تبـــادل المعطيـــات

ذ. محمد مونجي(*)

مدخل عام

من المعروف أن كل عمل ناجح يعتمـد علـى قـدر مناسب من المعلومات وهذا ينطبق على كافة الأنشطة في مختلف أنواع المعارف البشرية، وعلى هذا الأساس فإن المعلومات جزء أساسي من تقدم وازدهار الأمة ومن هنـا ينبغي وضع خطة وسياسة واضحـة للمعلومـات، علـى المستويين القطري والقومـي، تعكـس حاجـات كافـة الشرائح الاجتماعية وتغطي كافة الموضوعات الاقتصادية والثقافية والعلمية والتربويـة وغيرهـا مـن الموضوعـات، لتكون ركيزة من ركائز خطط التنمية القُطرية والقومية، وإن وضـع وصياغـة وتنفيـذ هـذه السياسـة تعـد مـن الخطوات والدعائم الأساسية لبنوك المعلومـات الوطنيـة والعربية، وتستطيع هذه البنوك والشبكات تأمين وصول المعلومات الموثقة إلى المخططين والعـاملين في النشـاطات والمجالات المطلوبة.

ونستطيع القول أن الهدف الرئيسي لشبكة ونظام المعلومـات علـى المسـتويين الوطـني والعربـي هـو خلـق مجتمع متعلـم ومتقـدم، مـزود بسـلاح العلـم والمعرفـة، كذلك تكون مصدرا للاتصال والتقدم الاجتمـاعي حيـث إن تطـور الفـرد اجتماعيـا و مـهنيا يعتمـد علـى مقـدار مناسب من المعلومات وعلـى نوعيـة محـددة منـها. إن أقطـار الوطـن العربـي، والمنظمـات العربيـة المختلفـة بحاجة إلى وضع خطط قصيرة ومتوسطة وبعيـدة المـدى تؤمـن تحقيـق تنسـيق وتعـاون بـين كافـة النشـاطات والمصادر في القطر الواحـد أو المنظمـة الواحـدة. وكذلـك على المستوى العربي الشامل، وبالتـالي تحقيـق الفـائدة لكافة المؤسسات والأقطار العربية.

إن أساليب وسبل التخطيط الحديث لشبكة ونظام المعلومات تتضمن تشخيص الاحتياجات والمصادر العامة والخاصـة، وتحديد الأهداف، المطلوب تحقيقها، ومـن ثـم تنظيـم وتقييـم الإمكانـات المتاحـة، أمـا مصـادر المعلومـات العامـة والخاصـة فمتعـددة، أهمـها مراكـز التوثيـق والمعلومـات في المجـالات المختلفـة والمراكـز الوثائقية والأرشيفات، والمجالات المختلفة التخصصات ومراكز البحوث والدراسات، التي تقوم باختيار وجمع المعلومات والمواد الثقافية والإعلامية المختلفـة بتنظيمـها وتحليلها وتهيئتها للاستخدام على المستويات القطريـة

(*) المركز الوطني للتوثيق – الرباط

والقومية بواسطة بنوك وشبكات المعلومات الخاصة.

وعلى أساس ما تقدم فإنه لا بد من إيجاد وعي بأهمية المعلومات وعلى كافة المستويات الاجتماعية والمهنية والقطرية والقومية. فبدون هذا الوعي لا يمكن تنفيذ شبكة أو نظام للمعلومات بالمستوى المطلوب.

من جانب آخر فإنه يطلب من العاملين في مراكز التوثيق والمعلومات والمتخصصين في هذا المجال أن يكونوا ملمين بدورهم الفعال في خطط التنمية القطرية والقومية، ومن الجانب الآخر فإنه يطلب من المخططين ورجال السياسة أن يكونوا على علم بأهمية هذا الدور المهم لأخصائي المعلومات والتوثيق وعلاقة أعمالهم بالأعمال والقطاعات الأخرى.

وبذلك يكون في الإمكان إيجاد الظرف والمناخ المناسبين لتستطيع مراكز التوثيق والمعلومات تقديم أقصى وأحسن ما عندها من الإسهام والمشاركة الفعالة في بنوك وشبكات المعلومات على مختلف المستويات والأصعدة.

ولا بد من الاعتراف هنا بأن المعلومات وتوثيقها لم تعط الاهتمام المطلوب والمناسب في العديد من الأقطار والمؤسسات العربية عند رسم خطط التنمية في الماضي.

وعلاجا لهذا النقص يجب تخطيط و تنظيم وتنسيق شبكات المعلومات القطرية والقومية إذا ما أريد لها أن تمارس وظائفها كعناصر فعالة في تطور الأمة الثقافي والتربوي والحضاري والعلمي وذلك بإلقاء نظرة على التجارب العالمية في هذا المجال.

وعلى المستوى العربي فإن شبكات المعلومات القطرية والعربية أصبحت محور ندوات ولقاءات تبشر بمستقبل أفضل في هذا المجال ولقد أدرك المعنيون في مجال المعلومات بأن هوية وشكل شبكة المعلومات القطرية والقومية تتألف من عدد من الأنظمة الفرعية والثانوية ذات تنسيق مركزي أو شبه مركزي سواء كان ذلك على مستوى المنظمات العربية العامة والمتخصصة القطاعية أو على المستوى القطري الذي يؤمن التنسيق الإقليمي أو القومي للمعلومات .

في الصفحات القادمة نسلط الأضواء على معايير وأسس قواعد بنوك المعطيات التي تعتمد على استخدام الحاسوب بمختلف أنواعه وأحجامه.

