

On the possibility of encoding non-text items as Unicode characters

William Overington

15 January 2009

There are a number of planes in the Unicode code point map that are presently unused. The rules for encoding characters include, as I understand it, that only text characters and symbols that can be used in text can be encoded. Yet, if the Unicode Technical Committee published a statement, possibly after a Public Review of the topic, that it would consider proposals to encode, at character code points, as if characters, items that are not text and are not symbols that can be used in text, would that result, after a period of time of, say, two or three years, in proposals that would greatly enhance the possibilities of using information technology?

I have experimented with various systems that use Private Use Area code points for things such as vector graphics, a portable interpretable object code and the eutofont font format. However, those are each an experimental system, trying out the concept: systems to be encoded in regular Unicode would need much research and development effort by many people in order to achieve the best systems possible. Yet, I think that such non-text items encoded as if they were characters could potentially be of great general usefulness in the future.

Thus this agenda item is to ask the Unicode Technical Committee to consider please whether the opportunity should be made available so that people may develop such systems with the knowledge that such systems could become encoded in regular Unicode.