البيانات الوصفية (الميتاداتا) للموارد الرقمية بالمستودعات المؤسسية للرسائل والأطروحات الإلكترونية.

إعداد
مصطفى محمد عبد الحميد مجاهد
باحث دكتوراه
مدير عام بالمكتبة المركزية
جامعة القاهرة سابقا

الملخص:

يتطلب الاحتفاظ بمكتبة من العناصر الرقمية الضرورية الحفاظ على البيانات الوصفية حول تلك الموارد، وتُعد البيانات الوصفية ضرورية للغاية للإدارة الناجحة لأي كيان رقمي، وهي أكثر شمولاً وتختلف عن البيانات الوصفية المستخدمة لإدارة مجموعات الأعمال المطبوعة والمواد المادية الأخرى، وقد قامت الكثير من المؤسسات بإنشاء معايير ميتاداتا أو مخططات ميتاداتا بغرض تنظيم وإدارة مجموعاتها ومواردها الرقمية من الرسائل والاطروحات الرقمية، قابلة للتشغيل البيني، ومتواءمة مع إحتياجاتها، والخصائص المميزة لمواردها الرقمية من الرسائل والاطروحات الرقمية، الغرض من هذه الورقة البحثية هو التحقيق في معايير البيانات الوصفية للرسائل والاطروحات الرقمية المتاحة في جميع أنحاء العالم ودراستها بالتفصيل، تقدم الورقة التحليل التفصيلي لهذه المعايير، معتمدة على المنهج الوصفي التحليلي قي جمع المادة العلمية.

وقد تضمنت الدراسة ماهية البيانات الوصفية (الميتاداتا)، قابلية التشغيل البيني بأنواعها المختلفة للرسائل والاطروحات الرقمية، مع معيار البيانات الوصفية للتشغيل البيني للرسائل الجامعية والأطروحات الألكترونية، معايير تجميع البيانات الوصفية، جداول الميتاداتا وطرق تنفيذها، ونختم بتنسيقات ملفات النصوص والصور والفيديو، وغيرها من التنسيقات المستخدمة في الرسائل والاطروحات الرقمية.

الكلمات الدالة : البيانات الوصفية؛ المستودعات المؤسسية؛ الميتاداتا؛ الرسائل والاطروحات الرقمية؛ ميتاداتا الرسائل الرقمية

تمهيد:

من أبرز مظاهر عصر الإنفجار المعرفي، تكنولوجيا المعلومات حيث بروز العصر الرقمي الحالي بوجود عدد هائل من مصادر المعلومات الرقمية المتاحة علي الانترنت، وتشمل هذه المصادر والموارد والتي تزايدت ونمت نموا مضطردا في السنوات القليلة الماضية علي أنواع متعددة ومختلفة من الأشكال الرقمية مثل صفحات ومواقع الويب، مقالات الدوريات الألكترونية المحكمة ذات المحتوى المفتوح للإطلاع، والدوريات ذات الوصول المفتوح والكتب الإلكترونية والمصادر الأخرى للانتاج الفكري والتراث الثقافي و نجد ان الكثير من هذه المصادر قد تم إنتاجها رقميا بالإضافة إلى المصادر الأخرى التئ ولدت رقميا أو تم تحويلها من شكل تقليدي مطبوع إلي شكل رقمي ومن أهمها الرسائل والأطروحات الرقمية.

وقد أصبح موضوع الإتاحة علي نطاق عالمي Global Access للمعلومات الرقمية والتقليدية هو الشغل الشاغل لكثير من المؤسسات التي تنتج أو تدير مصادر وموارد المعلومات الرقمية، مثل المكتبات والجامعات ومراكز البحوث والوزارات والمنظمات الخ، ونجد أن الكثير من هذه المؤسسات قد لجأت إلي استخدام نظام واحد لإدارة مواردها الرقمية بينما البعض الأخر من هذه المؤسسات يستخدم نظم متعددة قد لا تكون قابلة للتشغيل فيما بينها .

وقد قامت الكثير من هذه المؤسسات بإنشاء معايير ميتاداتا Metadata Standards أو مخططات ميتاداتا Metadata Schemes بغرض تنظيم وإدارة مجموعاتها ومواردها الرقمية، وتميل المؤسسات التي يجمعها تخصص موضوعي في مجال واحد من مجالات المعرفة إلي إنشاء وصيانة هياكل البيانات الخاصة بها والتي يتم تفصيلها لكي تخدم أنواع محددة من مصادر المعلومات وأنواع محددة من المستفيدين، حيث تعمل هذه المعايير والمخططات علي خدمة هذه الفئات المحددة من مصادر المعلومات والمستفيدين المرتبطة بهذه المؤسسات، ولكن ما أن تضطر هذه المؤسسات إلي المشاركة بمواردها أو تقاسمها علي نطاق أوسع، أو حين تحتاج لاستخدام هذه المعلومات لأغراض أخري مختلفة تظهر مشكلة (توافقية التشغيل) أو التشغيل البيني Interoperability والتشغيل البيني طبقا لتعريف النيسو (NISO) "المنظمة الدولية لمعايير المعلومات": (هو مقدرة نظم متعددة ذات أجهزة وبرمجيات وهياكل بيانات وواجهات مختلفة علي تبادل البيانات بأقل قدر من فقدان تلك البيانات أوالخدمة الوظيفية).

ويتطلب الاحتفاظ بمكتبة من العناصر الرقمية الضرورية الحفاظ على البيانات الوصفية حول تلك الموارد، وتُعد البيانات الوصفية اللازمة للإدارة الناجحة واستخدام الموارد الرقمية أكثر شمولاً وتختلف عن البيانات الوصفية المستخدمة لإدارة مجموعات الأعمال المطبوعة والمواد المادية الأخرى، فبينما قد تسجل المكتبة بيانات الوصف الوصفية المتعلقة بكتاب ما في مجموعتها، لن يتحلل الكتاب في سلسلة من الصفحات غير المتصلة إذا فشلت المكتبة في تسجيل البيانات الوصفية الهيكلية المتعلقة بتنظيم الكتاب، ولن يتمكن العلماء من عدم تقييم قيمة الكتاب إذا فشلت المكتبة في ملاحظة أن الكتاب أُنتج باستخدام مطبعة ما، لا يمكن قول الشيء نفسه بالنسبة لنسخة رقمية من نفس الكتاب، بدون البيانات الوصفية الهيكلية، تكون صورة الصفحة أو الملفات النصية التي تشتمل على العمل الرقمي قليلة الفائدة، وهذا ينسحب أيضا على الرسالة والأطروحة العلمية الرقمية، فبدون البيانات الوصفية التقنية المتعلقة بعملية الرقمنة، قد نكون غير متأكدين من مدى دقة انعكاس النسخة الرقمية الأصلية ولأغراض الإدارة الداخلية يجب أن تتمتع المكتبة بإمكانية الوصول إلى البيانات الوصفية التقنية المناسبة لتحديث البيانات وترحيلها بشكل دوري، مما يضمن استمرارية الموارد الرقمية القيمة.

وقد دفعت الرغبة في الوصول السلس والسهل لموارد المعلومات الرقمية الكثير من المؤسسات إلي تحويل بياناتهم وبرامجهم القديمة Legacy Data والتي تم تطويرها من اجل الاستخدام الداخلي، إلي معايير تكون متاحة بصورة أوسع وذلك من اجل العرض العام لهذه الموارد أو من اجل تقاسمها مع جهات أخرى أو بغرض توفير واجهة واحدة للبحث في قواعد بيانات متعددة في نفس الوقت[١] .

مشكلة الدراسة:

يتطلب التنسيق الجديد للرسائل والأطروحات أن تحدد المكتبة المعايير أو الضوابط التي يجب أن يطبقها المفهرس / أخصائي البيانات الوصفية أثناء عملية التقديم والنشر، وفي أي نقطة يجب تطبيقها، ونظرًا لأن إنشاء الأطروحات الإلكترونية والبيانات الوصفية لها سيتطلب استخدام معايير إضافية في المكتبة، فإن تنمية مهارات الموظفين هو مصدر قلق وتساؤل، كما هو الحال مع العديد من المكتبات البحثية، حتى في ظل وجود طاقم من المفهرسين ذوي الخبرة، سيتطلب نقل هذه المهارات إلى عملية تحسين البيانات الوصفية دبلن كور والمسير إلى مارك التدريب والممارسة، يعد التحكم في إنشاء البيانات الوصفية أكثر المجالات إهمالًا وخارج مهام سير عمل المكتبة التقليدية. ستأخذ هذه الدراسة في الاعتبار قيمة تحديد واستخدام نموذج قياسي كجزء من عملية تحسين البيانات الوصفية ومراقبة الجودة بها.

أهداف الدراسة:

الغرض من هذه الورقة البحثية هو التحقيق في معايير البيانات الوصفية للرسائل والاطروحات الرقمية المتاحة في جميع أنحاء العالم ودراستها بالتفصيل، تقدم الورقة التحليل التفصيلي لهذه المعايير.

منهج الدراسة:

أعتمدت الدراسة على المنهج الوصفي التحليلي في جمع البيانات حول موضوع الدراسة على الأدوات البحثية من خلال مسح الإنتاج الفكري حول موضوع الميتاداتا وأهميتها في دعم الوصول إلى مواقع الأرشيفات الوطنية على الإنترنت باللغات العربية والانجليزية، البحث الوثائقي على الخط المباشر، وفحص مواقع الأرشيفات على الشبكة العنكبوتية العالمية .

الدراسات السابقة:

عند البحث في أدلة الأطروحات الأكاديمية المجازة وماتحت الإجازة بالجامعات المصرية والعربية من خلال الدليل الببليوجرافي للإنتاج الفكري العربي في مجال المكتبات والمعلومات بطبعاته المتعددة منذ ١٩٧٦م وحتى الآن، وأيضا المرصد الببليوجرافي لإتحاد مكتبات الجامعات المصرية للإطروحات الأكاديمية وكذلك المواقع المتخصصة على الشبكة العنكبوتية وغيرها؛ لم يعثر الباحث على أطروحة اكاديمية عربية تتعامل مع موضوع البحث بأبعاده وجوانبه المختلفة في مجال الدراسات العربية، وقد أقتصر ماتم العثور عليه من الأطروحات الأكاديمية التي تقترب قليلا أو كثيرا من موضوع موضوع الإنتاج الفكري الصادر عن المستودعات المؤسسية للرسائل والاطروحات الرقمية مايلي:

- **السيد صلاح الصاوي**: الميتاداتا وأهميتها في دعم الوصول إلى المحتوى الأرشيفي الرقمي: دراسة تطبيقية على مواقع الأرشيفات الوطنية على الإنترنت[٢]

هدفت هذه الدراسة إلى تحليل وتقييم الميتاداتا في مواقع الأرشيفات الوطنية على الإنترنت، وبيان مدى حرص الأرشيفات الوطنية على توكيد الجودة في صياغة الميتاداتا مقارنة بالمعايير العالمية في هذا الصدد وبالممارسات المتبعة في الأرشيفات الوطنية في الدول المتقدمة (الولايات المتحدة، كندا، المملكة المتحدة، فرنسا، ...)، وقد شمل التحليل (٥٧) أرشيفاً. وكشفت الدراسة عن ضعف استخدام تيجان الميتاداتا في الأرشيفات الوطنية عامة والعربية منها بوجه خاص؛ فقد تضمن (٢٦) أرشيفاً بنسبة ٤٥,٦% من الأرشيفات موضوع الدراسة (١-٢) تاج فقط من تيجان الميتاداتا، وأن أكثر الأرشيفات إنشاءً لتيجان الميتاداتا الأرشيفاتالوطنية لكندا، والمملكة المتحدة، وإسبانيا، وأستراليا على الترتيب، وأكثر تيجان الميتاداتا استخداما في مواقع الأرشيفات الوطنية على الإنترنت تاج العنوان بنسبة ٢٦% من إجمالي تيجان الميتاداتا المستخدمة، وتاج "الوصف" بنسبة ١٠,٦%، وتاج "الكلمات المفتاحية" بنسبة ٨,٨%، وتاج "المُجمع"Generator بنسبة ٥,٨%. كما احتوت بعض تيجان الميتاداتا على قيم غير معبرة عن المضامين الموضوعية للمواقع. وقد اختتمت الدراسة بتقديم عدة توصيات يمكن أن تسهم في معالجة أوجه القصور التي كشفت عنها الدراسة، والتي تؤثر على الإفادة الكاملة من المحتوى الرقمي لمواقع الأرشيفات الوطنية على الإنترنت، وضمان سهولة، وسرعة الوصول إليها من قبل محركات البحث.

