

Title: **Emoji and Vendors: for consideration by the UTC**

From: **Mark Davis**

Date: **2017-06-28**

Draft: <https://goo.gl/uabhGB>

I propose that we explain more on our site the role that major vendors play in the selection of new emoji characters for Unicode, due to the fact that new emoji need to be widely available on all mobile phones and computers. The following provides background information for such explanations.

The selection process for most Unicode characters differs from that used for emoji characters. These differences are highlighted in the following sections, with an emphasis on the role that major vendors play in selection and implementation.

Section 1. Non-emoji characters, such as the letter A, or the [Nüshu characters recently added to Unicode 10.0](#), existed as text before their inclusion into Unicode. The process for those characters involves research into their fundamental ontology: when and how to determine when a set of glyphs can be represented by a single abstract character, and what are the values of the the over 100 different Unicode properties that determine the behavior of characters in text processing, input, and display. For complex scripts, it also requires specification of the characters' interactions: the most appropriate character-glyph model for those characters.

For non-emoji characters, implementation considerations that are important for vendors do play a strong role in the specification of behavior. Without a suitable character-glyph model and appropriate properties, characters would be unlikely to be implemented well, or at all. However, the decision to add Nüshu characters, for example, does not depend on whether or not major vendors plan to support them. They are encoded so that the languages and texts that use them can be interchanged and preserved, even if they never appear on mobile phones.

The primary goal for the Unicode encoding is to include all the characters of the world's languages, past, present, and future.

Section 2. Emoji characters are quite different. There was no pre-existing letter or symbol 🧑 that Unicode needed to add in order to represent text. There is an unlimited number of possible images of things, and there is no goal to encode them all, or even any large number of them. The goal for emoji is *not* to capture and encode images in the way that [Nüshu](#) characters were added.

Rather, the goal is to progressively add sets of emoji that will be popular, with the vehicle being major platforms (such as Android, iOS, Windows, Twitter, and Facebook) that can ensure wide distribution of the proposed emoji. Such platforms supply all of the capabilities described in [UTS #51](#), such as input, display, and editing. While there is no hard and fast line, for new emoji the goal is on the order of millions of daily active users, not thousands.

The vendors for such platforms have made it clear to the Consortium that they want no more than 50-100 new emoji characters each year. Essentially every mobile device ends up having to support them, which can have an impact on the memory of these devices, and on the complexity of the UI needed to support inputting the emoji. It would be pointless — and wasted effort — to include an emoji that did not end up being supported by major vendors. That emoji would simply displace a

different one that *could* be supported by vendors.

Thus for our final set of emoji each year, we work with major vendors involved in Unicode to prioritize the emoji candidates so that we get consensus from a majority of major vendors that each emoji in the prioritized set will be widely supported.

Background. The emoji selection process is described in [Submitting Emoji Proposals](#), with some more detail in [Comments in response to L2-17/147](#). The submission page describes how anyone can submit a proposal, what factors are used to assess the proposals, and the timeline. The submission form includes provision for people to provide as much data as they can to make the case for the expected popularity. The emoji subcommittee also looks at the “incremental value” of each addition; if an existing emoji can already approximately convey the intention of a proposed emoji, that makes the case less compelling.