المحاضرة الثالثة: تشفير المعلومات ومعالجتها

مقدمة

يقوم الحاسوب بالتعامل مع المعلومات (برمجيات وبيانات) بشكل ثنائي: سلسلة من الأصفار والآحاد (**001010101010101**)، ونظرًا لأنه من المستحيل تمامًا قراءة هذا الترميز بالنسبة لنا، يتم تحويل هذا الترميز عادةً إلى نظام سداسي عشري.
في الحياة اليومية، نستخدم النظام العشري، أي ترميز مكون من عشرة أرقام من **0** إلى **9**، أما في أجهزة الحاسوب، يتم تخزين كل شيء في شكل أولي وهو "البايت" **Byte** أو الأوكتي **Octet** : سلسلة من ثمانية أرقام ثنائية متتالية (إما صفر أو واحد) ؛

نظام اليونيكود **Unicode**

من الواضح أنه من المستحيل أن تقرأ العين البشرية سلسلة الآحاد والأصفار، لنكون قادرين على قراءة هذه القيم بسهولة أكبر، يتم تحويلها غالبًا إلى نظام الست عشري**Hexadecimal** .

كل بايت هو تشفير لأحد المحارف (**Character**) التي يتعامل معها الحاسوب.

مثال: **0100 0010** هي الطريقة التي يتم بها تشفير الحرف "**B**" في جهازك. ولكل رمز كتابي من رموز لغات العالم، لها تشفير معيّن. نسمي هذه التقنية في تشفير البيانات التي يتعامل معها الحاسوب باليونيكود **Unicode** بمعنى نظام التشفير العالمي.

نظام اليونيكود:Unicode هو معيار دولي لتمثيل النصوص والحروف والرموز المختلفة في جميع اللغات والكتابات المستخدمة حول العالم. يتيح اليونيكود تمثيل الحروف والرموز بأكثر من 128 ألف رمز مختلف، بما في ذلك الحروف المستخدمة في اللغات المختلفة والرموز المستخدمة في الرياضيات والعلوم والرموز التعبيرية مثل الإيموجي.

في نظام **Unicode**، تم تخصيص مجموعة واسعة من الشفرات الرقمية لتمثيل الحروف العربية والنصوص ذات الصلة. تم تعيين هذه الأكواد في المدى المعروف باسم "مجموعة الأحرف العربية" في **Unicode**.

تتراوح الأكواد الرقمية للحروف العربية في **Unicode** بين **U+0600** و **U+06FF** تحت مجموعة حروف وعلامات الشرق الأوسط. تشمل هذه المجموعة الحروف الأساسية والتشكيلات العربية والأرقام العربية والشرطة والأرقام المؤشرية والرموز الخاصة بالعربية والعديد من الرموز الأخرى المستخدمة في الكتابة العربية.

فمثلا حرف "ب" يُمثَّل بالأكواد التالية:
الست عشري: **0628**
الثنائي: **0000 0000 0110 1000**

حرف "ت" يُمثَّل بالأكواد التالية:
الست عشري: **A062**
الثنائي: **0000 0000 0110 1010**

فمن خلال ترميز الحروف العربية في **Unicode**، يتم تمكين الحواسيب والبرامج من التعامل بشكل صحيح مع النصوص والمحتوى العربي، وذلك بفضل توفير تمثيل شامل وموحد للحروف العربية في البيئة الرقمية.

وبالتالي فالبايت أو الأوكتي، فهي الوحدة الأساسية للمعلومات التي يتعامل معها الحاسوب، ولهذه الوحدة مضاعفات:

1 كيلوبايت KB يساوي 1,024 بايت

1 ميجابايت MB يساوي 2 ^{20} أي 2 مرفوعة للقوة 20 ويساوي 1,048,576 بايت

1 جيجابايت GB يساوي 2 ^{30} أي 2 مرفوعة للقوة 30 ويساوي 1,073,741,824 بايت

1 تيرابايت TB يساوي 2 ^{40} يساوي 1,099,511,627,776 بايت

حجم البيانات والبرامج : يقصد به المساحات التي تستهلكها تلك البيانات أو البرامج عند وضعها على قرص صلب وهو يقاس بوحدة البايت (الأوكتي) أو إحدى وحداتها المضاعفة. تكون البيانات والبرامج مشمولةً في ملفات معينة.

عند تخزين البيانات والبرامج في ملفات، يتم تخصيص مساحة محددة على وسائط التخزين لكل ملف. يعتمد حجم الملف على حجم البيانات أو الشفرة المصدرية الموجودة في الملف، بالإضافة إلى الهياكل التنظيمية اللازمة لتخزين الملف بشكل صحيح.

يجب ملاحظة أن حجم البرنامج يشمل عادة حجم الشفرة المصدرية والبيانات التشغيلية والمكتبات المستخدمة وغيرها من المكونات الضرورية لتشغيل البرنامج بشكل كامل.

أنواع الملفات: لكل ملف اسم يتضمّن امتدادا يحدد نوعية البيانات التي يشتمل عليها، هذا الامتداد هو عبارة عن توسيع يضاف إلى اسم الملف ويفيد في تحديد تنسيق ونوع المحتوى المخزن في الملف. وفيما يلي بعض أمثلة أشهر أنواع الملفات والامتدادات المرتبطة بها:

1 -ملفات النصوص:

ملفات النص العادي: يستخدم امتداد .txt أو .text.
ملفات الشفرة المصدرية: يستخدم امتدادات مثل .c، .cpp، .java، .py..