- **وسام يوسف بن غيدة**. المستودعات الرقمية المؤسساتية ودورها في إتاحة المحتويات الرقمية للمكتبات الجامعية الجزائرية على شبكة الانترنت[٣]: لخص الباحث ورقته بالقول "تعد المستودعات الرقمية المؤسساتية من أحدث التكنولوجيات المطبقة داخل المكتبات، خاصة المتواجدة منها بالمؤسسات الجامعية، حيث تكتسب هذه التكنولوجيا الحديثة أهمية كبيرة لدى هذا النمط من المكتبات، وهذا لاشتمالها على المهام الرئيسية للمكتبات في البيئة الرقمية، حيث تعمل على حفظ وإدارة مجموعاتها، وكذا إتاحتها دون عوائق مادية وقانونية على الويب، كما أكدت الكثير من الدراسات

العلمية بان المستودعات الرقمية سوف تشكل مستقبلا جزءا أساسيا في التكتلات الإقليمية بين المكتبات،سنحاول في هذه الدراسة تسليط الضوء على دور المستودعات الرقمية المؤسساتية في إتاحة المصادر الرقمية للمكتبات الجامعية الجزائرية على شبكة الانترنت".

- **مؤمن نشرتي**. نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في أطار التنقيب عن البيانات: دراسة تحليلية[4]: سعت هذه الدراسة إلى توضيح مفهوم التكامل المعرفي في سياق تقني، وأشهر المبادرات التقنية التي ساهمت في تحقيقه على صعيد المحتوى المتاح على الشبكة العنكبوتية، ثم تناولت الدراسة رصد القدرة على توظيف مبادرة الانطولوجيات في أطار التنقيب عن البيانات سعياً بذلك إلى الاستفادة منهما فيما يعرف بنظم الانطولوجيا للتنقيب على الويب Ontology-based Web mining لتحقيق التكامل المعرفي لمحتوى شبكة الويب، اعتمدت الدراسة على المنهج المسحي لرصد التحديات الخوارزمية واللغوية التي تواجه شبكة الويب في تحقيق التكامل المعرفي، كذلك تعتمد الدراسة على المنهج التحليلي في رصد وتحليل واقع قدرات الانطولوجيا في أن توظف لتحقيق التكامل المعرفي للمحتوى المتاح على الويب، وكان من اهم نتائج الدراسة هي قدرة التنقيب على الويب على تحقيق مستوى مرتفع من التشغيل المتبادل على صعيد محتوى الويب بأكمله لتوفير التكامل المعرفي.

- **عمر حسن عبد الرحمن**. التشغيل البيني للميتادات[5]: تستعرض هذه الدراسة أهم معايير الميتاداتا المستخدمة بغرض تبادل البيانات والكشف عن موارد المعلومات بشبكة الانترنت. ثم تلقى الضوء على مشكلة التشغيل البيني interoperability لهذه المعايير مع بعضها البعض من أجل تسهيل عملية تبادل البيانات بين نظم المعلومات المختلفة. وتتطرق الدراسة في هذا السياق إلى عملية إنشاء جداول الميتاداتا Metadata Crosswalks والتي يتم عن طريقها تحقيق عملية التشغيل البيني لمعايير ومخططات الميتاداتا المختلفة. كذلك توضح هذه الدراسة طرق تنفيذ جداول الميتاداتا و بيان أهم المشاكل التي تواجه عملية إنشاء هذه الجداول والتي تأتى في مقدمتها عدم وجود تطابق تام بين معايير ومخططات الميتاداتا المختلفة.

Potvin, Sarah, Thompson, Santi: An analysis of evolving metadata influences, standards, and practices in electronic theses and dissertations[6]

أستخدمت هذه الدراسة نهجًا مختلطًا لزيادة الوعي بالاختلافات بين الممارسات الحالية ومعايير البيانات الوصفية والمبادئ التوجيهية للرسائل والأطروحات الإلكترونية (ETDs)، التحليل متجذر في الأدبيات حول جودة البيانات الوصفية، وبخاصة البيانات الوصفية القابلة للمشاركة أو الموحدة ، وقابلية التشغيل البيني، مع الانتباه إلى تأثير الأنظمة والأدوات والممارسات على بيانات تعريف تاريخ الرسائل والاطروحات الرقمية، بحيث يوضع في الاعتبار الفلسفات التي وجهت تصميم العديد من معايير البيانات الوصفية، يخدم فحص قضايا قابلية التشغيل البيني الدلالي كتعبير عن الحاجة إلى تقدم مثالي أكثر قوة للأمام، متجذر في نماذج دورة حياة البيانات الوصفية ويهتم بالتعامل طويل الأمد مع الرسائل والاطروحات الرقمية والحفاظ عليها.

Mikael, Adam; Brace, Tim; Leggett, John; McFarland, Mark; Phillips, Scott (2007). Developing a Common Submission System for ETDs in the Texas Digital Library[7].

مكتبة تكساس الرقمية عبارة عن اتحاد جامعات منظم لتوفير بنية تحتية رقمية واحدة للأنشطة العلمية لجامعات تكساس فإمكانات النمو لمستودع واحد على مستوى الولاية تُعد الأكبر، ولتسهيل إنشاء هذه المجموعة الموحدة، وافقت المدارس على معيار بيانات وصفية مشترك يمثله مخطط MODS XML، وعلى الرغم من أن هذا يخلق خطًا أساسيًا لاتساق البيانات الوصفية، إلا أن هناك غموضًا في تفسير المخطط الذي يخلق تحديات قابلية الاستخدام والتشغيل البيني، لا يعالج المخطط مشكلات دقة الاسم ، وتحتاج بعض عناصر بيانات الوصف الوصفية إلى التناسق في التنسيق ومستوى الأهمية بحيث تعمل وظائف المستودع المشتركة بشكل متناسق عبر المجموعة، تقرر أن هناك حاجة إلى نقطة إلتقاط مشتركة للرسائل والاطروحات الرقمية لجمع البيانات الوصفية بطريقة متسقة وموثوقة، تم تشكيل مجموعة عمل تتألف من ممثلين من خمس جامعات لإنجاز تلك المهمة.

- Eden, B., McCutcheon, S., Kreyche, M., Maurer, M. B., & Nickerson, J. (2008). Morphing metadata: maximizing access to electronic theses and dissertations. Library Hi Tech.[8]

هدفت هذه الورقة البحثية إلى وصف العمل في مكتبات جامعة ولاية كينت وخدمات وسائل الإعلام لتعزيز وابتكارو تخزين الرسائل والأطروحات الإلكترونية (ETD) في مركز ETD في أوهايو لينك ، لإيجاد طرق فعالة لتمثيل هذه المواد العلمية الفريدة في كتالوج المكتبة، ولتعزيز وضع معايير كتالوج المكتبات على مستوى الولاية للرسائل والاطروحات الرقمية.

ماهي البيانات الوصفية ؟

كانت البيانات الوصفية تُستخدم تقليديًا في فهارس وبطاقات المكتبات حتى الثمانينيات، عندما حولت المكتبات بيانات الفهرس إلى قواعد بيانات رقمية OPAC، أصبحت البيانات الوصفية تمتلك أغراض مختلفة فهي تساعد المستخدمين في العثور على المعلومات ذات الصلة بموضوعاتهم واكتشاف الموارد (المصادر)، كما أنها تساعد في تنظيم الموارد الإلكترونية، وتوفير التعريف الرقمي، وأرشفة الموارد والحفاظ عليها، تسمح البيانات الوصفية للمستخدمين بالوصول إلى الموارد من خلال "السماح بالعثور عليها من خلال المعايير ذات الصلة، وتحديد تلك الموارد، وجمع الموارد المتشابهة معًا، والتمييز بين الموارد المختلفة، وإعطاء معلومات عن الموقع _ حتى في مشهد جديد في بض الأحيان بدأت بعض المؤسسات الحكومية الوطنية تَستخدم هذه البيانات لأغراض تحليل حركة المرور كما يمكن استخدامها للمراقبة الجماعية.

في العقد الأول من القرن الحادي والعشرين، ومع تزايد تخزين البيانات والمعلومات رقميًا، تم وصف هذه البيانات الرقمية باستخدام معايير البيانات الوصفية، يُقال أن الوصف الأول "للبيانات الوصفية" لأنظمة الكمبيوتر تمت ملاحظته من قبل خبراء مركز الدراسات الدولية التابع لمعهد ماساتشوستس للتكنولوجيا ديفيد جريفيل وستيوارت ماكنتوش في عام ١٩٦٧: "باختصار إذن لدينا بيانات بلغة كائن حول أوصاف موضوع البيانات ورموز الرموز الخاصة بـ البيانات، لدينا أيضًا عبارات بلغة وصفية تصف علاقات البيانات والتحولات، وهي العلاقات بين القاعدة والبيانات.

توجد معايير متفردة (مميزة) للبيانات الوصفية لمختلف التخصصات (على سبيل المثال، مجموعات المتحف، الملفات الصوتية الرقمية، المواقع الإلكترونية إلخ) إن وصف محتويات وسياق البيانات أو ملفات البيانات يزيد من فائدتها، فعلى سبيل المثال قد تتضمن صفحة الويب بيانات وصفية تحدد لغة البرنامج التي تمت كتابة هذه الصفحة به (على سبيل المثال HTML) والأدوات التي تم استخدامها لإنشائها والموضوعات التي تدور حولها الصفحة ومكان العثور على مزيد من المعلومات حول الموضوع، يمكن

لهذه البيانات الوصفية تحسين تجربة القارئ تلقائيًا وتسهيل عثور المستخدمين على صفحة الويب عبر الإنترنت.

في العديد من البلدان، تقوم المؤسسات الحكومية بشكل روتيني بتخزين البيانات الوصفية حول رسائل البريد الإلكتروني والمكالمات الهاتفية وصفحات الويب وحركة مرور الفيديو واتصالات IP ومواقع الهواتف المحمولة.

