



Autres trucs

- Accueil
- Seulement les RFC
- Seulement les fiches de lecture
- Mon livre
- « Cyberstructure »
- Ève

Recherche dans ce blog :

Version 18 d'Unicode

Première rédaction de cet article le 18 septembre 2026

Le 16 septembre 2026 est sortie la [version 18 d'Unicode](#). Une description officielle des principaux changements est [disponible](#) 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](#) 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
...
15.0      |    4489
15.1      |     627
16.0      |    5185
17.0      |    4803
18.0      |   13007
```

13 007 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 ?

```
ucd=> SELECT To_U(codepoint) AS Code_point, name FROM Characters
      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](#), qui nécessiteront un peu de mise à jour des bibliothèques logicielles.

En plus futile, il y a l'[habituelle arrivée](#) d'[emojis](#) comme le cornichon (comme dans [la chanson](#)) 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](#)). Voici par exemple, le [rial omanais](#) ر, une nouvelle lettre R utilisée pour le [xhosa](#), R, une marque utilisée pour l'[arabe](#), un

chiffre cunéiforme , une lettre jurchen , une lettre du petit sceau , la gomme  et le monarque .

[Version PDF de cette page](#) (mais vous pouvez aussi imprimer depuis votre navigateur, il y a une feuille de style prévue pour cela)

[Source XML de cette page](#) (cette page est distribuée sous les termes de la licence [GFDL](#))

---

Si vous aimez, vous pouvez payer avec [Bitcoin](#) : adresse 1HtNJ6ZFUc9yu9u2qAwB4tGdGwPQasQGax (ou voyez le [code QR](#)). Pour toute remarque sur ce blog, s'adresser par courrier à Stéphane Bortzmeyer <stephane+blog@bortzmeyer.org> ou sur le [fédivers](#) à [bortzmeyer@mastodon.gougere.fr](#). Je suis [les règles de Crocker](#) donc pas besoin de faire des excès de diplomatie. Ce blog est strictement personnel et les opinions exprimées ici n'engagent donc que moi, et notamment pas mon employeur présent ou mes employeurs passés ou mes éventuels employeurs futurs.