شفرة يونيكود

قبل أن نستطيع التوغل في الحل الذي يمكن أن تقدمه تقنية "يونيكود" يجب أن نفهم بعض المصطلحات الأساسية الخاصة بالموضوع. فأولا هنالك الحرف، وهو الوحدة الأساسية التي تستخدم في تبادل المعلومات. أما الصورة glyph فهي الشكل الذي يتخذه حرف معين عند استعراضه ضمن البرمجيات. وأما لوحة المحارف Character Set فهي مجموعة مرتبة من الأحرف، في حين أن الخط Font هو مجموعة مرتبة من الصور ذات النمط الواحد. ونشير إلى هذه المجموعة المرتبة بتشفير المحارف Character Encoding .

ومن المفهوم بوجه عام أن بعض التفاعل بين المعدات والبرمجيات ضروري لترجمة المحارف إلى صور تظهر على الشاشة. وتختلف عملية التفاعل هذه

اختلافا كبرا بين نظام وآخر، ونشير إليها بعملية الرسم rendering ، ويشترك في عملية الرسم العناصر التالية: نظام التشغيل، وتعريف اللغة وبيئتها، ولوحة المفاتيح وبرمجيات العرض، وبرمجيات معالجة الكلمات، وعمليات الإدخال والإخراج في المعدات.

لوحات المحارف و تشفيرها

في اللغات القائمة على الأحرف اللاتينية لا توجد أي مشاكل كبرى في عملية دعم البرمجيات للغات القائمة على هذه المحارف، حيث إنه قد تم تطوير لوحات المحارف لتكون شاملة تقريبا لمعظم الاختلافات بينها. فلوحة المحارفLatin-I مثلا، تتشابه حروفها تماما مع الصور الموجودة في اللغة مما يجعلها تخدم بشكل ممتاز كلوحة محارف تشفير للحروف.

ومن ناحية أخرى فإن عملية التحويل من لوحة المحارف إلى الخطوط يفشل بشكل كبير عند الانتقال إلى مقياس IS088596 وهو مقياس ا للاتينية– العربية. ففي هذه الحالة تشير الحروف إلى الحروف العربية دون الإشارة إلى الشكل الذي يتخذه الحرف خلال عملية الرسم عند تمثيله على الشاشة. ولذلك فإن الحرف الممثل في لوحة المحارف التابعة لمقياس ISO 88596 لن تمثل بشكل جيد أثناء عملية الرسم لأنها لا تحتوي على جميع الأشكال الضرورية لجميع المحارف العربية.

لحل هذه المشاكل وغيرها ابتدأ العمل عام 1988 بمعيار جديد للوحات المحارف أطلق عليه اسم "يونيكود". وهي شيفرة محارف جديدة، صممت لتشفير النصوص وتخزينها في ملفات حاسوبية. ويقوم تصميم معيار "يونيكود" على البساطة والتماسك الموجود في لوحات المحارف المستخدمة اليوم، وبالتحديد "آسكي ASCII" ، وLatin– I . إضافة إلى نسخة أكثر شمولا من لوحة محارف "آسكي " لكن "يونيكود" يوفر أكثر بكثير مما هو موجود في لوحة محارف "آسكي " حيث إنه يوفر للمستخدم القدرة على تشفير جميع الحروف المستخدمة في اللغات المكتوبة في العالم اليوم. ولاستيعاب ألوف الحروف المختلفة الموجودة في لغات العالم اليوم كان "يونيكود" يمثل الحروف ب 16 بت بدلا من مجموعة محارف "آسكي " بسعة 7 وبفضل هذا التوسع فإنه يمكن للوحة المحارف أن تستوعب أكثر من 65000 حرف. أكثر بكثير من ال 128 حرفا التي تتوافر في مجموعة محارف "آسكي " ورغم هذا الحجم الكبير من الحروف فإن عدد الحروف في لغات العالم المكتوبة يزيد خمس مرات عن عدد محارف "يونيكود". ومن هنا فإن تصميم "يونيكود" يعامل هذه المساحة بحرص. حيث لا تضيف الهيئة المسؤولة عن "يونيكود" المحارف بشكل جزافي، تحسبا للحاجة إلى حروف إضافية في المستقبل. ولذلك، فإننا نجد مثلا أن لوحة محارف "يونيكود" تحتوي على الحروف العربية الأساسية فقط، مع أخذ الحروف الأوردية والفارسية بعين الاعتبار. ومن الأمور المهمة الواجب تذكرها دوما هي أن "يونيكود" عبارة عن معيار وليس تقنية. ووظيفة هذا المعيار هي تحديد وتعريف الحروف ومواقعها ضمن لوحة المحارف

تمهيدا لتعامل البرمجيات أو الأجهزة معها. وبالتالي فإن الفروق الصغيرة في أشكال الحروف هي أمور يجب التعامل معها كمسائل متعلقة بالخطوط، وليس هناك حاجة لابتكار لوحة محارف جديدة للتعامل معها.

أضف إلى ذلك أنه رغم احتواء لوحة محارف "يونيكود" على حروف لغات معقدة (تصنف العربية أنها واحدة منها) فإن أيا من نظم التشغيل المستخدمة اليوم لا يمكنه أن يقوم بشكل آلي بعملية رسم rendering حروف "يونيكود" جميعها، ذلك أن "يونيكود" كما أشرنا هو مجرد وسيلة لتشفير المحارف وليس مكتبة للتطوير أو تقنية. ولذلك فإن "يونيكود" الحالية لا توفر ضمن لوحة محارفها أي دعم آخر للغة العربية مثل أن تأخذ بعين الاعتبار التغير في أشكال الحروف عند تغير موضعها في الكلمة. أو حركة المؤشر على الشاشة. وهي تترك ذلك للمطورين ليقوموا بتطوير الآليات اللازمة لربط الأحرف مع بعضها وتغير شكلها. وكل ما تفعله "يونيكود" هو أنها تقوم بتقسيم نظام الكتابة إلى عناصر شيفرة .code elements وهي في العادة أكثر الحروف استخداما في اللغة وللزيادة في الاقتصاد فإنه في حالة الاختلاف حول تعريف العنصر النصي تقوم "يونيكود" بتعريف عنصر الشيفرة اعتمادا على العنصر النصي الأساسي المستخدم. وبعد هذه الخطوة تقوم بتخصيص 16 لكل عنصر شيفرة. ويدعى كل من هذه الأرقام نقطة شيفرة code point تبتدئ جميعها ب+ U ابتداء بمحارف آسكي ومرورا بالإغريقية والعربية والعبرية. وثمة حقيقة أخرى يجب إدراكها وهي أن نظم التشغيل الشائعة اليوم يمكنها أن تعرض معظم اللغات البسيطة (اللغات اللاتينية على وجه العموم) إلا أن قليلا منها يمكنه التعامل مع اللغات المعقدة وما تتطلبه من حركة المؤشر على الشاشة، واختيار النص، والنسخ، واللصق، والتقنيات الرياضية المتعلقة بالمؤشرات في شيفرة البرمجة. وهنا تظهر الحاجة إلى استخدام أدوات برمجية ومكتبات لتمكين البرامج من التعامل مع هذه المسائل عند معالجة لغات معقدة.