والبيانات الوصفية تعني "بيانات حول البيانات" على الرغم من أن بادئة "meta" (من حرف الجر اليوناني والبادئة) تعني "بُعد" أو "بَعد"، إلا أنها تُستخدم لتعني "about" في نظرية المعرفة، يتم تعريف البيانات الوصفية على أنها البيانات التي توفر معلومات حول جانب واحد أو أكثر من البيانات يتم استخدامه لتلخيص المعلومات الأساسية حول البيانات التي يمكن أن تجعل تتبع البيانات المحددة والعمل معها أسهل، تتضمن بعض الأمثلة ما يلي: وسائل إنشاء البيانات، الغرض من البيانات، وقت وتاريخ الإنشاء، منشئ أو مؤلف البيانات، الموقع على شبكة الكمبيوتر حيث تم إنشاء البيانات، المعايير المستخدمة، حجم الملف، جودة البيانات، مصدر البيانات، العملية المستخدمة لإنشاء البيانات وغيرها.[٩]

قابلية التشغيل البيني Interoperability

مصطلح يستخدم لوصف النماذج المختلفة لتبادل البيانات بين أنظمة المعلومات، والتي يتم بناؤها على أساس الاتفاق على المعايير والبروتوكولات في عالم المستودعات المؤسسية (IRs)، غالبًا ما يمثل هذا المصطلح طريقة محددة للتشغيل البيني، أي تجميع البيانات الوصفية من خلال OAI-PMH (بروتوكول مبادرة الأرشيف المفتوح لحصاد البيانات الوصفية)، ومع ذلك في المجالات الأخرى لأنظمة المعلومات عبر الإنترنت (على سبيل المثال، مدونات الوسائط عبر الإنترنت، وأنظمة التعلم الإلكتروني، وشبكات الند للند وغيرها) يشير المصطلح إلى مناهج مختلفة لتكامل المصادر الموزعة[١٠]

ولكن الرسائل والأطروحات الإلكترونية (ETD) هي من حيث المحتوى سريعة النمو مقارنة بالأنواع الأخرى من المستودعات بالإضافة إلى ذلك، فإن تعدد الأساليب المقدمة لتجميع موارد تلك الرسائل والأطروحات جعلها نوعاً مثيرًا للاهتمام لتقييم أساليب التشغيل البيني، ووفقًا لـ Goncalves et al. (2000) هناك العديد من الخصائص الفريدة للرسائل والأطروحات الرقمية التي يجب أخذها في الاعتبار عند تقييم طرق التشغيل البيني، هذه الخصائص المتنوعة تشمل:

- **اللغة - Language** : لتوفر المحتويات بعدة لغات .
- **المقياس (حجم الوثيقة) - Scale** : مستندات بحجم الكتاب كبيرة، بالإضافة إلى ملفات الوسائط المتعددة ذات الصلة
- **التنسيق - Format** : توفر المحتوى الكامل بتنسيقات مختلفة (XML، PDF، HTML، إلخ) .
- **ببليوغرافيا - Bibliography** : تعدد المراجع الببليوغرافية .
- **البيانات الوصفية - Metadata** : مجموعات غنية من عناصر البيانات الوصفية التي تختلف في الجودة.
- **المستخدمين - Users** : تنوع في مدى ونطاق اهتمامات المستخدم (Goncalves et al.، 2000)[١١].

ونظرًا لتنوع البروتوكولات ولأن كل من هذه البروتوكولات مصممة لغرضها المحدد، ولتجنب المقارنة والتقييم، يتم التشغيل البيني بين تلك البروتوكولات ويتم التجميع والبحث الموزع والتشارك والنشر على النحو الذي يجعلها أكثر إيجابية ويقلل من الأثر السلبي لتنوعها .

الأنواع المختلفة من قابلية التشغيل البيني داخل ETDs:

عادةً ما يتم استخدام مصطلحات التكامل والتشغيل البيني معًا لوصف المعايير والأساليب والبروتوكولات التي من خلالها يمكن أن تتدفق حزم البيانات بين أنظمة المعلومات المختلفة (Cardoso and Sheth، 2005)[١٢] يشير التكامل إلى مجموعة من الأساليب لتسهيل تدفق البيانات بين أنظمة المعلومات القابلة للتشغيل البيني (.Chen et al، 2008)[١٣] .

وعلى الرغم من أن جميع نظم المستودعات المؤسسية مصممة لمواجهة تحديات متشابهة تقريبًا، إلا أن الهياكل المنخفضة المستوى منها يتم تطويرها بشكل ضعيف، لذلك فإن إجراء تحويل البيانات بدون فقدانها بين نظامي المستودعات عادة لا يكون عملية مباشرة، من ناحية أخرى على الرغم من التنوع في البنية منخفضة المستوى يمكن أن تندرج جميع الوظائف الرئيسية لحلول IR ضمن فئات استيعاب المحتويات وإدارتها ونشرها (.Tarrant et al، 2009)[١٤] ومن أجل تحسين قابلية التشغيل البيني في هذه المستويات المتعددة، يلزم أن تكون حلول ETD المختلفة متوافقة مع بروتوكولات ومعايير التشغيل البيني المقبولة على نطاق واسع.

تذكر دراسة (Hakimjavadi, H. And Noorman Masrek, M. 2013)[15] أن هناك أربع مجموعات من البروتوكولات ذات الوظائف المحددة لكل منها طبقا لطبيعة تكوينها والوظيفة التي أنشأت بمقتضاها وهي:

١. **بروتوكولات التجميع- Aggregation protocols**: تتضمن فئة التجميع بروتوكولات تُسهل عملية كشف البيانات الوصفية القابلة للمشاركة وألتقاطها وجمعها بحيث يمكن فهمها واستخدامها خارج بيئتها المحلية.

٢. **البروتوكولات النقابية (التشاركية)- Syndication protocols** : تسهل بروتوكولات المشاركة تدفق بيانات العميل والخادم من خلال إرسال "خلاصات" لمعلومات موقع الويب إلى المشتركين، مما يسمح لهم بجمع المعلومات التي يتم تحديثها بشكل متكرر من مصادر متعددة وتنظيمها وعرضها بسهولة.

٣. **نشر البروتوكولات- Publishing protocols**: نظرًا لأنه يتم إنشاء الـ IRs في المؤسسات العامة والخاصة والوطنية والدولية، يتعين على الباحثين اليوم إيداع أطروحاتهم في العديد من الـ IRs، وفي الوقت نفسه غالبًا ما يكون الباحثون على استعداد لتسجيل نتائج أبحاثهم في المستودعات الرقمية الموضوعية، وفي كثير من الحالات يرغبون فقط في توفير الوصول العام إلى عملهم من خلال مستودعات الوصول المفتوح للرسائل والأطروحات، في مثل هذه الحالات يتمثل الأسلوب العملي في إيداع المعلومات في المستودعات التي تنشر المحتوى تلقائيًا إلى المستودعات المؤسسية الأخرى من خلال آلية "المستودع إلى مستودع الإيداع" ستكون هذه الآلية قابلة للتطبيق إذا كانت جميع الأنظمة المعنية متوافقة مع بروتوكول النشر المتفق عليه.

٤. **بروتوكولات البحث الموزعة- Distributed search protocols** : بروتوكول Z39.50 : تم إصدار Z39.50 في أواخر السبعينيات كبروتوكول خادم العميل لاسترداد المعلومات من اتحاد الخوادم البعيدة، وتمت الموافقة عليه كمعيار ANSI / NISO Z39.50 في عام ١٩٨٨ (Lynch،

1991) كانت النقطة المهمة في هذا البروتوكول هي الفصل بين واجهة بحث المستخدم المرفقة سابقًا وآلية البحث في المكتبة بتطبيق Z39.50، كان برنامج جانب العميل قادرًا على التفاعل مع مجموعة واسعة من محركات بحث المكتبة من أجل توفير بيئة موحدة لخدمات استرجاع المعلومات.

وعلى الرغم من الاستخدام الواسع النطاق لـ Z39.50 في أنظمة المعلومات المعاصرة، إلا أنه يعد حلًا سابقًا للويب، ومع ذلك، فقد تم إجراء العديد من المحاولات لتحديث هذا البروتوكول، والتي كان أبرزها تقديم الإصدارات المحدثة من البروتوكول في عامي ١٩٩٥ و ٢٠٠٣، وإدخال ملف تعريف باث، كان هناك جهد آخر لجعل خدمات Z39.50 متاحة على نطاق أوسع وهو تطوير Z39.50) ZING (International) Next Generation) والذي أدى أخيرًا إلى إدخال البروتوكولات المزدوجة لـ SRU / SRW كبديل لـ Z39.50.

بروتوكول SRU / SRW

يتكون معيار البحث / الاسترداد عبر الويب (SRW) وهو من البروتوكولات المزدوجة لـ SRU و SRW، وهما أسلوبا وصول قائم على URL و SOAP على التوالي، تم تقديمه كبديل لـ Z39.50، وكان الغرض الرئيسي من SRU / SRW هو تسهيل البحث عن الموارد واسترجاعها من خلال طريقة تعتمد على الخادم - العميل وتقليل العوائق التي تواجه المستفيدين للوصول إلى نطاق أوسع من موفري خدمات البحث، باستخدام XML لتبادل البيانات و CQL كلغة استعلام، من الواضح أن تطبيق SRW أكثر وضوحًا من Z39.50 في كل من الخادم والعميل، وفيما يتعلق بهيكل خدمة الويب يتم تحديد SRW على أنه بروتوكول قائم على SOAP، بينما تتبع SRU مبادئ REST لتقديم خدمات الويب، على الرغم من الاختلاف في هيكلها، فإن وظيفة كلا البروتوكولين ومجموعة العمليات المدعومة هي نفسها تقريبًا، يتم تعريف العمليات الثلاث لـ Explain و Scan و Retrieve / Search في SRW للاستعلام عن الوظائف المدعومة في خادم بعيد، وإدراج المصطلحات المحددة في قاعدة بيانات بعيدة، وإرسال الاستعلامات واسترداد النتائج من خادم البحث على التوالي، في حين أن SRU هو بروتوكول أكثر قابلية للقراءة ويسهل تنفيذه، فإن SRW قادر على قبول استفسارات للعملاء وآلية إدخال أكثر تعقيدًا .

وتخلص دراسة (Hakimjavadi, H. And Noorman Masrek, M 2013) والتي أجريت على ثمانية من المستودعات المؤسسية للرسائل والأطروحات الرقمية أن هدف المقارنة بين البروتوكولات المختلفة المتخصصة في وظائف مختلفة ليست هي الهدف بحد ذاته، ومع ذلك فقد أظهرت أن هناك بعض السمات المشتركة في هذه البروتوكولات التي تجعلها قابلة للاستخدام في مجالات أخرى غير تخصصاتها.

