

# تعريب نظام اليونيكود

ائتلاف يونيكود هو منظمة غير ربحية تأسست لتطوير وتشجيع استخدام معيار يونيكود، تشمل العضوية على طيف واسع من الشركات والمنظمات في صناعة الحوسبة ومعالجة البيانات. تأسست عام 1991 (ما هو الترميز الموحد يونيكود؟، [الموقع الرسمي ليونيكود](#)). نفس السنة التي اجتمع فيها أعضاء المنظمة مع ممثلين لمنظمة المعايير الدولية (آيزو)، طال هذا الاجتماع عدة أشهر حيث وحدت جهود المنظمين لتنتهي هذه الاجتماعات بظهور الشكل الحديث لليونيكود، الذي يمثل بطريقتين اثنتين:

(يو سي إس-4) و (يو سي إس-2)، حيث يعبر الرقم عن عدد البايت في كل حرف، 4 بايت في الأول و 2 في الثاني، و تقسم ال 32 بت الى 32,000 مستوى، يحوي كل مستوى على 65,000 حرف، و يحتوي التمثيل الثاني (الثنائي البايت) على جميع المستوى صفر (<https://www.unicode.org/versions/Unicode1.0.0/V2ch01.pdf>).

وصل عدد الأحرف (2024/2/13) 149,813 حرفاً، حسب آخر تحديث (يونيكود 15.1.0).

بما أن اللغة العربية أحد أكبر اللغات في العالم بعدد المتحدثين لها الذي يبلغ 550 مليون نسمة حول العالم، نحو 300 مليون منهم تعد العربية لغتهم الأم ونحو 250 مليوناً يتخذونها لغة ثانية ([اللغة العربية.. أصلها وتاريخها وعدد الناطقين بها، الجزيرة](#))، و أهميتها التاريخية و الدينية و كونها أحد اللغات الرسمية للأمم المتحدة، حيث [تتوفر كل الوثائق الرسمية بالعربية](#) (مكتبة داج همرشولد)، فنحن بحاجة الى نظام يوحد الرسم و الترميز للكتابة العربية لأجهزة الحاسب، و جاء اليونيكود حلاً لهذا و لكثير من اللغات أيضاً.

اللغة العربية من أول اللغات المضافة في معيار يونيكود، حسب الإصدار يونيكود 1.0.0 حيث تلي العبرية بالسطر 060C الى السطر 06F9 و تسبق الكتابة الديفانجارية (أحد الكتابات لشبه القارة الهندية)، في الإصدار الأخير (يونيكود 15.1.0) تقع العربية في المستوى صفر (المستوى الأساسي BMP - BasicMultilingualPlane)، تشغل المساحة ما بين 0600 و 06FF، تستعمل العربية أيضاً لكتابة العديد من اللغات الأخرى، حيث تستعمل العربية كأساس لبناء نظم الكتابة للغاتٍ أخرى، كالأوردو و الفارسية و البشتونية و أخرى ([Middle East-I Modern and Liturgical Scripts](#)).

## Figure 9-1. Directionality and Cursive Connection

Memory representation: ه ه ه ه ه

After reordering: ه ه ه ه ه

After joining: ه ه ه ه ه

تمثل الأحرف العربية في الذاكرة كسواها من البيانات الرقمية بدءاً من اليسار الى اليمين، و يجب استعمال خوارزمية يونيكود ثنائية الاتجاه، لترتب الأحرف قبل الطباعة، كل حرف عربي يمثل برمز واحد، مهما تعددت اشكاله حسب ما قبله و ما بعده، و يمثل في جدول اليونيكود بشكله و هو منفرد، تشكّل الأحرف عن طريق كتابة الحركات كرموز منفصلة ذات ترميز خاص، حيث لا يعبر كل حرف ذا تشكيل كرمز مستقل، و لكن كسلسلة من الرموز التي يشكلها اليونيكود باستعمال خوارزمياته، يقسم التشكيل الى:

-الحركات

-التنوين

-الشدة

-السكون

مع إن الألف المهموزة رمز مميز ذا ترميز مميز (U+0654, U+0655)، الا أن نظام اليونيكود يقدم طريقة لإضافة الهمزة لتسهيل الكتابة لمتحدثي اللغات الشرقية<sup>[1]</sup>، يمكن أيضاً أن تُستعمل مدة الألف كحركة (U+0653) لكتابة بعض النصوص كالقرآن، حيث تظهر علامة المد على غير الألف، و لكن مثل الألف المهموزة، يتوفر رمز مستقل لهذه الحالة من الألف (U+0622).

[1] اللغات شرق جزيرة العرب، التي تستعمل النظام العربي للكتابة شمولاً أو جزءاً.