The following is a proposal for documenting and ensuring two invariants between properties.

1. [:Ideographic:] is a superset of [:Unified_Ideograph:].

I propose that we document that the containment as an invariant that programmers can depend on. (It would be a bizarre to have something be a [:Unified_Ideograph:] yet not be [:Ideographic:].)

BTW, [:Ideographic:] contains 1,016 characters not in [:Unified_Ideograph:], namely [〆〇一丿丿]{壹-殻塚晴-羽姫諸逸都飯-〇-〇-〇-〇}. Most of these are compatibility (not NFKC), except for 11 outliers.

The 11 NFKC characters are: [〆〇一]

昶 U+3006 IDEOGRAPHIC CLOSING MARK
〇 U+3007 IDEOGRAPHIC NUMBER ZERO

9 HANGZHOU NUMERALs (all but 3)

The 1,005 non-NFKC characters are: [々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々際に]

1,002 CJK COMPATIBILITY IDEOGRAPHs (all)
3 HANGZHOU NUMERALs:

□ U+3038 HANGZHOU NUMERAL TEN toNFKC=十
□ U+3039 HANGZHOU NUMERAL TWENTY toNFKC=卄
□ U+303A HANGZHOU NUMERAL THIRTY toNFKC=卅

2. [:scx=Han:] is a superset of [:Ideographic:]

I propose documenting the containment as an invariant that programmers can depend on. All of the Ideographic are really Han Ideographic. Remember that [:scx=Han:] is the set of characters with Han as at least one of its SCX values: the characters in that set may have other scripts also.

BTW, [:scx=Han:] contains 603 characters not in [:Ideographic:], namely [々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々際に]. Most of these are compatibility (not NFKC), except for 210 outliers.

The 210 NFKC characters are: [々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々々際に]

113 CJK RADICALs (all but 2).
36 CJK STROKES (all)
18 COUNTING ROD UNIT DIGITS (all)
2 IDEOGRAPHIC ANNOTATION characters (Kanbun)
41 misc punctuation/symbols

The **393** non-NFKC characters are: [母ー一龠毁חלק1月12月13日1日1日]

- 母 U+2E9F CJK RADICAL MOTHER
- □ U+2EF3 CJK RADICAL C-SIMPLIFIED TURTLE

214 KANGXI RADICALs (all)
14 IDEOGRAPHIC ANNOTATION characters (Kanbun, all but 2)
36 PARENTHESIZED IDEOGRAPHs (all)
49 CIRCLED IDEOGRAPHs (all)
12 IDEOGRAPHIC TELEGRAPH MONTHs (all)
25 IDEOGRAPHIC TELEGRAPH SYMBOL FOR HOURs (all)
31 IDEOGRAPHIC TELEGRAPH SYMBOL FOR DAYs (all)
4 SQUARE ERA NAMEs (all)
1 SQUARE CORPORATION (all)
5 HALFWIDTH punctuation variants

3. [:scx=han:] is *almost* a superset of [:name=/IDEOGRAPH/:]

There is no particular requirement that it be a complete superset—we all know that the names are not necessarily indicative. However, since there are only 16 outliers, namely [〜〜〜〜〜〜], and they are (overwhelmingly) used in writing systems that use Han (with possibly one exception), we could consider whether any of them should also be in [:scx=han:].

The **12** NFKC characters are: [〜]

12 IDEOGRAPHIC DESCRIPTION CHARACTERs (all)

Arguably these should be [:scx=Han:] and [:Ideographic:], since that’s their fundamental function.

The **4** non-NFKC characters are: [\u3000・°]

All are [:scx=Common:] and [:Ideo=N:]

- U+3000 IDEOGRAPHIC SPACE
- U+FE11 PRESENTATION FORM FOR VERTICAL IDEOGRAPHIC COMMA
- U+FE12 PRESENTATION FORM FOR VERTICAL IDEOGRAPHIC FULL STOP
- U+FE51 SMALL IDEOGRAPHIC COMMA

Arguably the last 3 should be [:scx=Han:]; their corresponding NFKC characters are already [:scx=Han:]:

- U+3001 IDEOGRAPHIC COMMA
- U+3002 IDEOGRAPHIC FULL STOP