2- ملفات الصور:

صورة بامتداد JPEG: يستخدم امتداد .jpg أو .jpeg.
صورة بامتداد PNG: يستخدم امتداد .png.
صورة بامتداد GIF: يستخدم امتداد .gif.
صور بامتداد BMP

3- ملفات الصوت:

ملفات الصوت MP3: يستخدم امتداد .mp3.
ملفات الصوت WAV: يستخدم امتداد .wav.
ملفات الصوت FLAC: يستخدم امتداد .flac.

4-ملفات الفيديو:

ملفات الفيديو MP4: يستخدم امتداد .mp4.
ملفات الفيديو AVI: يستخدم امتداد .avi.
ملفات الفيديو MKV: يستخدم امتداد .mkv.

5-ملفات الوثائق:

ملفات Adobe PDF: يستخدم امتداد .pdf.
ملفات Microsoft Word: يستخدم امتداد .doc أو .docx.
ملفات Microsoft Excel: يستخدم امتداد .xls أو .xlsx.

6-ملفات الويب :
تشير إلى الملفات المستخدمة في تطوير وعرض صفحات الويب. تتضمن ملفات الويب العديد من الأنواع المختلفة التي تعمل معًا لتشكيل صفحة ويب كاملة. إليك بعض أمثلة أهم ملفات الويب وامتداداتها:

- ملفات HTML (Hypertext Markup Language): يستخدم امتداد .html أو .htm. يتضمن هذا النوع من الملفات هيكل الصفحة والعناصر المختلفة فيها.

- ملفات CSS (Cascading Style Sheets): يستخدم امتداد .css. يستخدم ملف CSS لتحديد تنسيقات التصميم والمظهر الخاص بالصفحة.

- ملفات JavaScript: يستخدم امتداد .js. يستخدم ملف JavaScript لإضافة السلوك التفاعلي والوظائف الديناميكية إلى صفحة الويب.

- ملفات الأنماط (Stylesheets): تشمل ملفات الأنماط المستخدمة في تصميم صفحات الويب مثل ملفات Bootstrap (.css) وملفات Sass (.scss، .sass) وملفات LESS (.less) وملفات أخرى.

هذه مجرد أمثلة قليلة لأنواع الملفات والامتدادات المشهورة. هناك العديد من أنواع الملفات وامتداداتها المختلفة والتي تختلف اعتمادًا على نوع البيانات المخزنة في الملف.

ضغط البيانات Data compression :

هو عملية تشفير المعلومات حيث تأخذ حيزا قليلا من المساحة لتحسين كفاءة النقل والتخزين، هناك نوعان رئيسيان لضغط البيانات:

- ضغط البيانات بدون خسائر: يتم ضغط البيانات دون فقدان أي معلومات. تستخدم هذه التقنية في العديد من تنسيقات الملفات مثل ZIP وRAR، وتستخدم أيضًا في البرامج المضغوطة وصور الأقراص المضغوطة. تعتمد تقنيات ضغط البيانات بدون خسائر على الاستفادة من النمط والتكرار في البيانات لتحقيق ضغط فعال.

- ضغط البيانات مع خسائر: تستخدم هذه التقنية في حالات حيث يمكن التخلص من بعض المعلومات الدقيقة التي لا تؤثر بشكل كبير على الجودة النهائية للبيانات. يتم استخدام ضغط البيانات مع خسائر في تنسيقات الصوت والفيديو، مثل MP3 وJPEG. يتم تطبيق خوارزميات معقدة لإزالة التفاصيل غير الضرورية والحفاظ على الجودة العامة للبيانات.

خاتمة

في ختام هذا الدرس، ندرك أن التشفير ونظام اليونيكود وحجم البيانات وضغطها هي مفاهيم هامة في عالم تكنولوجيا المعلومات الحديث. إنها تتعلق بطرق تمثيل وتخزين البيانات وتأمينها، وتلعب دورًا حاسمًا في تحقيق التواصل وتعزيز التوافق وتوفير كفاءة عالية لمعالجة ونقل البيانات.

من خلال استخدام أنظمة التكويد المناسبة مثل اليونيكود، نحن قادرون على تمثيل المحتوى اللغوي والرموز والأحرف من مختلف اللغات والثقافات بطريقة قياسية وموحدة. وهذا يساهم في تحقيق التواصل الفعال وحل المشاكل الناجمة عن التوافق اللغوي.

أما بالنسبة لحجم البيانات، فإنه يشكل تحديًا مستمرًا في عالم يتزايد فيه حجم المعلومات بشكل هائل. نظرًا لزيادة كمية البيانات المنتجة والمتاحة، فإن تخزينها ومعالجتها يتطلب حلولًا فعالة ومستدامة. وهنا تأتي تقنيات ضغط البيانات وتجزئتها وتخزينها المناسب للعب دورها في تقليل حجم البيانات وتحسين كفاءة النقل والتخزين.

في النهاية، يمكننا القول بأن فهم التشفير واليونيكود وحجم البيانات أمور حاسمة للتعامل مع التحديات الحديثة في مجال تكنولوجيا المعلومات.