"كان الغرض من هذه الدراسة الإستطلاعية هو تقييم حالة ثمانية بروتوكولات في بيئات مستودعات ETD وعلى الرغم من مقارنة هذه البروتوكولات الثمانية في أربع فئات منفصلة، أشارت النتائج إلى أنه يمكن استخدام تلك البروتوكولات بالتبادل، بغض النظر عن فئتها المحددة (داخل الفئات الأربع) على سبيل المثال بينما يقع ATOM ضمن فئة المشاركة، فهو مرشح محتمل لاستبدال OAI-PMH، وهو البروتوكول المعروف جيدًا لتبادل البيانات الوصفية، أشارت نتائج الاستطلاعات الثلاثة إلى أن ATOM مقبول على نطاق واسع كبروتوكول RESTfull حديث في أنظمة إدارة ETD ومع ذلك يعتمد عدد قليل من برامج المستودعات المؤسسية عليه وهو في الأصل يعد تخصيصه مناسبا لخلاصات ATOM المنشورة، يعتبر الافتقار إلى التخصيص بدوره عقبة أمام استبدال بروتوكول OAI-PMH المتقدم بـ ATOM".

وقد جاء في ملخص تلك الدراسة المعنونة (Evaluation of interoperability protocols in repositories of electronic theses and dissertations) وهي من أبرز الدراسات التي أطلع عليها الباحث في هذا المجال مانصه:

" أشار تحليل النتائج التي تم الحصول عليها من ثلاثة استطلاعات إلى أنه على الرغم من عيوبه، لا يزال OAI-PMH هو بروتوكول التشغيل البيني الأكثر استخدامًا بين موفري ETD وكذلك مطوري برامج ETD والمنفذين، بالإضافة إلى تقييم البروتوكولات وإيجاد نقاط ضعفها وقوتها، أدت هذه الدراسة إلى بعض النتائج الفرعية:

لوحظ أنه لا توجد دائمًا علاقة مباشرة بين وظيفة وشعبية بروتوكولات التشغيل البيني.

لا يتم استخدام معظم البروتوكولات التي تم تقييمها بكامل إمكاناتها في مستودعات ETD، الاستثناءات هي OAI - PMH و z39.50. لا يزال الأول مستخدمًا على نطاق واسع، بينما لا يُتوقع رؤية واضحة لمستقبله، ولم يعد الأخير يلعب دور بروتوكول في الـ IRs.

يمكن استخدام ثلاثة بروتوكولات لـ ATOM و PMH و ORE (وفي بعض الحالات RSS 2.0) بالتبادل كمعايير للتشغيل البيني في المستودعات، المشاكل الرئيسية في هذا الصدد تكمن في عدم وجود قابلية كافية للتخصيص التي توفرها منتجات برامج IR ونقص التوثيق المباشر."

ويعد تطوير بروتوكول جديد لأنظمة المعلومات عملية تستغرق وقتًا طويلاً، ولا تؤدي بالضرورة إلى قبول واسع من قبل المجتمع، على الرغم من قدراتها وكفاءتها، وهذا هو السبب في الأعتماد على استخدام البروتوكولات الحالية بكامل إمكاناتها، بدلاً من بناء بروتوكولات أخرى في المستقبل، أصبحت مبدأ مقبولاً اليوم؛ تشير الدراسة إلى أن المعايير الكاملة لبروتوكولات التشغيل البيني لا تُستخدم في مستودعات ETD، بينما يتوفر الأساس التقني لاستخدام هذه البروتوكولات في مناطق جديدة من قابلية التشغيل البيني، يمكن استنتاج أن تطوير أدوات جديدة للاستفادة من هذه القدرات هو المفضل لتطوير بروتوكول جديد.

من المعروف أن المكتبات لديها تقليد قوي في اقتراح المعايير لبقية مجتمع المعلومات، نظرًا لأن خدمات المكتبة يتم دمجها بشكل متزايد في خدمات الإنترنت، فقد حان الوقت الآن للاستفادة من هذه التجارب لبقية العالم عبر الإنترنت في الـ IRs. في هذا الصدد ومن أجل تحسين قابلية التشغيل البيني بين المستودعات المتماثلة، ينبغي اعتماد الحد الأدنى لعدد البروتوكولات الأبسط المستخدمة على نطاق واسع"[١٦].

معيار البيانات الوصفية للتشغيل البيني للرسائل الجامعية والأطروحات الألكترونية ETD-MS:

البيانات الوصفية هي المصطلح المستخدم لوصف البيانات حول البيانات، وفقًا لـ NDLTD بالنسبة للعناصر الموجودة في مستودعات الوصول المفتوح، تتكون البيانات الوصفية عادةً من مرجع ببليوجرافي كامل وملخص وكلمات رئيسية ومعلومات مماثلة حول الكائن الرقمي (العناصر) الرقمية ذات الصلة، مثل ETDs تشير إرشادات UGC-ETD إلى أن الأنواع الثلاثة التالية من البيانات الوصفية مرتبطة بالكيانات الرقمية:[١٧]

- **البيانات الوصفية (بيانات الوصف):** وتتضمين المحتوى أو الوصف الببليوغرافي الذي يتكون من كلمات رئيسية وواصفات موضوعية.
- **البيانات الوصفية الإدارية أو التقنية**: تتضمن تفاصيل عن المصدر الأصلي، وتاريخ الإنشاء، وإصدارة الكائن الرقمي، وتنسيق الملف المستخدم، وتقنية الضغط المستخدمة، وعلاقة الكيان الرقمي، وما إلى ذلك. قد تكون البيانات الإدارية موجودة داخل الكيان الرقمي أو خارجه، وهي مطلوبة لفترة طويلة لضمان إدارة المجموعة الرقمية علي المدى الطويل.

- **البيانات الوصفية الهيكلية**: عناصر داخل العناصر الرقمية تُسهل التنقل، على سبيل المثال جدول المحتويات، والفهرس على مستوى الإصدار أو مستوى الصوت، والصفحة في كتاب إلكتروني، إلخ.

قامت المؤسسات التي تتعامل مع الرسائل الجامعية والأطروحات بتطوير معاييرها الخاصة أو تكييف معايير البيانات الوصفية الحالية تبعاً ل NDLTD، أكبر اتحاد في العالم لبرامج ETD والذي يحتفظ بفهرس اتحادي دولي للرسائل والأطروحات الرقمية من جميع أنحاء العالم، قد حدد "مجموعة قياسية من عناصر البيانات الوصفية"[١٨] المستخدمة لوصف رسالة أو أطروحة إلكترونية وهو ليس بديلاً عن مخطط البيانات الوصفية المستخدم بالفعل في أي جامعة، ولكن يجب استخدامه كدليل إرشادي لتطوير مسار متقاطع بين معايير البيانات الوصفية المحلية ومعيار موحد يُستخدم لمشاركة المعلومات حول الرسائل والأطروحات الرقمية، تستند المجموعة إلى عمل مبادرة دبلن الأساسية للبيانات الوصفية ومجموعة المؤهلات المقترحة التي قدمتها مجموعة عمل DC Date.

معايير تجميع البيانات الوصفية

بروتوكول مبادرة الأرشيف المفتوح لحصاد البيانات الوصفية (OAI-PMH) -

" Harrod's Librarian"s Glossary & Reference Books mentions, OAI-PMH " يذكر أن"OAI-PMH هو البروتوكول الذي يوفر إطارًا للتشغيل البيني مستقل عن التطبيق لمبادرة الأرشيفات

OAI-PMH هو بروتوكول تم تصميمه لجعل البيانات الوصفية المقروءة آليًا متاحة على نطاق واسع للاستخدام، وقد ترسخ تطوير بروتوكول OAI-PMH في أواخر عام ١٩٩٩ في سانتافي، ولاية نيو مكسيكو لمعالجة مشاكل عالم المستودعات الإلكترونية، نظرًا لأن خوادم تلك المستودعات والكيانات الرقمية أصبحت أكثر شيوعًا، فكان من الصعب دعم البحث عبر مستودعات متعددة، تحتاج المستودعات إلى قدرات أكبر لتحديد الأوراق التي تم إيداعها في مستودعات أخرى والوصول إليها تلقائيًا، يوفر بروتوكول مبادرة الأرشيفات المفتوحة لحصاد البيانات الوصفية[١٩] إطارًا للتشغيل البيني مستقل عن التطبيق يعتمد على حصاد البيانات الوصفية، وهناك فصلين من المشاركين في إطار OAI-PMH:

يدير مقدمو البيانات الأنظمة التي تدعم OAI-PMH كوسيلة لكشف البيانات الوصفية؛ ويستخدم مقدمو الخدمات البيانات الوصفية التي يتم جمعها عبر OAI-PMH كأساس لبناء خدمات ذات قيمة مضافة.

يتم نقل البيانات الوصفية المخزنة في قاعدة بيانات موفري البيانات بشكل مجمّع إلى قاعدة بيانات البيانات الوصفية لموفري الخدمة، يتم نقل البيانات الوصفية في سلسلة من الطلبات والاستجابات بين مزودي البيانات وموردي الخدمة / الحاصد؛ يعتمد بروتوكول OAI-PMH على إطار عمل HTTP transaction للتواصل بين الحاصدة والمستودع، يمكن تقديم الطلبات باستخدام طرق HTTP GET أو POST، يتم ترميز جميع الردود الناجحة في XML، وتتم الإشارة إلى جميع ردود الاستثناءات والتحكم في التدفق بواسطة أكواد حالة HTTP.

وتلعب جداول الميتاداتا Metadata Crosswalks دورا محوريا في تحقيق هذه الأغراض :

- تحويل البيانات Data Conversion من معيار إلي معيار أخر جديد او مختلف.
- جمع البيانات Data Harvesting وإعادة تنظمها Repackaging من مصادر متعددة.
- دمج مصادر وموارد بيانات متنوعة.

معايير الميتاداتا لـ Metadata Standards ETDs

توجد العديد من معايير الميتاداتا Metadata Standards لمقابلة احتياجات المستفيدين المختلفة من المعلومات، ومن أكثر هذه المعايير والمخططات طبقا لمعدلات الأخذ بها وسهولة التعامل معها وذيوعها وأنتشارها بين برامج بناء المستودعات الرقمية للرسائل والأطروحات الرقمية :

١) الدبلن كور Dublin Core: إن Dublin Core ™، المعروف أيضًا باسم Dublin Core ™ Metadata Element Set، هو مجموعة من خمسة عشر عنصرًا "أساسيًا" (خصائص) لوصف الموارد (عناصر الكيانات الرقمية)، تم توحيد تلك المعايير ليصبح دبلن كور الموجد .

تشير كلمة "دبلن" إلى دبلن، أوهايو، الولايات المتحدة الأمريكية حيث نشأ المخطط أثناء ورشة عمل OCLC / NCSA Metadata الدعائية لعام ١٩٩٥، التي استضافتها OCLC (المعروفة في ذلك الوقت باسم مركز مكتبة الكمبيوتر عبر الإنترنت)، وهي عبارة عن اتحاد مكتبات مقره دبلن، بالشراكة مع المركز الوطني لتطبيقات الحوسبة الفائقة (NCSA)، وتشير كلمة "الأساسية" إلى مصطلحات البيانات الوصفية على أنها "واسعة النطاق وعامة يمكن استخدامها لوصف نطاق واسع من الموارد"، تم إنشاء دلالات دبلن كور والمحافظة عليها من قبل مجموعة دولية متعددة التخصصات من المتخصصين في المكتبات وعلوم الكمبيوتر وترميز النص والمتاحف وغيرها من المجالات ذات الصلة بالمعرفة والممارسة.