وتتوافر مثل هذه الأدوات في الوقت الحالي لدى عدد من الشركات مثل "غاما برودكشنز" "ستاربلاس غلوب". و "زنك ".

وقد يعتبر البعض هذا الأمر عجزا في "يونيكود" كمعيار، إلا أن وجهة القائمين على الأمر هي أن "يونيكود" يوفر البنية الأساسية اللازمة لعملية توطين المنتجات وما على المطورين إلا توفير الآليات والبرمجيات اللازمة التي تقوم بعملية رسم هذه المحارف والتعامل معها. وهو أمر أسهل كثيرا من جعل نظام التشغيل بأكمله يقوم بعملية رسم محارف "يونيكود " كاملة.

جدير بالذكر أن واضعي معيار "يونيكود" يسعون منذ سنين لجعله المعيار العالمي الأساسي للمحارف ووضعوا خطة أولية لمعيار جديد اسمه ISO 10646 إلا أن اتفاقية حديثة بين اتحاد "يونيكود" واللجنة العاملة على تعريف المعيار ISO 10646 نصت على دمج هذين المعيارين بحيث يصبح معيار "يونيكود" جزءا من لوحة

المحارف الخاصة من ISO 10646 وللتكيف مع هذا المعيار الجديد سيحتوي الطرح الجديد من يونيكود على 4600 حرف جديد منها 600 مخصصة للغات الصينية واليابانية والكورية CJK . إضافة لأكثر من 1000 حرف جديد للغات أخرى.

" جافا" تحتضن العربية

جافا هي أول لغة برمجة تجعل من معيار "يونيكود" نمط البيانات الأساسي لها. وإن كانت لا تتعامل مع رسم الحروف أو الآليات الأخرى التي تعد من مسؤوليات نظام التشغيل الذي تعمل من خلاله، مع ذلك، فان العديد من البرامج المكتوبة بلغة "جافا" لا تزال بحاجة إلى التعامل مع النصوص المكتوبة باستخدام لوحات محارف أخرى. وبالتالي فإن "جافا" تقوم بتوفير مجموعة من الأصناف Classes لتحويل العديد من شيفرات الحروف القياسية من و إلى "يونيكود". وبذلك تقوم برامج "جافا" التي تحتاج للتعامل مع نصوص بغير تنسيق "يونيكود" بتحويل هذه البيانات إلى "يونيكود"، ومن ثم معالجتها بهذه الصفة، ثم تحويل النتائج مرة أخرى إلى شيفرة الحروف الأولى.

ويقوم صنفان بتوفير الواجهة التي تستخدمها محولات شيفرات المحارف. وهذان الصنفان هما Java.io.Byte to charconverter ، وهو عبارة عن قاعدة للأصناف التي تقوم بتحويل الحروف المشفرة بشيفرات خارجية إلى "يونيكود". والصنف Java.io.char to Byte converter وهو قاعدة للأصناف التي تقوم بتحويل محارف "يونيكود" الى محارف خارجية. وبما أنه من النادر أن يكون هنالك تطابق كامل بين تشفير المحارف فإنه عندما يواجه أحد المحولات حرفا في الشيفرة المصدرية لا يوجد له مقابل في الشيفرة المستهدفة فإنه يقوم إما بإظهار رسالة خطأ، أو يقوم باستخدام حروف بديلة يتم تعريفها مسبقا عوضا عن الحروف التي لا يتم تمييزها.

إلا أن الوضع ليس بهذه السهولة بالنسبة للغة العربية. فهذه الأصناف لا تشمل اللغة العربية، أضف إلى ذلك أنه في الوقت الذي تدعي فيه شركة "صن " أن جافا لغة كونية يمكن تشغيل برامجها في أي بيئة تشغيل إلا أنها لا توفر إلى الآن أمرين هامين: الأول، هو آلية لرسم المحارف العربية ودمجها، إضافة إلى قدرات مثل الكتابة من اليمين إلى اليسار، أضف إلى ذلك ما لاحظناه من أن آلة "جافا" الافتراضية المصممة لنظام "ماكنتوش " مثلا يمكنها أن تميز جميع الخطوط المستخدمة في النظام، في حين أن هذا القول لا ينطبق على آلة "جافا" الافتراضية المصممة لوينـدوز 95 ولتفادي هذه المشكلة يمكن كتابة أصناف قادرة على عرض الخطوط والتعامل معها، إلا أن نقطة الضعف في هذا الأسلوب هي أن ذلك سيؤدي إلى وضع أعباء إضافية على النظام بسبب الوقت اللازم لترجمة الشيفرة البايتية للغة "جافا" إضافة لكون برامج وبريمجات جافا تقوم بشكل منظم بعملية مسح للذاكرة لتنظيفها مما يؤدي إلى زيادة بط ء النظام.

