

Version 18 d'Unicode

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 18 septembre 2026

<https://www.bortzmeyer.org/unicode-18-0.html>

Le 16 septembre 2026 est sortie la version 18 d'Unicode <<https://blog.unicode.org/2026/09/announcing-unicode-standard-version-180.html>>. Une description officielle des principaux changements est disponible <<https://www.unicode.org/versions/Unicode18.0.0/>> mais voici ceux qui m'ont intéressé particulièrement. (Il n'y a pas de changement radical.)

Pour explorer plus facilement la grande base Unicode, j'utilise un programme qui la convertit en SQL <<https://www.bortzmeyer.org/unicode-to-sql.html>> et permet ensuite de faire des analyses variées. Faisons quelques requêtes SQL :

```
ucd=> SELECT count(*) AS Total FROM Characters;
      total
-----
      172873
```

Combien de caractères sont arrivés avec la version 18?

```
ucd=> SELECT version,count(version) FROM Characters GROUP BY version ORDER BY version::float;
...
15.0    |    4489
15.1    |     627
16.0    |    5185
17.0    |    4803
18.0    |   13007
```

13[Caractère Unicode non montré ¹]007 nouveaux caractères, c'est beaucoup par rapport aux précédentes versions, mais ce sont presque tous des sinogrammes du petit sceau. Autrement, quels sont ces nouveaux caractères?

1. Car trop difficile à faire afficher par L^AT_EX

```

ucd=> SELECT To_U(codepoint) AS Code_point, name FROM Characters WHERE version='18.0' ORDER BY Codepoint;
code_point | name
-----+-----
...
U+20C3     | UAE DIRHAM SIGN
U+20C4     | OMANI RIAL SIGN
...
U+A7E2     | LATIN CAPITAL LETTER R WITH LONG LEG
...
U+10EF3    | ARABIC SMALL HIGH NOON WITH FATHA
...
U+12477    | CUNEIFORM NUMERIC SIGN ASH TIMES ONE DISH TENU
...
U+18E17    | JURCHEN CHARACTER-18E17
...
U+1FA8B    | METEOR
U+1FA8C    | ERASER
U+1FA8D    | NET WITH HANDLE
U+1FACC    | MONARCH BUTTERFLY
U+1FADD    | PICKLE
...
U+3D005    | SEAL CHARACTER-3D005
...

```

On y voit quelques nouveaux caractères, ainsi que les trois nouvelles écritures de cette version d'Unicode, les anciens chiffres cunéiformes, le jurchen et le petit sceau. Et il y a aussi quelques changements de comportement <https://www.unicode.org/versions/Unicode18.0.0/#UTS_Changes>, qui nécessiteront un peu de mise à jour des bibliothèques logicielles.

En plus futile, il y a l'habituelle arrivée <<https://www.unicode.org/emoji/charts-18.0/emoji-released.html>> d'emojis comme le cornichon (comme dans la chanson <<https://www.youtube.com/watch?v=WYMSjg3RxYk>>) et le météore. Le trognon de pomme n'a toujours pas été intégré (il avait déjà été retiré d'Unicode 17).

Si vous avez les bonnes polices de caractères, vous allez pouvoir voir quelques exemples (sinon, le lien mène vers Uniview <<https://r12a.github.io/uniview/>>). Voici par exemple, le rial omanais [Caractère Unicode non montré] <<https://r12a.github.io/uniview/?char=20c4>>, une nouvelle lettre R utilisée pour le xhosa, [Caractère Unicode non montré] <<https://r12a.github.io/uniview/?char=a7e2>>, une marque utilisée pour l'arabe, un chiffre cunéiforme [Caractère Unicode non montré] <<https://r12a.github.io/uniview/?char=1247e>>, une lettre jurchen [Caractère Unicode non montré] <<https://r12a.github.io/uniview/?char=18e42>>, une lettre du petit sceau [Caractère Unicode non montré] <<https://r12a.github.io/uniview/?char=3e949>>, la gomme [Caractère Unicode non montré] <<https://r12a.github.io/uniview/?char=1fa8c>> et le monarque [Caractère Unicode non montré] <<https://r12a.github.io/uniview/?char=1facc>>.