ابتداءً من عام ٢٠٠٠، ركز مجتمع Dublin Core على "ملفات تعريف التطبيق" – وهي الفكرة القائلة بأن سجلات البيانات الوصفية ستستخدم Dublin Core مع المفردات المتخصصة الأخرى لتلبية متطلبات التنفيذ الخاصة- خلال ذلك الوقت - كان عمل اتحاد (RDF)، في مرحلة النضج كجزء من مجموعة ممتدة من مصطلحات البيانات الوصفية DCMI، أصبحت Dublin Core واحدة من أكثر المفردات شيوعًا للاستخدام مع RDF، مؤخرًا في سياق حركة البيانات المرتبطة.

توفر مبادرة دبلن الأساسية للبيانات الوصفية DCMI[1] منتدى مفتوحًا لتطوير معايير البيانات الوصفية القابلة للتشغيل البيني على الإنترنت لمجموعة واسعة من الأغراض ونماذج الأعمال، تشمل أنشطة DCMI مجموعات العمل، والمؤتمرات وورش العمل العالمية، ومعايير الاتصال، والجهود التعليمية لتعزيز القبول الواسع لمعايير وممارسات البيانات الوصفية؛ في عام ٢٠٠٨ انفصلت DCMI عن OCLC وتم دمجها ككيان مستقل.

في الوقت الحالي، تتم مراجعة جميع التغييرات التي تم إجراؤها على معيار دبلن كور من قبل مجلس استخدام DCMI في سياق سياسة مساحة أسماء (DCMI-NAMESPACE) DCMI)، تصف هذه السياسة كيفية تعيين المصطلحات وتضع أيضًا حدودًا لمقدار التغييرات التحريرية المسموح بها للتسميات والتعريفات وتعليقات الاستخدام.

يمكن استخدام بيانات تعريف دبلن كور لأغراض متعددة، بدءًا من وصف الموارد البسيط إلى دمج مفردات البيانات الوصفية لمعايير البيانات الوصفية المختلفة، إلى توفير إمكانية التشغيل البيني لمفردات البيانات الوصفية في سحابة البيانات المرتبطة وتطبيقات الويب الدلالي وهو معيار مبسط ومختصر يتكون من خمسة عشر عنصراً أنشأت أساساً لمساعدة المؤلفين لوصف مصادرهم الرقمية على الشبكة العنكبوتية وتتمثل تلك العناصر في : العنوان، المُنشئ للموضوع (المسئول عن المحتوى الموضوعي)، بيانات الوصف، الناشر، المساهم، التاريخ، النوع، الشكل، المعرّف، المصدر، اللغة، العلاقة، التغطية والحقوق.

تضمن معيار دبلن كور في الأصل مستويين: بسيط ومؤهل، يتكون دبلن كور البسيط من ١٥ عنصرًا؛ يتضمن دبلن كور المؤهل ثلاثة عناصر إضافية (الجمهور والمصدر وحامل الحقوق)، بالإضافة إلى مجموعة من تنقيحات العناصر (تسمى أيضًا المؤهلات) التي يمكن أن تنقح دلالات العناصر بطرق قد تكون مفيدة في اكتشاف الموارد.

منذ عام ٢٠١٢، تم دمج الاثنين في شروط البيانات الوصفية DCMI كمجموعة واحدة من المصطلحات باستخدام نموذج بيانات .[RDF] وتلك العناص تفصيلا هي:

يتكون الإصدار ١,١ من DCMES الأصلي من ١٥ عنصر بيانات وصفية، تم تعريفها بهذه الطريقة في المواصفات الأصلية:

- المساهم - "كيان مسؤول عن تقديم مساهمات إلى المصدر(المورد الرقمي)".
- التغطية - "الموضوع المكاني أو الزمني للمصدر، والتطبيق المكاني للمصدر، أو الولاية القضائية التي بموجبها يكون المصدر ذا صلة بها"
- المنشئ - "كيان مسؤول بشكل أساسي عن إنشاء المصدر"
- التاريخ - "نقطة أو فترة زمنية مرتبطة بحدث في دورة حياة المصدر.
- الوصف - "حساب المصدر"
- التنسيق - "تنسيق الملف أو الوسيط المادي أو أبعاد المصدر".
- المعرّف - "إشارة (متفردة) لا لبس فيها إلى المصدر ضمن سياق معين."
- اللغة - "لغة المصدر".
- الناشر - "كيان مسؤول عن إتاحة المصدر"
- العلاقة - "مصدر ذو صلة".
- الحقوق - "معلومات حول الحقوق الموجودة في المصدر وعبره"
- المصدر - "مصدر ذو صلة يُشتق منه المصدر الموصوف".
- الموضوع - "موضوع المصدر".
- العنوان - "اسم يُعطى للمصدر".
- النوع - "طبيعة أو نوع المصدر".

كل عنصر من عناصر دبلن كور اختياري ويمكن تكراره، أنشأ DCMI طرقًا قياسية لتحسين العناصر وتشجيع استخدام مخططات الترميز والمفردات، لا يوجد ترتيب محدد في دبلن كور لتقديم العناصر أو استخدامها، أصبح دبلن كور أحد معاييرالمواصفات القياسية العالمية في ال NISO، مثل Z39.85، و IETF RFC 5013 في عام ٢٠٠٧، ومعيار ISO 15836 في عام ٢٠٠٩ ويستخدم كعنصر بيانات على مستوى القاعدة لوصف موارد التعلم في البيانات الوصفية ISO / IEC 19788-2 لمصادر التعلم (MLR)[٢٠]

دبلن كور المؤهل (المطور) يحتوي على بعض التطوير والتنقيحات للمصطلحات النموذجية والإضافية لتوسيع عناصره الخمسة عشر ولتتماشى مع التغيرات والمستجدات في وصف البيانات.

٢) معيار مارك MARC Standard:

معايير مارك (الفهرسة المقروءة آليًا) (Machine-readable cataloging MARC) هي مجموعة من التنسيقات الرقمية لوصف العناصر المُفَهرسة بواسطة المكتبات، مثل الكتب وغيرها من مصادر المعلومات؛ طورته عالمة الكمبيوتر الأمريكية هنرييت أفرام مارك في الستينيات من خلال العمل مع مكتبة الكونجرس لغرض إنشاء سجلات يمكن قراءتها بواسطة أجهزة الكمبيوتر ومشاركتها بين المكتبات، بحلول عام ١٩٧١، أصبحت تنسيقات مارك هي المعيار الوطني الأمريكي لنشر البيانات الببليوغرافية، وبعد عامين أصبح معياراً دولياً، هناك العديد من إصدارات مارك قيد الاستخدام في جميع أنحاء العالم، وأكثرها انتشارًا هو مارك ٢١، الذي تم إنشاؤه في عام ١٩٩٩ كنتيجة لمواءمة تنسيقات مارك الأمريكية والكندية، و UNIMARC تتم إدارة UNIMARC من قبل اللجنة الدائمة لUNIMARC التابعة للاتحاد الدولي لجمعيات ومؤسسات المكتبات (IFLA)، وتستخدم على نطاق واسع في أوروبا، تتضمن مجموعة معايير مارك ٢١ الآن تنسيقات للتسجيلات الاستنادية وتسجيلات المقتنيات وجداول التصنيف، بالإضافة إلى تنسيق التسجيلات الببليوغرافية.

تحدد معايير مارك ثلاثة جوانب من تسجيلة مارك: التعيينات الميدانية في كل تسجيلة، وهيكل التسجيلة، والمحتوى الفعلي للتسجيلة نفسها.

تم تصميم معيار مارك MARC لوصف مختلف مصادر المعلومات الورقية ونقل أو تحويل البيانات من نظام إلى نظام آخر، ويعتمد مارك على المواصفة الأمريكية الخاصة بتبادل البيانات الببليوجرافية Z39.2 وكذلك يعتمد على معيارالأيزو ISO 2709 الخاص بتبادل البيانات، وتوجد خمسة أنواع من البيانات لأشكال مارك ٢١ وهى:

- **البيانات الببليوجرافية**: لتشفير(وصف) أشكال البيانات الببليوجرافية في التسجيلات الخاصة بأوعية المعلومات، وتصف التسجيلات الببليوغرافية الخصائص الفكرية والمادية للمصادر الببليوغرافية (كتب وتسجيلات صوتية وتسجيلات فيديو وما إلى ذلك).
- **البيانات الاستنادية**: لتشفير البيانات الاستنادية المُجمعة في التسجيلات الاستنادية التي تم إنشاؤها للمساعدة في ضبط محتوى حقول التسجيلة التي تخضع للضبط الاستنادى، وتوفر التسجيلات الاستنادية معلومات حول الأسماء الفردية والموضوعات والألقاب الرسمية، يؤسس التسجيلة الاستنادية شكلاً معتمدًا لكل عنوان، مع مراجعة (إحالة) حسب الحاجة من الأشكال الأخرى للعنوان.
- **بيانات المقتنيات**: لتشفير عناصر البيانات في تسجيلات المقتنيات التي تُظهر المقتنيات وبيانات الموقع لأوعية المعلومات الموصوفة في التسجيلات، وخاصة بالنسخ عن مصدر المكتبة (رقم القاعة، موقع الرف، وحدات التخزين المحفوظة) وما إلى ذلك.
- **المعلومات المجتمعية**: لتشفير البيانات في التسجيلات التي تحتوى على معلومات عن الأحداث والبرامج والخدمات حتى يمكن تكامل هذه التسجيلات مع التسجيلات المجتمعية، وتسجل معلومات المجتمع سجلات مارك التي تصف وكالة تقدم الخدمة، مثل مأوى المشردين المحليين أو مزودي المساعدة الضريبية وغيرها.
- **بيانات التصنيف**: لتشفير عناصر البيانات المتعلقة بأرقام التصنيف ورؤوس الموضوعات المتصلة بها.

ويسجل التصنيف تسجيلات مارك التي تحتوي على بيانات التصنيف، على سبيل المثال، تم تشفير تصنيف مكتبة الكونجرس باستخدام تنسيق تصنيف مارك ٢١.

MARCXML هو مخطط XML يعتمد على معايير مارك ٢١ الموحد، تم تطوير MARCXML من قبل مكتبة الكونجرس وتبنته مع آخرين كوسيلة لتسهيل مشاركة المعلومات الببليوغرافية والوصول إليها عبر الشبكة العنكبوتية ويتسم بسهولة التحليل بواسطة أنظمة مختلفة باستخدامه كتنسيق تجميعي، كما هو الحال في حزم البرامج مثل MetaLib، وعلى الرغم من أن هذه الحزمة تم دمجها في مواصفات DTD أوسع، تضمنت أهداف تصميم MARCXML الأساسية:

بساطة المخطط، المرونة والقابلية للتوسع، التحويل بدون فقد أو خسارة وعكس للبيانات من مارك، عرض البيانات من خلال أوراق أنماط XML، يسجل مارك التحديثات وتحويلات البيانات من خلال تحويلات XML، وجود أدوات التحقق من صحة البيانات.

