الدراســـــــــات

المحقق الإلكتروني للنصوص العربية القديمة: الواقع والمأمول

أ.د. شريف كامل شاهين
أستاذ المكتبات والمعلومات ووكيل كلية الآداب
بجامعة القاهرة
sherifshn@yahoo.com

مستخلص:

قدمت التكنولوجيا على مدار السنين حلولاً عظيمة لمعظم المشكلات التي واجهت الإنسان؛ بغرض تيسير الأداء وتسهيله، وتوفير الوقت والسرعة والقضاء على بيروقراطية العمليات والإجراءات الورقية وتداخلها وأحيانًا كثيرة تكرارها دون حاجة تبرر ذلك.

ولم تكن المخطوطات والنصوص القديمة بعيدة عن دائرة الاهتمام والبحث والتطوير، فقد ظهرت أدوات جديدة متطورة لقراءة المخطوطات القديمة تمكّن من الحفاظ على النصوص القديمة وتحسينها رقميًّا.Preserved and digitally enhanced :

ولذلك طرح الباحثون تساؤلاً مهمًّا كان الدافع وراءه الاحتياج الفعلي للباحثين، وهو ماذا بعد الصورة الرقمية؟

والمتتبع للتقدم في مجال التصوير الرقمي للنصوص القديمة يلحظ الاتجاه نحو تكويد أو ترميز النصوص القديمة؛ من أجل تيسير بحثها ومقارنتها، ودراستها دراسة نقدية تكاملية من أجل تحقيقها وتقديمها للباحثين في قالبٍ معرفي متكاملٍ يراعي الأبعاد المختلفة: كاللغة والموضوع والمكان والزمان.

ولا يخفي على أحد ما تلعبه التكنولوجيا في مجالات اللسانيات الحاسوبية وتحليل النصوص وتنقيبها وإدارة وتحليل اللغة وأبعاد المكان والزمان، من خلال برمجيات لا تعرف أية قيودٍ في هذه الأبعاد.

ويُعد هذا التقدم تحديًا كبيرًا أمام التصوير الرقمي للنصوص، والدافع الأكبر وراء المزيد من الأبحاث والتجارب في مجال تمييز الحروف الحديثة والقديمة سواء المطبوعة أو المخطوطة باليد، أو حتى المقروءة صوتيًّا.

وعلى الجانب الآخر، منذ أن وضع برجستراسر وهارون والمنجد متونهم الذائعة الصيت في أصول نقد النصوص ونشرها وتحقيق النصوص وقواعد تحقيق المخطوطات في ذلك الزمن المبكر من القرن الماضي، وهو ما يمكن أن يكون عاملاً من بين العوامل التي أدت إلى وجود الفجوة الفكرية أو الثقافية بين القائمين على تحقيق النصوص ونشرها ورجال تكنولوجيا المعلومات والاتصالات.

إن الهدف من هذا البحث الوصفي التأصيلي لفكرة جديدة قد ترى النور على أرض الواقع وفي معامل المكتبات والمراكز والمؤسسات الثقافية "المحقق الإلكتروني" .

وتستهدف الدراسة المتأنية المتعمقة لمبررات واحتياجات هذه الأداة العصرية لمهمة ونشاط بحثي قديم، يُقصد بها أن تكون أداة لكل محقق للنصوص التراثية العربية.

وباختصار، هذه الدراسة تصميم منطقي متكامل لنظام معلومات داعم لمعظم مراحل وخطوات عملية تحقيق النصوص القديمة العربية، يستفيد من إمكانات تقنيات ترميز النصوص الإلكترونية، والربط فيما بينها لأغراضٍ مختلفة.

ويمكن لنظام المحقق الإلكتروني تقديم العون في الخطوات الأساسية في عملية التحقيق التي تبدأ بالتفكير في اختيار النص، مرورًا بتوثيقه، وآليات البحث عن نُسخه، وقراءة النص وتحريره ونقده ودراسته وخدمته. وقد أثبتت تكنولوجيا المعلومات والاتصالات أن لديها الحلول العملية السهلة لهذه الإجراءات، شريطة توافر النصوص في البيئة الرقمية. وتوجد أنظمة تربط بين النص الأصلي والنص المحقق، والنص المترجم إلى لغة أخرى. وتجدر الإشارة إلى تعاون الأكاديمية النمساوية للعلوم مع جامعة سالزبرج؛ من أجل إعداد برنامج لتحرير النصوص القديمة إلكترونيًّا، وُيعرف بـ -Classical Text Editor.

كلمات مفتاحية:

تحقيق النصوص – المحقق الإلكتروني – تكويد النصوص – النصوص الفائقة – مستودعات النصوص الإلكترونية – رقمنة النصوص – الصور الرقمية – التعرف الضوئي على الحروف – مبادرة ترميز النصوص TEI– الوصف الأرشيفي المرمز EAD.

القسم الأول: الإطار المنهجي

١ – ١ إرهاصات مشكلة البحث:

يشهد العالم طفرة في مجال تحليل النصوص الرقمية وأدواتها (Textual analysis tools)، خاصة المتاح منها على الشبكة العنكبوتية (الويب)، وتستفيض معظم المشروعات العالمية في المجال في استعراض الإجراءات والعلميات السابقة للتحليل، وتشمل عمليات تكويد أو ترميز النصوص Text Coding ، فضلاً عن اللغات المستخدمة لهذا الغرض والمعايير والمبادرات والمشروعات العالمية التي وضِعت من أجلها. وقد حرص الباحث في سياق الخطوات السابقة البحث عن اللغة العربية وحضورها الفاعل.