ولعل ما ورد أعلاه يصح على JIKL.1 لكن

أوضاع "جافا" واللغة العربية ستكون أفضل كثيرا في المستقبل القريب. حيث تقول مصادر شركة "صن " إن توجهها حاليا هو نحو إنتاج الأصناف اللازمة للتعامل مع اللغتين العربية والعبرية. فضلا عن مجموعة من الحلول التي توفرها شركات أخرى للتعامل مع "جافا" باللغات المعقدة.

ففي أيار/ مايو الماضي أعلنت شركة صن أنها قامت بالحصول على ترخيص باستخدام تقنيات طورتها شركة "تاليجينت"، و التي تملكها "أ. ب. م." وتشمل هذه التقنيات تقنية التصميم والكتابة ثنائية الاتجاه للتعامل مع النصوص العالمية إضافة إلى القدرات الرسومية اللازمة للتعامل مع تغير الأشكال الهندسية. وستكون هذه التقنية جزءا من واجهة تطوير البرامج java 2 DAPI، وحزمة تطوير جافاJIK وتقوم هذه التقنيات بتقوية "قدرات حزمة AWT الخاصة بلغة "جافا" والتي هي جزء مهم من مجموعة أصناف "جافا" الأساسية. وفي هذه الحالة فإن التعامل مع اللغات المعقدة كاللغة العربية سيكون على مستوى النظام نفسه.

ومن الحلول الأخرى التي طرحت في نهاية أيلول الماضي تقنية JEF من شركة "بيت ستريم " والمصممة بالكامل بلغة "جافا" وبفضل هذه التقنية يمكن لبرامج "جافا" أن تعرض أيا من الخطوط المستخدمة في اللغات العالمية أو في أي تشفير للمحارف (بما في ذلك "يونيكود") دون الحاجة لاستخدام أي أصناف أو برامج إضافية. وستسهل هذه التقنية عمل المطورين الذين يسعون لتطوير برامج جافا وواجهات استخدامه. ويعتمدJET على حزمة "ترودوك " للتطوير، وهي عبارة عن شيفرة مصدرية تحل مشاكل الخطوط في الوثائق الإلكترونية القابلة للنقل. وتتكون حزمة تطوير "ترودوك " من عنصرين أساسيين؛ الأول هو مسجل أشكال الحروف CSR ومشغل أشكال الحروف." CS ويقوم CSRبتسجيل أشكال الحروف (الصور) ثم يضغطها في هيكلية كبيرة الضغط. ويدعى الملف الناتج بمصدر الخطوط النقال P f R.ويقوم CSP باستعادة البيانات منPIR وإعادة خلق أشكال الحروف بسرعة كبيرة.

وبالطبع فإن هذه الحلول جميعها تبشر بقرب التعريب الكامل لجافا. إلا أن السؤال الكبير الذي يطرح نفسه هنا هو: ماذا عن الأداء؟ جميع الحلول التي قرأنا عنها تضع عبئا كبيرا على ذاكرة الأجهزة وتقلل من سرعة الأداء. وقد لا يكون هذا الموضوع خطرا في صناعة يتواصل فيها انخفاض أسعار الأجهزة مع ازدياد قوة المعالجة فيها. لكنه بالتأكيد موضوع مهم للمؤسسات التي لا ترغب بترقية أجهزتها الحالية، أو تلك التي تعتمد في عملها على معلومات مخزنة على أجهزة قديمة. إن الحل لذلك يكمن في تسريع عمليات المعالجة في لغة "جافا" وتحسين هيكلتها بحيث يتم تقليل الوقت اللازم لترجمة الشيفرة البايتية. و"صن " وخلفاؤها ليسوا غافلين عن هذه المسألة. ومن هنا كانت فكرة إنشاء مركز لتعديل وتصدير لغة جافا (انظر أخبار الشبكات العدد الماضي)، والذي وضع هذه الأهداف

جميعها نصب عينيه. أضف إلى هذا عمل الشركة على تطوير نظام فوري لتوليف البيانات JIT مما سيدعم سرعة التنفيذ في بيئة جافا.

وفي حالة حدوث تحسينات جذرية على هذه الأصعدة في لغة "جافا" فإنها ستكون، على الأقل بالنسبة للعالم العربي، الحصان الأسود الذي سيتمكن بواسطته المطورون العرب الدخول بتقنياتهم إلى الأسواق العالمية، حيث يمكن لبرامج "جافا" أن تعمل ضمن أي بيئة تشغيل بما في ذلك البيئات الأجنبية. ولا شك في أن لدينا الكثير مما يمكن أن نساهم به وخصوصا في مجال اللغويات وتقنيات الذكاء الاصطناعي. ولا شك في أن استخدام "جافا" سيحل مشكلة كبيرة للمؤسسات العربية التي تجد نفسها مضطرة للتعامل بعدة لغات خارج العالم العربي وداخله. وفي حالة توافر"جافا" وتطبيقاتها بالعربية فإن الحمل الناتج عن دعم عدة تطبيقات وبلغات مختلفة داخل المؤسسات العربية سيكون أقل كثيرا.

لهذه الأسباب فإن العالم يبشر بأن يكون من الأسواق التي تتحقق فيها تصورات مصممي "جافا" والبيئة التي تتجلى فيها هذه اللغة وجميع الأماني التي صممت لتحقيقها، علما أن " صن" تدعي بوجود 200 مطور في المنطقة.

المعايير و التراكيبNORMES ET FORMATS

تعتبر الموصفات وأشكال التبادل بمثابة الأداة الأساسية في مجال الوصف الببليوغرافي. ولذلك تعد التسجيلات الوسيلة الوحيدة للتعرف المادي على الوثائق، والولوج الى المعلومات التي تحتويها.