٣) معيار نقل وتشفير الميتاداتا (METS) Metadata Encoding and Transmission Standard[21]:

أُنشئ هذا المعيار لتغطية الحاجة إلى وجود معيار هيكل بيانات لوصف المواد الرقمية المعقدة، والـ (METS) عبارة عن مخطط تم بناءه على لغة الـ XML ويهدف إلى إنشاء وثائق XML تعبر عن بنية الموارد الموجودة بالمكتبات الرقمية، والميتاداتا الوصفية والإدارية والهيكلية المرتبطة بهذه الموارد، بالإضافة إلى أسماء ومواقع الملفات التي تتكون منها هذه الموارد الرقمية، ولغة الـ XML(eXtensible Markup Language) هي لغة ترميز الوثائق على الانترنت بغرض تعريف بمعلوماتها البنيوية وهى لغة تستخدم بواسطة الحاسوب لتعريف المعلومات المخفية عن بنية الوثيقة، إن الميتاداتا اللازمة لإدارة واستخدام الموارد الرقمية بنجاح تختلف وتبدو أكثر اتساعاً من تلك اللازمة لإدارة مجموعات الأعمال المطبوعة، فنجد أنه بالنسبة للموارد الرقمية فإن الميتاداتا الهيكلية ضرورية لضمان أن الملفات التي تمت رقمنتها بصورة منفصلة قد تم بناءها بصورة صحيحة كما هو الحال في صفحات من كتاب تمت رقمنته، لذلك نجد أن الميتاداتا الفنية ضرورية لمعرفة معلومات عن عملية الرقمنة لكي يتمكن الباحثون من معرفة ما إذا كانت النسخة الرقمية تمثل النسخة الورقية بصورة دقيقة، كما تبدو الحاجة للميتاداتا الفنية الأخرى لتسهيل عملية الانتقال الدوري للبيانات من نظام إلى آخر من أجل المحافظة على الموارد ذات القيمة العالية .

وتتكون وثيقة METS من سبعة أقسام رئيسية:

- **رأس METS** - يحتوي رأس METS على بيانات وصفية تصف وثيقة METS نفسها، بما في ذلك معلومات مثل المنشئ والمحرر وما إلى ذلك.
- **بيانات الوصف الوصفية** - يشير اسم بيانات الوصف الوصفية إلى البيانات الوصفية الخارجية لوثيقة METS (على سبيل المثال، سجل مارك في الأوباك أو أداة مساعدة العثور على EAD المحفوظة على خادم WWW)، أو يحتوي على بيانات وصفية مضمنة داخليًا، أو كليهما، قد يتم تضمين مثيلات متعددة من البيانات الوصفية الخارجية والداخلية في قسم البيانات الوصفية الوصفية.
- **البيانات الوصفية الإدارية** - يوفر قسم البيانات الوصفية الإدارية معلومات تتعلق بكيفية إنشاء الملفات وتخزينها، وحقوق الملكية الفكرية، والبيانات الوصفية المتعلقة بكائن المصدر الأصلي الذي اشتُق منه كائن المكتبة الرقمية، والمعلومات المتعلقة بمصدر الملفات التي يتألف منها كائن

المكتبة الرقمية (على سبيل المثال، علاقات الملفات الرئيسية / المشتقة، ومعلومات النقل / التحويل)، كما هو الحال مع البيانات الوصفية الوصفية، قد تكون البيانات الوصفية الإدارية إما خارجية بالنسبة لوثيقة METS أو مشفرة داخليًا.

- **قسم الملف** - يسرد قسم الملف جميع الملفات التي تحتوي على المحتوى الذي يشتمل على النسخ الإلكترونية للمورد الرقمي، يمكن تجميع عناصر <file> ضمن عناصر <fileGrp> لتوفير تقسيم الملفات حسب

إصدار المورد الرقمي.

- **الخريطة الهيكلية** - الخريطة الهيكلية هي قلب وثيقة METS، وهي تحدد الهيكل الهرمي لعنصر المكتبة الرقمية، وتربط عناصر تلك البنية بملفات المحتوى والبيانات الوصفية التي تتعلق بكل عنصر.
- **الروابط الهيكلية** - يسمح قسم الروابط الهيكلية في METS لمنشئي METS بتسجيل وجود الروابط التشعبية بين العقد في التسلسل الهرمي المبين في الخريطة الهيكلية، هذا له قيمة خاصة في استخدام METS لأرشفة مواقع الويب.
- **السلوك** - يمكن استخدام قسم السلوك لربط السلوكيات القابلة للتنفيذ بالمحتوى الموجود في كائن METS. يحتوي كل سلوك ضمن قسم السلوك على عنصر تعريف واجهة يمثل تعريفًا تجريديًا لمجموعة السلوكيات التي يمثلها قسم سلوك معين، يحتوي كل سلوك أيضًا على عنصر آلي يحدد وحدة نمطية من التعليمات البرمجية القابلة للتنفيذ التي تنفذ وتدير السلوكيات المحددة بشكل تجريدي من خلال تعريف الواجهة.

في المجمل يوفر مخطط METS آلية مرنة لتشفير بيانات الوصف الوصفية والإدارية والهيكلية لكائن مكتبة رقمية، وللتعبير عن الروابط المعقدة بين هذه الأشكال المختلفة من البيانات الوصفية، لذلك يمكن أن يوفر معيارًا مفيدًا لتبادل عناصر المكتبة الرقمية بين المستودعات الرقمية، بالإضافة إلى ذلك، يوفر METS القدرة على ربط كائن رقمي بالسلوكيات أو الخدمات، كما تم التوضيح سابقا والميزات الرئيسية للمخطط، ولكن من الضروري إجراء فحص شامل للمخطط والوثائق المضمنة فيه لفهم النطاق الكامل لإمكانياته.

٤) مخطط وصف كائن الميتاداتا Metadata Object Description Schema (MODS) :

وهو مخطط ميتاداتا وصفي مشتق من معيار مارك ٢١ ويهدف إلى حمل بيانات مختارة من تسجيلات موجودة لمارك ٢١ أو لتمكين إنشاء تسجيلات أصلية لعرض وصف موارد المعلومات وتستخدم اللغة بدلاً عن الأرقام المستخدمة في حقول مارك ٢١، وتستخدم لغة الـ XML في التعبير عن هذا المخطط، وتعتبر عناصر الـ (MODS) أكثر ثراءًا من عناصر الدبلن كور ولكنها أقل تعقيداً من عناصر مارك ٢١ ويركز هذا المعيار على الوصف التفصيلي للموارد الإلكترونية وهذا يمنحه ميزة على مخططات الميتاداتا الأخرى، وتعتبر عناصر الـ MODS أكثر ثراءاً من عناصر الدبلن كور، كما أن عناصره أكثر توافقاً مع البيانات الببليوجرافية المكتبية بالمقارنة مع الدبلن كور، كما انه أبسط في التطبيق من شكل مارك ٢١ الببليوجرافي وباستخدامه للغة الـ XML فإن الـ MODS يوفر تقنيات وإضافات أكثر من تلك المُعدة بالمارك، فمثلا يوفر استخدام صفة هوية "ID" اختيارية لتسهيل عملية الربط على مستوى العناصر وإمكانية تحديد اللغة وطريقة الكتابة.

٥) الوصف الأرشيفي المشفر The Encoded Archival Description EAD :

تم تطوير الوصف الأرشيفي المشفر EAD لترميز البيانات الموجودة بوسائل الإيجاد Finding aidsللوثائق الأرشيفية من أجل إمكانية البحث فيها وعرضها إلكترونياً، وتعتبر وسائل الإيجاد في الأرشيف والمجموعات الخاصة أداة مهمة لوصف المواد، وتختلف هذه الوسائل من تسجيلات الفهرس في أنها أطول وتحتوي على الكثير من التوضيحات، كما أنها مهيكلة بصورة عالية بطريقة هرمية وتبدأ هذه الوسائل بوصف المجموعة كاملة مع إيضاح نوع المواد التي تحتويها وسبب أهميتها، كما تقوم بتقديم سيرة ذاتية عن المؤلفين، كما توصف مجموعة السلاسل التي تنظم فيها المجموعات بالإضافة إلى قيامها بتقسيم محتويات الصناديق والملفات التي تتكون منها المجموعة الأرشيفية.

ويستخدم الـ EAD بصورة واسعة في المكتبات الأكاديمية والجمعيات التاريخية والمتاحف التي تحتوي على مجموعات خاصة ضخمة ونجد أن الكثير من هذه المجموعات يحتوي على مواد فريدة لا توجد في أماكن أخرى، وغالباً فإن مثل هذه المواد التي توجد ضمن مجموعة خاصة لا يتم فهرستها منفردة، كما هو الحال في مجموعات المكتبات التقليدية، وبإنشاء وسائل إيجاد مبنية على الـ EAD فإن المكتبات والمتاحف تزيد من عملية التوعية لمجتمع الانترنت بوجود هذه المجموعات الفريدة.

٦) مبادرة تشفير النص Text Encoding Initiative (TEI):

وهي مشروع دولي يهدف إلى تطوير موجهات أو أدلة لترميز النصوص الإلكترونية مثل الروايات والمسرحيات والشعر، والغرض الأساسي منه هو دعم عملية البحث في مجال العلوم الإنسانية، وبالإضافة إلى تحديد كيفية تشفير النص الخاص بعمل ما، فإن الموجهات الخاصة بالـ TEI تقوم بتحديد جزء خاص في رأس الوثيقة Header يتم تضمينه في المصدر المعين والذي يشتمل على الميتاداتا الخاصة بهذا العمل، وتتكون المعلومات الببليوجرافية الأساسية من عناصر شبيهة بتلك الموجودة بفهارس المكتبات، بالإضافة إلى وجود عناصر أخرى خاصة بتسجيل تفصيلات عن كيفية كتابة النص وتحديده، وكيفية أجراء عملية الترميز، وما هي المراجعات التي تمت بالإضافة إلى معلومات أخرى غير بيولوجرافية .

بالإضافة إلى هذه المعايير التي تم استعراضها أعلاه فهنالك العديد من معايير الميتاداتا المتخصصة في مجال موضوعي معين مثل معيار الISO 19115 والخاص بتوفير معلومات مكانية وزمانية للبيانات الجغرافية. والمعلوم أن مخططات ومعايير الميتاداتا المختلفة يمكن أن تستخدم لوصف مصدر معين من مصادر المعلومات ولخدمة مجموعات مختلفة من المستفدين.

وبسبب وجود هذا العدد الكبير من معايير ومخططات الميتاداتا فقد ظهرت الحاجة إلى وجود خاصية التشغيل البينيInteroperability للميتاداتا يمكن بواسطتها لنظم المعلومات المختلفة من تبادل الميتاداتا فيما بينها دون أن يتسبب ذلك التبادل في فقدان أي قدر من المعلومات المتبادلة، ويعتبر التشغيل البيني للميتاداتا Metadata interoperability أهم أسس تطبيق مخططات الميتاداتا باعتبار أنه يُسهل من عملية تبادل وتشاطر البيانات التي تم إعدادها وفقاً لمخططات ميتاداتا مختلفة، وبذلك يُمكّن التشغيل البيني للميتاداتا من إجراء عمليات بحث عبر النطاقات الموضوعيه والبرامجيه المختلفة. Cross Domain Searching.

