Title: Proposal to add HTMLElement to UAX #45
Author: William T. Nelson
Date: 2021-09-23

UTC-01005「ヲ」 and 「ヰ」 are under consideration for disunification from U+5F50「ヲ」. This document is a proposal to facilitate this disunification by adding 「ヲ」 to UAX #45 with the glyph and data shown below.

<table>
<thead>
<tr>
<th>Glyph</th>
<th>RS</th>
<th>TS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ヲ</td>
<td>58.0</td>
<td>3</td>
</tr>
</tbody>
</table>

Additionally, it is advisable that the UAX #45 record for UTC-01005「ヲ」 have its status field changed from U to N, and that its Unicode code point, U+5F50, be removed.

**Evidence for 「ヲ」**

「ヲ」 is used as ① a variant form of KANGXI RADICAL SNOUT 「ヲ」 (radical 58) and ② as a component deriving from either historical 「ヲ」 (snout) or 「ヰ」 (hand). The following dictionary entry differentiates the "snout" and "hand" origins, but using indistinguishable (and so unifiable) glyphs.

The source is 大字典 by UEDA Kazutoshi (上田万年) as reprinted by Harvard UP in 1942.
Evidence for 「ヨ」 (continued)

The source is 角川 新字源。
「ヲ」は、動物の頭の形。これを部首として、「ヲ」を含む漢字が集められる。「ヲ」は、「又（＝て）」の変形。冠になるときは、特に「けいがしら」「いのこがしら」と呼ぶ。「けいがしら」の名称は、「ヲ」の音から。「いのこがしら」の名称は、「ヲ」が、豚やイノシシの頭を表したことから。

The source is 漢字源。
Evidence for 「ㄜ」 (continued)

The source is 漢字海。
Evidence for 「猬」 (continued)

**KEIGASHIRA**

4204  ▶

<table>
<thead>
<tr>
<th>4-3-1</th>
<th>Radical</th>
<th>U2E95</th>
</tr>
</thead>
<tbody>
<tr>
<td>kald D</td>
<td>S3-3-0</td>
<td>F0000</td>
</tr>
</tbody>
</table>

radical variant → see 猬 at 4205

**KEIGASHIRA**

4205  ▶

<table>
<thead>
<tr>
<th>4-3-1</th>
<th>Radical</th>
<th>U5F50</th>
</tr>
</thead>
<tbody>
<tr>
<td>kald D</td>
<td>S3-3-0</td>
<td>F0000</td>
</tr>
</tbody>
</table>

**RADICAL**

*keigashira* ‘pig’s head’, used for character classification

Variants: 猬 *keigashira* (猬 彙 象)

The source is the *Kodansha Kanji Dictionary.*
Additional background

In Unicode 13 and later, CJK RADICAL SNOT TWO in the CJK Radicals Supplement block has the same actual shape as 「ヨ」.

\[
\begin{align*}
\text{2E95} & \quad \equiv \quad \text{CJK RADICAL SNOT TWO} \\
& \quad \to \quad \text{5F50} \quad \equiv
\end{align*}
\]

However, CJK RADICAL SNOT TWO looked like UTC-01005 「ヨ」 in Unicode 12.1 and earlier versions of the standard.

\[
\begin{align*}
\text{2E95} & \quad \equiv \quad \text{CJK RADICAL SNOT TWO} \\
& \quad \to \quad \text{5F50} \quad \equiv
\end{align*}
\]

Fonts display CJK RADICAL SNOT TWO inconsistently, but hopefully the encoding of 「ヨ」 and 「ヨ」 as separate unified ideographs will allow for these basic shapes to be used without ambiguity in the future.