إن ضرورة الولوج المريح على الصعيد العالمي إلى المعلومات الببليوغرافية وتنمية التنسيق بين الوحدات المشتغلة بالمعلومات، قد نادت باكرا إلى المعايير. وتم إنجازها أولا على مستوى وطني أو لغوي، ثم وضعت واصفات وقواعد للفهرسة، واستمر تحيينها بواسطة لجان مهنية في بلدان عديدة. وبالنظر لانتشار اللغة الإنجليزية، فإن قواعد الفهرسة الإنكلو أمريكية (AACR) قد حظيت باهتمام كبير وترجمت ثم استعملت باللغتين الفرنسية والإسبانية.

وضاعف استعمال المعلوميات الحاجة إلى المعايير إلى درجة أنه لكي تتم قراءة التسجيلات الببليوغرافية بواسطة الآلة، فإنه أصبح من اللازم، ليس فقط تطبيق القواعد بحزم على الصعيد الذهني، بل تقديمها بدقة شديدة وموحدة باستخدام حروف وأبعاد محددة، ويعطي تطبيق تلك القواعد إلى جانب التجسيد المادي للأبعاد شكلا محددا.

هكذا بذلت جهود لوضع مواصفات عالية تحت رعاية الفيدرالية الدولية لجمعيات المكتبات والمكتبيين(IFLA) ، قادت إلى بلورة وصف ببليوغرافي دولي موحد (ISBD) ، لاستعماله أولا بالنسبة للكتب (ISBD-M)، ثم بعد ذلك بالنسبة للدوريات (ISBD-S) ، وسهرت نفس المؤسسة بعد ذلك على بلورة معايير خاصة بمجالات أخرى منها السمع-بصريات (ISBD-NBM) والخرائط والتصاميم (ISBD-CM) والكتابة الموسيقية (ISBD-Music)

والكتب القديمة (ISBD-A)، وهناك معايير أخرى قيد الدرس خاصة بالملفات المقروءة بواسطة الآلة (MRP)مثلا، أو بالجزئيات المكونة لها أو بالاستشهادات البيبليوغرافية. (CP).

وهذه المعايير هي عبارة عن مجموعة من القواعد المحددة لتقديم المعطيات البيبليوغرافية وعلامات للتنقيط تفيد في التعرف عليها. ويسمح تطبيق تلك المعايير بإجراء "مراقبة بيبليوغرافية عالمية.(CBU) وهي عبارة عن وصف موحد من طرف جميع المؤسسات المكلفة بإنتاج البيبلوغرافيات الوطنية للوثائق المنشورة داخل كل بلد، حتى يتمكن الجميع من استعمالها، مسهلة بذلك الوصول إلى تطبيق لبرنامج UAP- (Universal Availability of Publication) الدخول العالمي للوثائق المنشورة.

وبطريقة موازية" قام فريق عمل تابع (international Council of Scientific Unions) ببلورة دليل مراجع خاص بالوصف البيبليوغرافي المقروء بواسطة الآلة، الذي قامت "اليونسكو" بنشره، وهو يحدد نماذج معينة من التسجيلات البيبليوغرافية ومن تركيبات التبادل، وكل تسجيلة متفق عليها يمكن تطبيقها على ثلاث مستويات: على الكتب، جماعية وتحليلية (انظر فيما بعد)، وهذا يجعلها قابلة للاستعمال من كل مراكز التوثيق والمصالح الفرعية التي تقوم بتسجيل وثائق أصغر من البيبليوغرافيات الوطنية (مثل فصل من كتاب، مقالة في دورية، مداخلة واحدة من بين أعمال مؤتمر، الخ..)

وتأخذ عملية معيرة الأشكال مجراها ونموذج MARC الأمريكية في مركز هذه الجهود، حيث اعتبر منطلقا لكثير من أشكال التبادل البيبليوغرافية لدى كثير من المكتبات، وذلك بسبب العدد الهائل للمراجع المتوافرة في هذا النموذج. MARC-BNB, (Intermarc Marcal) وقد اختارت (IFLA) سنة Unimarc 1977، بصفتها شكلا لمسك وعرض المعطيات، وكذلك بصفتها شكلا لتبادل المعطيات. وتتوجه Unimarc أساسا للمكتبات. ومن ناحية ثانية فإن أشكالا أخرى عديدة قد جرت بلورتها لفائدة أنماط توثيقية ولمصالح فرعية. وبالنظر لكون الأنساق المعلوماتية التي وضعت هذه المعايير من أجلها، تستجيب لعراقيل متعددة، فإن عملية وضع معايير خاصة بالأشكال لا تميل لبلورة شكل موحد، وإنما تحاول أن تجعل الأشكال العديدة متلائمة مع بعضها، بحيث يصبح الانتقال من إحداها إلى الأخرى عملية سهلة وآلية. وهذا ممكن بمجرد أن يكون محتوى وبنية الحقول منسجما، ويكون بالوسع التعرف على كل منطقة بطريقة واضحة. (بواسطة علامات متفق عليها تؤدي إلى معرفة محتوى الحقل، وبداية ونهاية الحقول والحقول الفرعية). وبهذا الصدد كلفت IFLA واليونسكو فريق عمل بدمج الأشكال الرئيسية للتبادل، فيما سمي بالتركيبة الشكل للتواصل (CCF)، وهكذا، قامت مكتبة (داغ همرشولد مثلا) بتطوير نسق جديد خاص بها انطلاقا من هذا الشكل المشترك(UNBIS)، وهو شكل خاص لمعالجة وثائق وأدبيات الأمم المتحدة.

تحتوي التسجيلة عموما على شفرة (كود) تحيل على الوحدة الإعلامية التي انتجتها، وعلى سنة الإنتاج ورقم تسلسلي . ويعطي هذا الرقم في الغالب وفق ترتيب متسلسل تاريخيا، وربما داخل حزمة تحتوي كل منها على مجموعة من الأرقام يسلمها المنسق الدولي لكل وحدة مشاركة في الشبكة.