جداول الميتاداتا Metadata Crosswalks

تتم عملية التشغيل البيني للميتاداتا بما يعرف بال Crosswalks وهي عباره عن جداول تتم بواسطتها عملية مقارنة وتحليل Mapping لإثنين أو أكثر من مخططات الميتاداتا لتمكين مجموعة من المستخدمين لاحد مخططات الميتاداتا من استخدام مخطط آخر، وبكلمات أخرى فإن الجدول

Crosswalkهو عبارة عن عملية تحويل transformation تم تطبيقها على مجموعة عناصر أحد مخططات الميتاداتا(ويسمى المصدر) لينتج عن هذه العملية حفظ تعديل مقابل لهذه العناصر في عناصر مخطط آخر من مخططات الميتاداتا (ويسمى الهدف).

وطبقاً للتعريف الصادر من الشبكة الكندية للمعلومات التراثية Canadian Heritage Information Network)[22] فإن جداول الميتاداتا تستخدم للتنقل بين مجموعات عناصر الميتاداتا المختلفة، حيث أن العناصر أو الحقول الموجودة في احد مخططات الميتاداتا يتم مقابلتها مع العناصر أو الحقول الخاصة بمخطط ميتاداتا آخر، والتي تحمل نفس المعنى أو معنى مشابه، وتسمى هذه العملية ايضاً بالمقابلة السيمنطيقية Semantic Mapping، وللتدليل على ذلك نضرب المثل التالي: لنفترض أن الأرشيف يحتوي على تسجيل مارك في الكتالوج الخاص به يصف مخطوطة، إذا قام الأرشيف بعمل نسخة رقمية من تلك المخطوطة وأراد عرضها على الويب مع المعلومات من الفهرس، فسيتعين عليه ترجمة البيانات من تسجيلة كتالوج مارك إلى تنسيق مختلف مثل مخطط وصف كائن البيانات الوصفية وهو للعرض في صفحة ويب، ونظرًا لأن مارك يحتوي على حقول مختلفة عن MODS، يجب اتخاذ قرارات حول مكان وضع البيانات في MODS، هذا النوع من "الترجمة" من تنسيق إلى آخر يسمى غالبًا "تعيين البيانات الوصفية" أو "تعيين الحقول"، ويرتبط بـ "تعيين البيانات" و "الخرائط الدلالية".

تمتلك Crosswalks أيضًا العديد من القدرات الفنية، أنها تساعد قواعد البيانات باستخدام مخططات البيانات الوصفية المختلفة لمشاركة المعلومات، كما أنها تساعد جامعي البيانات الوصفية في إنشاء الفهارس التعاونية. و تُمكن محركات البحث من البحث في قواعد بيانات متعددة في وقت واحد باستعلام واحد.

عملية المقابلة و التحويل Mapping

وهي المنهجية القائمة على القواعد والتي تدعم التعيين التلقائي لحقول البيانات بين مجموعات البيانات المختلفة في بيئة معالجة البيانات، إذا تعذر تعيين حقل أو مطابقته استنادًا إلى الاسم وحده (على سبيل المثال، تطابق متطابق)، فإن المنهجية تستخدم قواعد لتحديد نوع الحقل، بناءًا على اسم الحقل، ثم يتم استخدام نوع الحقل المحدد للمطابقة، يمكن استخدام المنهجية لمطابقة الحقول التي تبدو مختلفة صوتيًا ولكنها مناسبة للمطابقة، وللتدليل على ذلك، يتم تحديد القواعد في شكل: تعبير عادي = نوع، هنا يسرد عضو التعبير العادي السلاسل النصية أو السلاسل الفرعية للحقل، يتم ترتيب القواعد في تفضيل تنازلي وفقًا لاحتمال أن تحدد قاعدة معينة حقلًا بشكل صحيح، بهذه الطريقة تسمح المنهجية بمهمة تعيين الحقول من مجموعة بيانات إلى أخرى لتكون آلية بالكامل.[٢٣]

وهناك مجموعة من خصائص الميتاداتا يجب وضعها في الاعتبار عند إجراء عملية المقابلة والتحليل Crosswalkin Mapping طبقاً لوثيقة ال NISO (Content Metadata Standards (Issues in 1998 هذه الخصائص قد تحتوى على الآتي :[٢٤]

تعريف سيمنطيقى Semantic definition لكل عنصر من عناصر الميتاداتا.

ما اذا كان عنصر من العناصر الزامياً أم اختيارياً، أم انه الزامى طبقاً لشروط معينة.

ما اذا كان هناك عنصر ما قد يتكرر ظهوره بنفس التسجيله الموجودة.

التقييدات (المحددات) بسبب تنظيم عناصر الميتاداتا نسبياً مع بعضها البعض (مثلاً العلاقات الهرمية).

التقييدات المفروضة على قيمة عنصر ما (مثلاً نص حر – مدى رقمى معين – تاريخ – مفردة مقيدة)

طرق تنفيذ جداول الميتاداتا

هنالك طريقتان يتم استخدامهما في عملية إنشاء الجداول:

١. **الطريقة الأولى** : هي طريقة الجداول المطلقة Absolute Crosswalking وهذه الطريقة تتطلب تقابلاً متطابقاً (كاملا) بين العناصر أو الحقول في مخطط الميتاداتا الذي يمثل المصدر والمخطط الآخر الذي يمثل الهدف، وفي حالة عدم وجود حقول متطابقة بين المخططين لاتتم عملية إنشاء الجداول Crosswalks بين مخططات الميتاداتا، إن الجداول المطلقة تضمن تماثلية أو شبه تماثلية العناصر بين مخططات الميتاداتا ولكنها لاتعمل بصورة جيدة عند عملية تحويل البيانات Data conversion بين نظام وآخر، والمشكلة هنا تتمثل في أن قيم البيانات التي لايتوفر لها مقابل سوف يتم اسقاطها من الجدول، وخاصة عندما يكون مخطط الميتاداتا الذي يمثل المصدر يمتلك بنيه أغنى من تلك الخاصة بمخطط الميتاداتا الذي يمثل الهدف (كما هو الحال في إنشاء جدول للمقابلة من مارك إلى دبلن كور) كما في جدول.(١)

جدول رقم (١) جدول للمقابلة من مارك إلى دبلن كور

MARC field		Dublin Core element
210 Abbreviated Title	→	Title
222 Key Title	→	Title
240 Uniform Title	→	Title
242 Translated Title	→	Title
245 Title Statement	→	Title
246 Variant Title	→	Title

جدول رقم(٢) جدول للمقابلة من حقول MARC إلى دبلن كور

Dublin Core Elements	MARC Fields
Title	130, 240, 245, 246
Creator	100, 110, 111
Contributor	100, 110, 111, 700, 710, 711*
Subject / Keyword	600, 610, 630, 650, 651, 653
Description	Notes 500, 505, 520, 562, 583
Publisher	260 b
Relationship	581, 700 t, 730, 787, 776
Date	008/ 07-10 260 c

٢. **الطريقة الثانية** : وهي التي يتم استخدامها للتغلب على هذه المشكلة هي طريقة الجداول النسبية Relative Crosswalking والتي يتم بموجبها مقابلة عناصر مخطط ميتاداتا المصدر مع عنصر واحد على الأقل من عناصر مخطط ميتاداتا الهدف، بغض النظر عن كون العنصرين متعادلين من ناحية المعنى أم غير متعادلين. ومن الواضح أن هذه الطريقة البديلة تعمل بصورة أفضل في حالة إجراء المقابلة والتحليل Mapping من مخططات معقدة إلى أخرى بسيطة (مثلاً من المارك إلى الدبلن كور وليس العكس) كما في الجدول (٣).

جدول رقم (٣) جدول للمقابلة من دبلن كور إلى مارك

Dublin Core element		MARC field
Title	→	245 Title Statement
Title	→	245 Title Statement
Title	→	245 Title Statement
Title	→	245 Title Statement
Title	→	245 Title Statement
Title	→	245 Title Statement

تنشأ مشاكل المطابقة الأخرى عندما:

يحتوي مخطط واحد على عنصر واحد يجب تقسيمه مع وضع أجزاء مختلفة منه في عدة عناصر أخرى في المخطط الثاني (تعيين "واحد إلى متعدد")

يسمح أحد الأنظمة بتكرار عنصر أكثر من مرة بينما يسمح النظام الآخر فقط لهذا العنصر بالظهور مرة واحدة مع مصطلحات متعددة فيه.

المخططات لها تنسيقات بيانات مختلفة (مثل John Doe أو Doe، John)

عنصر في مخطط واحد مفهرس ولكن العنصر المكافئ في المخطط الآخر ليس كذلك.

قد تستخدم المخططات مفردات مختلفة مضبوطة.

المخططات تغير معاييرها بمرور الوقت.

بعض هذه المشاكل ببساطة غير قابلة للحل، كما تقول كارين كويل في " Crosswalking Citation Citation Metadata: The University of California's Experience"[٢٥] كلما زادت خبرة البيانات الوصفية التي نمتلكها، كلما أصبح من الواضح أن كمال البيانات الوصفية لا يمكن تحقيقه، وسيصاب أي شخص يحاول ذلك بخيبة أمل شديدة، عندما يتم عبور البيانات الوصفية بين مصدرين أو أكثر غير مرتبطين، ستكون هناك عناصر بيانات لا يمكن التوفيق بينها بطريقة مثالية، مفتاح النجاح في عبور البيانات الوصفية هو المرونة الذكية، من الضروري التركيز على الأهداف المهمة والاستعداد لتقديم تنازلات من أجل الوصول إلى نتيجة عملية للمشروعات".

تنسيقات ملفات ETDs:

جامعة Virginia Tech، أول جامعة تنشئ برنامج للرسائل والأطروحات الرقمية أوصت بقائمة من تنسيقات الملفات للرسائل والأطروحات الإلكترونية[٢٦] .

تنسيقات النصوص

١. **PDF (.pdf)**: يعد PDF اختصارًا لتنسيق المستندات المحمولة وهو التنسيق الأكثر استخدامًا لإرسال النص الكامل للرسائل والأطروحات الرقمية، يستخدم PDF لتقديم المستندات بطريقة مستقلة عن برامج التطبيقات والأجهزة وأنظمة التشغيل، وأفضل استخدام لملف PDF لتخزين الرسومات المتجهية (الجرافيك) (أي الرسومات المرسومة باستخدام الخطوط والمنحنيات بدلاً من وحدات البكسل).

٢. **(ppt) PPT :** يمكن أيضًا تضمين عرض تقديمي PPT أو Power-Point في الرسائل والأطروحات الرقمية

٣. **JPEG (.jpeg) أو .jpg :** تعد JPEG أو Joint Photographic Experts Group طريقة شائعة الاستخدام للضغط اللامع للصور الرقمية التي يتم إنتاجها بشكل خاص بواسطة التصوير الفوتوغرافي الرقمي، يمكن ضبط درجة الضغط مع الأخذ بعين الاعتبار حجم التخزين وجودة الصورة.

