Title:	**Comments on draft Named Composite Entities**
Doc. Type:	Expert contribution
Source:	Peter Constable, Microsoft
Action:	For consideration by UTC
References:	L2/03-400
Distribution:	UTC members

Document L2/03-400 provides a draft data file for "Named Composite Entities". I do not know the exact details on what led to the preparation of this data file. I see serious concerns in a file of this form being published as part of the standard. It has exactly the same format as UnicodeData.txt, with entries such as the following:

```
F0000;KHMER CONSONANT SIGN COENG KA;Lo;0;L;17D2 1780;;;;N;;;;;
```

Each record is for a presentation form. The intent appears to be to document certain presentation forms, but in this draft file, they are presented in exactly the same format one would expect to see an encoded character in the UCS. This makes it appear that these are encoded characters for presentation forms.

Moreover, there are numeric identifiers assigned to each record that are valid USVs, and in particular, code points in the plane-15 private use area. This makes it appear that the Unicode Consortium has, in fact, encoded characters in the private use area.

I have no objection to providing a list of presentation forms, including unique identifiers and the character sequences used to represent them. We must not give the impression, however, of having encoded these presentation forms, or of having assigned characters within any of the private use areas.