ويمكن أن تحتوي التسجيلة على علامات تسمح بالتعرف هل هذه التسجيلة جديدة أم أنها تعوض تسجيلة سابقة عليها، وإن كانت ضمن حزمة من التسجيلات التي تصف نفس الوثيقة، فإن هذه العلامات تتيح التعرف على التسجيلات المتصلة بها (ترجمة أو طبعة جديدة إلخ.).

وتتقدم هذه الإفادات غالبا في بداية التسجيلة، ولا تظهر عند طبع الملفات والفهارس لأنها تكون مفيدة على الخصوص بالنسبة لعمليات المراقبة والمعالجة.

ويظهر الرقم النهائي دائما على التسجيلة، لأنه يكون بمثابة الرابط الرئيسي بينها وبين المنتوجات والخدمات التوثيقية الأخرى (الكشاف index ، طلبات الحصول على نسخ، إلخ...). وتمكن المؤشرات الدالة على نوع الوثيقة من معرفة النوع الوثائقي الذي يتم التعامل معه، وحيث إن كل نوع من هذه الأنواع الوثائقية يتطلب معالجة خاصة به، فإن العلامات تمكن من اختيار سلامة التسجيلة ومطابقتها للمعايير.

فنعرف مثلا، أن تسجيلة متعلقة بأطروحة ما، يجب أن تحتوي على حقل العنوان، وحقل المؤلف، والتاريخ، ولغة الأطروحة والتوريق والجامعة، ثم وجود الأطروحة أو عدمه، وتجري مراقبة توافر هذه الحقول في جميع التسجيلات الخاصة بالأطروحات. كما يسمح لنا هذا المؤشر بالقيام بالمراقبة الإحصائية. ولا تقوم أنظمة الترقيم عادة بحصر كل أنواع الوثائق، لكنها ترمز لبعضها (ستة أنواع بالنسبة لنظام ICSU-AS/Unisist مثلاً)، باعتبارها وثيقة يمكنها أن تندرج في هذه الأصناف الستة. والشيء المهم بالذات هو أن المعالجة الخاصة بكل نوع ينبغي أن تطبق على الوثيقة.

وتستخدم بعض الأنساق مؤشرات بيبليوغرافية تماما كما هو الشأن بالنسبة للمؤشرات الخاصة بأنواع الوثائق. وهي تفيد في الإشارة إلى توافر عناصر قد تكون لها فائدة خاصة في الوثيقة المعالجة، دون أن يكون هناك تناول مفصول لهذه العناصر. مثل الخرائط، والمعطيات الرقمية، والملخصات، إلخ) وهكذا يتم وضع هذه العناصر في حقل خاص بها، وهو حقل الملاحظات (كعدد المراجع، والفترة الزمنية التي تغطيها تلك المراجع مثلا). وتوضع المؤشرات غالبا في مطلع التسجيلة.

وتذكر المؤشرات ذات الطابع البيبليوغرافي بالتدقيق الجزء الذي قمنا بوصفه من الوثيقة المعالجة، لان الكتاب يتكون بالطبع من عدة فصول يشكل كل واحد منها وحدة تامة، وينبغي بالتأكيد معالجة الكتاب برمته لأنه الوعاء المادي الوحيد للمعلومات المقصودة، لكن قد يكون كذلك من المجدي وصف بعض الفصول على الأقل إن لم نصف مجموعها، بصفتها

وحدات ذهنية نسميها "وحدات توثيقية"، وينطبق نفس الشيء على المنشورات المتسلسلة والمقالات التي تتضمنها، وعلى أعمال المؤتمرات، ومختلف المداخلات إلخ... ومن جهة ثانية، فإن بعض الوثائق رغم كونها مفصولة ماديا عن بعضها، إلا أنها تشكل كلا مترابطاً، ماديا وذهنيا (مثل الأجزاء المتعلقة بموسوعة، أو المستلاتfasucils الخاصة بتقرير تقني، أو دوريات، الخ.) لذلك فنحن مضطرون إلى التمييز بين ثلاثة مستويات بيبليوغرافية، وهي كالتالي، حسب ترتيب تنازلي لأهميتها:

– المستوى التحليلي، عندما تتناول التسجيلة جانبا من وثيقة خاصة كخريطة من كتاب "الأطلس" أو مقالة من دورية، أو فصلا من كتاب.

– المستوى المونوغرافي عندما تتعرض التسجيلة لوثيقة منفردة بصفة إجمالية– كتاب أو أطلس خرائط، أو معايير أو براءة اختراع إلخ...

– المستوى الجماعي عندما تهتم التسجيلة بمجموعة معينة من الوثائق (مثل كتاب من عدة أجزاء، أو دورية مسلسلة الصدور إلخ)

وبطبيعة الحال، فإن المستوى التحليلي لا يقع اللجوء إليه إلا بصفة استثنائية، عندما لا نتوفر على الوثيقة التي استخرجناها من الوحدة التوثيقية المطلوب معالجتها (نتوفر على فصل ولا نتوفر على الكتاب الأصلي، نتوفر على مقالة ولا نتوفر على الدورية التي سبق نشره بها). وعادة ينبغي أن تكون التسجيلة ذات المستوى التحليلي مصحوبة بتسجيلة أخرى من المستوى المونوغرافي أو الجماعي ومع ذلك فهي بمفردها تتضمن المؤشرات أي تسمح لنا بالتعرف على الوثيقة المصدر، ويكفينا الإطلاع عليها للحصول على المعطيات الكافية.

