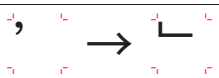
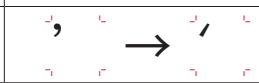
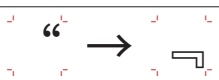
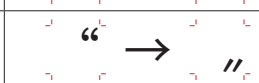
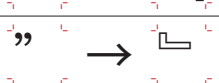
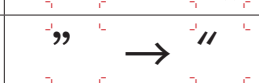
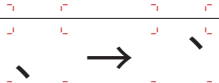
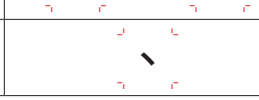
The font instances of the OTCs also include the 'locl' GSUB feature, but its presence represents an alternate method for accessing the glyphs for the non-default languages that does not involve selecting a different font instance of the OTC.

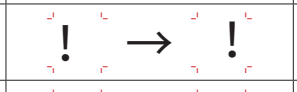
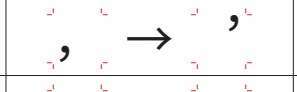
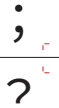
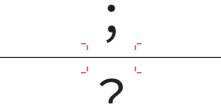
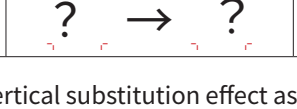
Note that in addition to using an application that supports the 'locl' GSUB feature, such as Adobe InDesign or modern browsers (examples include Chrome, Edge, Firefox, and Safari), the text must also be properly language-tagged at the character, paragraph, or document level.

Also note that all font resources, including the region-specific subset OTFs, include the 'locl' GSUB feature. For the region-specific subset OTFs that obviously do not include glyphs for the ideographs of the non-supported regions, the 'locl' GSUB feature instead operates only on a small number of glyphs for punctuation by tailoring them for CJK use. See the table in the “Proportional & Half-Width CJK Punctuation” section on page 18 for a complete listing of these special glyphs.

The 'vert' GSUB Feature

The 'vert' GSUB feature includes substitutions that may be different for each language, which apply to the following code points:

Unicode	Simplified Chinese	Traditional Chinese	Japanese	Korean
U+2018[1]				
U+2019[1]				
U+201C[1]				
U+201D[1]				
U+3001				

Unicode	Simplified Chinese	Traditional Chinese	Japanese	Korean
U+3002				
U+FF01				
U+FF0C				
U+FF0E				
U+FF1A				
U+FF1B				
U+FF1F				

1 To achieve the same vertical substitution effect as Traditional Chinese for Japanese or Korean, the 'fwid' GSUB feature must first be applied to change the default proportional glyphs to their full-width forms.

OpenType GPOS Features

Eight GPOS features, 'halt', 'kern', 'mark', 'palt', 'vert', 'vhal', 'vkrn', and 'vpal', are included in all font resources, and their details are listed below (see the OpenType Feature Registry for additional information):

- The 'halt' and 'vhal' GPOS feature are identical across all weights.
- The 'palt' and 'vpal' GPOS features for ExtraLight through Medium are identical, as are those for Bold and Heavy. These features cover the glyphs for kana, some full-width punctuation, some full-width symbols, full-width digits, and full-width Latin.
- The 'kern' GPOS feature includes weight-specific kerning pairs for proportional Latin, Greek, and Cyrillic glyphs, along with weight-independent kerning pairs for kana and some punctuation. The 'vkrn' GPOS feature includes only weight-independent kerning pairs for kana and some punctuation.
- The 'mark' GPOS feature is used for handling bopomofo tone mark placement.
- The 'vert' GPOS feature is used to adjust the glyphs for U+20DD, U+20DE, U+302A through U+302D, U+3099, U+309A, U+3127, U+31B4 through U+31B7, U+31BB, and those that represent vowels and trailing consonants of combining jamo, which have zero-unit horizontal advances, such that their positions and vertical origins are appropriate for vertical writing.

Noto Sans CJK Differences

Other than by name, the fonts for the Google-branded version of this typeface design, Noto Sans CJK, differ in the following ways:

- The *name.ID=0* (*Copyright notice*) string does not include a reference to the “Source” name.
- The *name.ID=7* (*Trademark*) string specifies a Google trademark.
- The *name.ID=11* (*URL Vendor*) string specifies a Google URL.
- A *name.ID=12* (*URL Designer*) string was added that specifies an Adobe URL.
- Although localized menu names are not specified, localized 'name' table strings are included, in terms of *name.IDs* 1, 2, 4, 16, and 17, though the actual strings are identical to the English-language ones.
- The *OS/2.usWeightClass* value is set to 100 for Thin (*Source Han Sans ExtraLight* uses 250) per Noto CJK Issue #86.
- The *OS/2.achVendID* tag is set to *GOOG* (*Source Han Sans* uses *ADBO*).