ملفات الصور:

١. **(gif.) GIF** : تنسيق تبادل الرسومات أو GIF هو تنسيق صورة نقطية تم تقديمه بواسطة CompuServ في عام ١٩٨٧، ومنذ ذلك الحين أصبح استخدامًا واسع النطاق على WWW نظرًا لدعمه الواسع وإمكانية نقله، من الأفضل استخدامه لتخزين صور جودة الشاشة التي لا تحتوي على العديد من الألوان، عادةً ما تكون ملفات GIF صغيرة جدًا، ولكن لا يمكنها إعادة إنتاج نطاق الألوان اللازم لإعادة إنتاج الصور الفوتوغرافية التي يمكن أن يقوم بها JPEG.

٢. **TIFF (.tif)** : TIFF، اختصار لـ Tagged Information File Format، هو تنسيق أرشيفي، مما يعني أنه لا يضحي بجودة الصورة من أجل تقليل أحجام الملفات، يتم دعم تنسيق TIFF على نطاق واسع عن طريق تطبيقات معالجة الصور، عن طريق النشر وتطبيقات تخطيط الصفحة، والمسح الضوئي، والفاكس، ومعالجة الكلمات، والتعرف الضوئي على الأحرف والتطبيقات الأخرى، وهو تطبيق ممتاز لتخزين صور مفصلة وعالية الجودة.

٣. **PNG (.png)** : رسومات الشبكة المحمولة (PNG) هي تنسيق ملف نقطي يدعم ضغط البيانات بشكل أقل، تم إنشاؤه كبديل محسّن وغير حاصل على براءة اختراع، لـ GIF وهو تنسيق ضغط الصور الأكثر استخدامًا على الإنترنت ميزة نسبية نظرًا لأنه يمكنه تخزين صور عالية الألوان، يمكن استخدامه أيضًا لتخزين محتوى الصور الفوتوغرافية.

٤. **MPEG (.mpeg)** : تعد مجموعة Moving Picture Experts Group أو MPEG هي أقدم وأشهر تنسيق مدعوم للأفلام، يتم استخدام تنسيق MPEG بشكل شائع كتنسيق إخراج من أدوات UNIX المساعدة التي تنشئ محتوى فيديو.

٥. **كويك تايم (.mov .qt)** : QuickTimeهو إطار عمل للوسائط المتعددة قابل للتوسيع طورته شركة Apple Inc، وهو قادر على التعامل مع تنسيقات مختلفة من الفيديو الرقمي والصورة والصوت والصور البانورامية والتفاعلية، دعم QuickTime جيد على نظامي Macintosh و Windows ولكن ليس كثيرًا على نظام UNIX.

ملفات الفيديو

١. **AVI (.avi)** :تحتوي ملفات(AVI) Audio Video Interleave على كلٍ من بيانات الصوت والفيديو في حاوية ملف تتيح تشغيل الصوت مع الفيديو المتزامن، وهو أكثر تنسيق مدعوم من Windows وغير مدعوم بشكل جيد في التنسيقات الأخرى.

٢. **WAV (.wav)** :تنسيق ملف الصوت الموجي أو WAVE هو معيار افتراضي لملفات صوت Windows، كما أنه مدعوم من منصات أخرى، يستخدم بشكل أساسي لملفات الصوت الخام وغير المضغوطة.

٣. **AIFF (.aif)** : تنسيق ملف تبادل الصوت (AIFF) هو معيار تنسيق ملف صوتي يُستخدم لتخزين بيانات الصوت لأجهزة الكمبيوتر الشخصية وأجهزة الصوت الإلكترونية الأخرى، تنسيق AIFF هو مكافئ خاص بماكنتوش لتنسيق WAV لا يتم دعمه بشكل جيد على جميع الأنظمة الأساسية مثل تنسيق WAV.

ملفات الصوت

١. **MPEG-3 (.mp3)** : يزيل تنسيق MPEG-3 (أو MP3) البيانات الصوتية التي لا تدركها الأذن والدماغ بشكل قوي، وبالتالي ينشئ ملفات بجودة معقولة أصغر بعشر مرات من البيانات الأولية نفسها، تعد ملفات MP3 جيدة لتخزين مقاطع طويلة من محتوى الصوت حيث لا تتطلب الجودة العالية.

تنسيقات ملفات أخرى

٢. **ملفات مضغوطة-.zip** : إنه تنسيق ملف أرشيف يدعم ضغط البيانات بدون فقدان بعضها، قد يحتوي ملف .zip على ملف أو مجلد واحد أو أكثر.

٣. **HTML Hyper** : لغة ترميز النص (HTML) هي اللغة التي تستخدم بها صفحات الويب وقد تم تصميمه بواسطتها، تم تعريف هذا المعيار بواسطة شبكة الويب العالمية كونسورتيوم (W3C) يسمح HTML بإنشاء مستندات الويب بسهولة، HTML هو ملف نصي عادي ويمكن استخدام أي محرر نصوص بسيط مثل Notepad لإنشاء مستندات HTML.

٤. **XML eXtensible :**لغة الترميز (XML) توفر تمثيل منظم للبيانات التي يمكن تنفيذها على نطاق واسع وسهلة النشر XML هي مجموعة فرعية من SGML (لغة الترميز المعيارية العامة)، تم تعديلها وتحسينها للتسليم عبر الويب، تم تعريف هذا المعيار من قبل اتحاد شبكة الويب العالمية (W3C)، ويمكن استخدام XML لتنسيق البيانات ونقلها بطريقة سهلة ومتسقة، XML هو أيضًا تنسيق نصي عادي.

المراجع

١) عمر حسن عبد الرحمن .التشغيل البيني للميتاداتا -. Cybrarian Journal -.ع ٢١، ديسمبر ٢٠٠٩ .- تاريخ الاطلاع <٢٠٢٠/٨/٢٢ -.>متاح في < :
http://journal.cybrarians.info/index.php?option=com_content&view=article&id=489:2011-08-13-21-17-13&catid=144:2009-05-20-09-53-29&Itemid=62 <

٢) صاوي، السيد صلاح. الميتاداتا وأهميتها في دعم الوصول إلى المحتوى الأرشيفي الرقمي : دراسة تطبيقية على مواقع الأرشيفات الوطنية على الإنترنت .- Cybrarians Journal.- ع ٤٢، يونيو ٢٠١٦ .- تاريخ الاطلاع <٢٠٢٠/٩/٢٢> .- متاح في: <
http://www.journal.cybrarians.info/index.php?option=com_content&view=article&id=733:ssawy&catid=290:studies <

٣) بن غيدة، وسام يوسف. المستودعات الرقمية المؤسساتية ودورها في إتاحة المحتويات الرقمية للمكتبات الجامعية الجزائرية على شبكة الانترنت.- Cybrarians Journal.- العدد ٤٥، مارس ٢٠١٧ .- تاريخ الاطلاع: <٢٠٢٠/٦/١٥> .- متاح في: <

http://www.journal.cybrarians.info/index.php?option=com_content&view=a
<rticle&id=773:wbinghida&catid=302:papers

٤) نشرتي، مؤمن. نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في أطار التنقيب عن البيانات: دراسة تحليلية. ١.- Cybrarians Journal.- ع ٣٤، مارس ٢٠١٤ .- تاريخ الاطلاع < ٢٠٢٠/٦/١١>.- متاح في:
http://www.journal.cybrarians.info/index.php?option=com_content&view=a
<rticle&id=765:mnasharty&catid=267:researches

٥) عمر حسن عبد الرحمن. التشغيل البيني للميتاداتا.- Cybrarian Journal .- ع ٢١، ديسمبر ٢٠٠٩ .- تاريخ الاطلاع <٢٠٢٠/٨/٢٢ >.- متاح في: >
http://journal.cybrarians.info/index.php?option=com_content&view=article
&id=489:2011-08-13-21-17-13&catid=144:2009-05-20-09-53-
> 29&Itemid=62

6) Potvin, S., & Thompson, S. (2016). An analysis of evolving metadata influences, standards, and practices in electronic theses and dissertations.Access date 15/7/2020._ Available at https://uh-ir.tdl.org/handle/10657/1341

7) () Mikeal, Adam; Brace, Tim; Leggett, John; McFarland, Mark; Phillips, Scott (2007). Developing a Common Submission System for ETDs in the Texas Digital Library. Available electronically from http : / /hdl .handle .net /1969 .1 /5679.

8) Eden, B., McCutcheon, S., Kreyche, M., Maurer, M. B., & Nickerson, J. (2008). Morphing metadata: maximizing access to electronic theses and dissertations. Library Hi Tech.Access date 11/7/2020.- Available at: https://www.emerald.com/insight/content/doi/10.1108/07378830810857799/full/html?casa_token=c--fv_HzgmEAAAAA:XFrgSieE7QrNTZqlCsMi0Mt6cVmVPa1ucWu5OuH4VTpLwil-

9) Access date18/7/2020.available at:: https://web.archive.org/web/20110629082434/http://www.bls.gov/ore/pdf/st000010.pdf

10) Hakimjavadi, H. and Noorman Masrek, M. (2013), "Evaluation of interoperability protocols in repositories of electronic theses and dissertations", Program: electronic library and information systems, Vol. 47 No. 1, pp. 34-59.access date18/7/2020.available at: https://doi.org/10.1108/00330331211296303

11) Goncalves, M.A., France, R.K., Fox, E.A. and Doszkocs, T.E. (2000), “MARIAN searching and querying across heterogeneous federated digital libraries”, Proceedings of the First DELOS Network of Excellence Work-

shop on Information Seeking, Searching and Querying in Digital Libraries, Zurich, Switzerland

12) Access date18/7/2020.available at:https://www.emerald.com/insight/content/doi/10.1108/00330331211296303/full/html#b8

13) Access date18/7/2020.available at: https://www.emerald.com/insight/content/doi/10.1108/00330331211296303/full/html#b9

14) Access date18/7/2020.available at:https://www.emerald.com/insight/content/doi/10.1108/00330331211296303/full/html#b44

15) Hakimjavadi, H. and Noorman Masrek, M. (2013), "Evaluation of interoperability protocols in repositories of electronic theses and dissertations", Program: electronic library and information systems, Vol. 47 No. 1, pp. 34-59. Access date18/7/2020.available at: https://doi.org/10.1108/00330331211296303

16) Hakimjavadi, H. and Noorman Masrek, M. (2013), "Evaluation of interoperability protocols in repositories of electronic theses and dissertations", Program: electronic library and information systems, Vol. 47 No. 1, pp. 34-59. Access date18/7/2020.available at: https://doi.org/10.1108/00330331211296303

17) Electronic Thesis Online (India)UGC (Submission of Metadata and Full-text of Doctoral Theses in Electronic Format) Regulations 2005.Retrieved 18/4/ 2016 from http://www.wvu.edu/~thesis/Presentations/ugc.pdf

18) ETD-MS: Interoperability Metadata Standard for Electronic Theses and

19) Dissertations.Retrieved 24/8/2019 from http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html

20) -Electronic Thesis Online (India)UGC (Submission of Metadata and Full-text of Doctoral Theses in Electronic Format) Regulations 2005.R

21) e(22) Electronic Thesis Online (India)UGC (Submission of Metadata and Full-text of Doctoral Theses in Electronic Format) Regulations 2005.Re df

22) ETD-MS: Interoperability Metadata Standard for Electronic Theses and

23) Dissertations. Retrieved 24/8/2019 from HYPERLINK "http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html" http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html s%20records.