وتمس تلك العلامات عمليا كل الحقول التي تتضمنها التسجيلات، وهي: الحقل الخاص بالمؤلف، أو مسؤولية ما جاء في الوثيقة، وقد يكون فردا بعينه (شخص مادي) أو هيأة (شخص معنوي أو المؤلف الجماعي) كما قد يكون غير محدد أو مجهول. وفي هذا الحقل وحده تتعدد المعايير المتعلقة بترتيب وكتابة واختيار جميع الأنواع الفرعية من كل واحد من الأصناف الثلاثة المذكورة، بل إن عملية المعيرة أصبحت تشمل حتى أدق تفاصيل العلاقة بين نوع الوثيقة ونوع المؤلف والجوانب القانونية المتصلة بتنظيم العلاقة بين المؤلفين والمؤسسات التي تتحمل جزءا من المسؤولية.

ثم حقل العنوان، بعلاقة مع نوع الوثيقة، ومع مختلف المستويات أو الطول الذي يكون عليه العنوان، وهل هو رئيسي أو فرعي، أو مترجم أو موازي، أو جزء من عنوان مشتق من أعمال ندوة أو مؤتمر، أو عنوان دورية متسلسلة، أو عنوان جديد لدورية كانت تصدر بعنوان آخر قديم.

ثم حقل الطبعة، وهو يتضمن كل المعلومات المتعلقة بالطبعة وما يلحق بها من توضيحات– ثم حقل العنوان المادي Adresse ، أو البيبليوغرافي، الذي يتضمن معلومات عن إنتاج الوثيقة من الناحية المادية بحيث يشتمل على مكان النشر، واسم الناشر، وتاريخ

النشر أو الطبع، ومكانه، واسم الطابع إن لم يذكر اسم الناشر ومكانه. ثم حقل التوريقCollation الذي يشتمل على وصف للمكونات المادية للوثيقة، مثل أقسام الكتاب (أجزاء، مجلدات، مستلات)، وحجم الوثيقة بالسنتمتر، وعدد الصفحات، والرسوم والصور والجداول إن وجدت، والمراجع وعددها وطبيعتها ونوعها وهل هناك كشافات أم لا- ثم حقل خاص عن السلسلات إن كانت الوثيقة تصدر ضمن سلسلة معينة (ISBD-M) ثم حقل الملاحظات الذي يثبت فيه ما يتصل بلغة الوثيقة أو لغاتها، ولغات الملخصات الواردة ضمنها. وما إلى ذلك من الحقول الرئيسية مثل ISBN و ISSN.

الوصف البيبليوغرافي للوثائق السمعية البصرية

يعرف وصف هذه الوثائق نوعا من الصعوبة، بحيث يفرض تقديم معلومات متعلقة بالوثيقة نفسها (المحتوى)، أو بالعناصر الخاصة بخزنها أو بالوثائق المرفقة بها، إضافة إلى ضرورة خلق عناوين (إبداعها) بالنسبة للصور مثلا كما تأتى الصعوبة كذلك أحيانا من الحيرة بين أهمية وصف الوثيقة معزولة (صورة) أو تقديم وصف عن الاستطلاع بكامله، أي عن مجموعة الصور التي تؤخذ لموضوع واحد، وتأتي أيضا من صعوبة حصر الوحدة الجزئية، مثلا هل هي التصميم، أم مجموعة الصور التي تؤخذ لوصف تصميم واحد، أم هي الفلم بكامله؟

كما أن المؤلف سيتم نسيانه في الغالب ما عدا بالنسبة للصور حيث إن معظم الوثائق السمعية البصرية هي وثائق يتم إنتاجها بصفة مشتركة بين عدد كبير من الأشخاص ويعتبر المنتج بمثابة مؤلف رئيسي. وبالنسبة للموسيقى الكلاسيكية يعطى اسم الموسيقار بينما في الأغاني يسجل اسم المغني. ثم يضاف إلى هذا اختلاف طريقة الطبع من وثيقة إلى أخرى، وغالبا ما يجري طبع أو إخراج نفس المحتوى بطرق متباينة إلخ.

ومن هنا فان المعمول به هو تسجيل اسم وعنوان الناشر أو المنتج، ويذكر اسم الموزع، ومن ناحية أخرى تحظى الخاصيات الفيزيائية للوعاء الحامل لمحتويات الوثيقة باهتمام كبير، لأنه يساعد على تحديد طريقة الاستعمال، مثل نوعية الجهاز الذي سيسمح بالاطلاع على المحتوى الكافي في الوعاء، وتوضع هذه المعطيات في حقل التوريق، ويضاف بالنسبة للأفلام عدد الجرارات (Bobines) أو البكرات، والأشرطة: والسفطات (cassetes)، والمدة التي يستغرقها الفلم، والطول بالمتر، والحجم أي العرض بالمليمتر أو البوصة (pouce) ونوعية الصور، والصوت، والتلوين، ومساحة العرض، والوثائق المكتوبة التي تصاحب الفلم (Script)، و تحتوي الملاحظات على معلومات أخرى مفيدة مثل مشاكل إنتاج الفلم والظروف المحيطة.

الهيكل المادي لتركيبة التراسل المشتركة

تشكل تركيبة التراسل المشتركة تنفيذا محددا للمواصفة الدولية ايزو (المواصفة العربية رقم(668). تتكون كل تسجيلة من تسجيلات ت ت م من أربعة أجزاء رئيسية:

- مرقومة التسجيلة.
- الدليل
- حقول البيانات
- فاصل التسجيلات

عند بناء مجلد مادي (شريط أو قرص ممغنط) من التسجيلات البيبلوغرافية لا بد من إيلاء هيكل هذا المجلد العناية اللازمة. وسنغطي أدناه كلا من هذه الموضوعات.