- Although glyphs for U+2252 and U+25C8 are included, named *uni2252* (CID+858) and *uni25C8* (CID+1261), respectively, the Format 4 and 12 'cmap' subtables do not include mappings for U+2252 and U+25C8, per Noto CJK Issue #24 and Noto CJK Issue #23, respectively.

Changes

Version 2.000

Build Date: November 2, 2018. Built By: Dr. Ken Lunde (小林劍). Release Date: November 19, 2018.

Compared to the previous release—Version 1.004 that was released on 2015-06-16—a large number of changes were made, far too many to list here. Listed below are some of the more significant changes that were made in this release:

- A second flavor of Traditional Chinese, for Hong Kong and supporting the HKSCS-2016 standard, was added, which increased the total number of font resources by 16, from 72 to 88.
- 155 new mappings have been added to the CMap resources. 66 are from BMP code points, 22 are from Plane 1 code points, and the remaining 67 are from Plane 2 code points. Among the 67 new Plane 2 code points, 57 are from Extension B, two are from Extension C, three are from Extension E, and the remaining five are from Extension F.
- As a result of removing approximately 1,750 glyphs in order to make room for approximately 1,750 new glyphs, the CID assignments of the glyphs necessarily—and drastically—changed. The CID assignments of exactly 200 glyphs are unchanged from Version 1.004: 0–107, 2570–2633, 47223–47232, 47262–47272, 47281–47286, and 65484.
- The Traditional Chinese form of the Radical #162 辶 component was improved.
- The URO is complete up through U+9FEF (Unicode Version 11.0).
- The glyphs for some of the kana were tweaked.
- The glyphs and support for bopomofo, along with their tone marks, were improved. This involved adding the 'GDEF' (*Glyph Definition*) table, the 'mark' (*Mark Positioning*) GPOS feature, and the 'ruby' (*Ruby Notation Forms*) GSUB feature.
- The language and script declarations in the 'locl' and 'vert' GSUB features were improved.
- The 13-page glyph synopsis PDFs for the 500 pre-composed high-frequency hangul syllables have been incorporated into the Unicode-base glyph synopsis PDFs, and are bookmarked under the “Korean” bookmark.
- Blank placeholder glyphs for U+32FF, *uni32FF* (CID+2184) and *uni32FF-V* (CID+65359), are included. This character has been reserved for the two-ideograph square ligature that represents the name of Japan’s forthcoming new era which starts on 2019-05-01, and will be the only character added in Unicode Version 12.1.
- Like *Source Han Serif*, the CIDFont and CMap resources do not include XUID arrays.
- Like *Source Han Serif*, there are no mappings for the range U+0000 through U+001F.
- Like *Source Han Serif*, the code points that correspond to Halfwidth Jamo variants map to glyphs that correspond to code points in the Hangul Compatibility Jamo block. In other words, the glyphs for half-width jamo have been removed.
- Like *Source Han Serif*, the 'name' table does not includes any Macintosh (*PlatformID=1*) strings.
- Like *Source Han Serif*, the Regular weight is now style-linked to the Bold weight. This means that the Bold weight may not appear in the font menu, particularly when using applications that support style-linking as a way to make text bold.
- Like *Source Han Serif*, the 'vert' GPOS feature is included.
- Like *Source Han Serif*, the deprecated 'hngl' (*Hangul*) GSUB feature is not included in the fonts and font instances whose default language is Korean.

Known Issues

Please report all issues in the GitHub repository so that they can be properly tracked and addressed, and for greater visibility among the user community. The Wiki also conveys some useful information about upcoming releases. Also, be sure to thoroughly check the closed issues prior to submitting a new issue, being sure to exercise the search feature.

Because these fonts exercise several architectural limits, particularly the ones that include 65,535 glyphs, some environments may have difficulties using them properly, sometimes due to implementation limits or poor assumptions. If this is the case, please report such issues so that they can be recorded and tracked. You are also strongly encouraged to contact the developer of such environments to report the same.

General

- While not an issue per se, all of the fonts include placeholder (aka blank) glyphs for U+32FF, *uni32FF* (CID+2184) and *uni32FF-V* (CID+65359), which is the code point that has been reserved for the two-ideograph square ligature form of the name of Japan's forthcoming new era that takes effect on 2019-05-01. U+32FF is expected to be included in Unicode Version 12.1. The purpose of including the placeholder glyphs is to facilitate a dot-release shortly after the official announcement of the era name is made.
- The glyphs for the four CJK Unified Ideographs Extension G ideographs, which are made accessible via the 'ccmp' GSUB feature using their IDSes, will be mapped from the appropriate Plane 3 code points as soon as their code points have been deemed stable enough to implement.

Western

- None

Chinese—Simplified & Traditional

- None

Simplified Chinese

- None

Traditional Chinese—TW & HK

- None

Traditional Chinese—TW

- None.

Traditional Chinese—HK

- None

Japanese

- None

Korean

- None

That is all.