مرقومة التسجيلة

تبدأ كل تسجيلة من تسجيلات ت ت م بمرقومة ثابة الطول من 24 محرفا تكون محتوياتها على النحو التالي:

موضع المحرف	المحتـــوى
0 إلى 4	طول التسجيلة. يتضمن طول التسجيلة المرقومة والدليل وحقول البيانات وفاصل التسجيلة (إن استعمال 5محارف لطول التسجيلة يسمح بتسجيلات يصل طولها إلى 99999 محرفا).
5	وضع التسجيلة. يشير إلى وضع التسجيلة باستعمال رمز مأخوذ من قائمة رموز وضع التسجيلة المبينة في القسم 1.4. خ (خال) espace . موضع المحرف غير مستخدم.
7	المستوى البيبليوغافي للمادة الهدف باستعمال رمز مأخوذ من قائمة رموز المستوى البلوغرافي المبينة في القسم 2.4.
8	خ (خال). موضع المحرف هذا غير مستخدم.
9	خ (خال). موضع المحرف هذا غير مستخدم.
10	"2" طول المؤشر. يبين هذا عدد محارف المؤشر في كل حقل.
11	"2" طول محدد الحقل الفرعي. يبين هذا عدد المحارف المستعملة كمحدد للحقل الفرعي.
12 إلى 16	العنوان الأساس للبيانات. الموقع ضمن التسجيلة الذي يبدأ عنده حقل البيانات الأول بالنسبة للمحرف الأول من التسجيلة الذي هو موضع المحرف المسمى"0" (صفر).
17 إلى 19	خ (خال). مواضع المحارف هذه غير مستخدمة.
20	4 طول "طول حقل البيانات " في الدليل. (يسمح استخدام 4 محارف بحقول بيانات يصل طولها الى 9999 –محرفا).
21	5 طول "موضع محرف البدء" في الدليل.
22	2 طول القسم المعرف للتنفيذ من كل مدخل في الدليل. في المحرفين المستخدمين يكون الأول لمحدد الجزء والثانى لمحدد الورود، وكلاهما موضحان أدناه.
23	خ (خال) . موضع المحرف هذا غير مستخدم

الدليل

الدليل هو جدول يحتوي على عدد متغير من مداخل من أربعة عشر محرفا، تنتهي بمحرف فاصل حقل. ويناظر كل مدخل في الدليل ورود حقل بيانات في التسجيلة. وينقسم إلى خمسة أجزاء:

المميز

- طول حقل البيانات
- موضع محرف البدء
- محدد الجزء
- محدد الورود

المميزTAG

رمز من ثلاثة محارف يحدد اسم حقل البيانات المناظر لمدخل الدليل. إن حقول البيانات مشروحة في القسم 2.3 يسبق كلا منها محددها.

طول حقل البيانات ((Length of datafield)

رقم من أربع خانات يبين عدد المحارف التي يحتلها الحقل، بما في ذلك المؤشرات وفاصل الحقل لكن ليس رمز فاصل التسجيلة إذا كان حقل البيانات هو الحقل الأخير في التسجيلة.

طول محرف البدء
(Starting character position)

عدد من خمس خانات يبين موضع المحرف الأول من حقل البيانات بالنسبة للعنوان الأساس للبيانات، أي المحرف الأول من أول حقل من حقول البيانات.

محدد الجزء ((Segment identifier)

محرف واحد (يتم اختياره من0–9 و أ– ي)
يسمى الحقل بأنه عضو من جزء معين. لمناقشة تجزئة التسجيلات.

محدد الورود ((Occurrence identifier)

محرف واحد (يتم اختياره من 0–9 و أ– ي)
يميز الورودات المعددة للحقول ضمن نفس جزء التسجيلة التي تحمل نفس المميز.

ويكون المدخل الواحد في الدليل منظما على النحو التالي:

محدد الورود	محدد الجزء	موضع محرف البدء	طول الحقل	المميز
محرف واحد	محرف واحد	5 محارف	5 محارف	محرفان

حقول البيانات

يتكون كل حقل للبيانات مما يلي:

- مؤشرات
- حقل فرعي واحد او اكثريكون كل منها مسبوقا بمحدد حقل فرعي.
- فاصل حقل.

ا لمؤشرات (Indicators)

بايتان (محرفان) محجوزان للاستخدام كما هو معرف لكل حقل للبيانات وقد يوفران مزيدا من المعلومات عن محتوى حقل البيانات، أو حول العمل المطلوب في عمليات معينة لمعالجة البيانات.

الحقول الفرعية

فاصل حقل	حقل فرعي	محدد الحقل الفرعي	المؤشرات
محرف واحد	متغير	محرفان	محرفان

أما حقل البيانات ذو الحقلين الفرعيين فيكون منظما على النحو التالي:

فاصل حقل	الحقل الفرعي	محدد الحقل	الحقل الفرعي	محدد الحقل	مؤشرات
	الثاني	الفرعي الثاني	الأول	الفرعي الأول	
محرف واحد	متغير	محرفان	متغير	محرفان	محرفان

فاصل الحقل

فاصل الحقل (الحرف 13/1 في مواصفة ايزو رقم 646) هو المحرف الأخير من التسجيلة، ويلي فاصل الحقل لحقل البيانات الأخير في التسجيلة.

يتكون الحقل الفرعي من محدد حقل فرعي متبوعا بسلسلة بيانات تنتهي إما بمحدد حقل فرعي آخر أو فاصل حقل. يتكون محدد الحقل الفرعي من علم محدد حقل فرعي (المحرف15/1 من مواصفة ايزو رقم646 المشار إليها هنا بالمحرف @) متبوعا بمحرف واحد آخر.

فاصل الحقل

فاصل الحقل(المحرف14/1 في مواصفة ايزو646) يشكل المحرف الأخير من كل حقل بيانات. يكون حقل البيانات ذو الحقل الفرعي الواحد منظما على النحو التالي:

هيكل الوسط المغنط

إن على أطراف اتفاقية التبادل الاتفاق على المقاس وكثافة التسجيل إلخ للوسط المادي الذي سيم تبادله. ولا بد من الرجوع إلى مواصفة ايزو رقم1001 من اجل الاسترشاد حول طريقة وضع المرقومة للشريط وهيكل الملف.