كانت البدايات من عام ١٩٤٩ حتى أوائل السبعينيات من القرن العشرين، حيث بدأ الكاهن اليسوعي الإيطالي الأب روبرتو بوسا Roberto Busa، ما يُعد حتى يومنا هذا مهمة ضخمة تمثلت في إعداد كشاف مفردات an index verborum يتضمن كافة الكلمات الواردة في أعمال سانت توماس الأكويني St Thomas Aquinas وغيره من المؤلفين المعاصرين، وقد بلغ مجموع الكلمات حوالي ١١ مليون كلمة تعود إلى لاتينية العصور الوسطى. والجدير بالذكر أنه في عام ١٩٩٢ ظهر قرص مليزر لأعمال سانت توماس الأكوينيSt Thomas Aquinas متضمنة بعض ملامح النصوص الفائقة Hypertextual features ملحقًا به دليل للمستخدم باللغات اللاتينية والإنجليزية والإيطالية. (Graham, 2013)

وفي خلال الفترة من السبعينيات حتى منتصف الثمانينيات، تضاعفت الجهود وأثمرت المزيد من النصوص الإلكترونية، وشهدت إصدار النسخة الثانية البريطانية من برمجيات تكوين فهارس الكلمات إلكترونيًّا COCOA concordance program التي تعمل على الحاسبات الضخمة. كما شهدت هذه الفترة التركيز على الحوسبة الأدبية واللغوية Literary and linguistic computingz، مع بعض التركيز على "اللغويات (اللسانيات) "Linguistic ، وما يُعرف بلغويات مجموعة كاملة من النصوص Corpus linguistics.

واستمرت التطورات الجديدة من منتصف الثمانينيات حتى أوائل التسعينيات، لتشهد التطورات المهمة في حوسبة العلوم الإنسانية. ويمكن أن يُعزى بعضها إلى تكنولوجيتين جديدتين هما: الحاسب الشخصي، والبريد الإلكتروني؛ فقد كان جهاز أبل ماكنتوش أكثر جاذبية للمستخدمين في العلوم الإنسانية؛ لسببين: الأول- كان لديها واجهة المستخدم الرسومية Graphical User Interfaceعلى أجهزة الحاسبات الشخصية، والثاني- يأتي جهاز ماكنتوش مع برنامج يسمح ببناء نصوصٍ فائقة Hypertexts بدائية بسهولة، يُعرف بالهيبركارد

.HyperCard

وفي عام ١٩٩١م، قدَّم باتريك كونر Patrick Conner منصة العمل Beowulf Workstation التي تسمح للمستخدم بإدخال النص مع الروابط Links، والشروح اللغوية والسياقية مختلفة الأنواع والترجمة أيضًا. وفي عام ١٩٨٦، ظهرت طريقة جديدة لترميز النصوص الإلكترونية على الساحة، وهي لغة الترميز العامة المعيارية (Standard Generalized Markup Language (SGML)، التي مهدت الطريق للخروج بمعيارٍ متفق عليه لترميز النصوص الإلكترونية Text Encoding Initiative (TEI)، ويُعرف بمبادرة ترميز النصوص. كما تتراوح عملية الرقمنة من الإدخال الآلي للنص عن طريق لوحة المفاتيح إلى البطاقات المثقبة إلى ترميز النصوص، إلى المسح الضوئي للوثائق، إلى ملفات الصور، والتسجيل الرقمي للصوت والصور المتحركة... إلخ. وأمام هذا التاريخ الطويل لنشأة وتطور التعامل الإلكتروني مع النصوص باللغات المختلفة، كان لا بد من السؤال: ماذا عن النصوص العربية القديمة المخطوطة باليد وأوائل المطبوعات؟ هل يمكن الاستفادة من هذا التاريخ الطويل للمعالجة الإلكترونية للنصوص باللغات المختلفة وانتشار المعاجم وتحليل النصوص ونسبتها لأفرادٍ بعينهم وأزمنة وحقب تاريخية محددة في عملية أو صنعة التحقيق للنصوص العربية القديمة؟

١ــ ٢ مشكلة البحث وأهميتها:

تكمن مشكلة البحث في افتقاد التراث الوثائقي العربي لمشروعٍ عصري يستهدف تقديم منظومة تعتمد على التكنولوجيات الحديثة لدعم ومساندة عملية فكرية (صنعة) تُمارَس حتى الآن بالطرق اليدوية التقليدية متجاهلة كافة أشكال التكنولوجيا، سواء ما يتصل منها بترميز المحتوى وربطه وإتاحة الوصول إليه بالأجهزة الإلكترونية المختلفة من أي مكان، وفي أي وقت، وهي تحقيق النصوص القديمة.

ولا يزال التحقيق يستهلك الوقت والجهد والمال من جانب الأفراد والمؤسسات؛ بحثًا عن النصوص والنسخ الأصلية والمترجمة وغيرها في مكتبات العالم، وتجميعها ومقارنتها والبحث في المصادر المرجعية وغيرها من الخطوات والمهام ذات الصلة. وهناك بعد آخر للمشكلة يكمن في محدودية ثقافة الرقمنة أو التحول الإلكتروني لمصادر التراث الوثائقي، واقتصارها على الصورة الرقمية أحيانًا، وأحيانًا أخرى الصورة الرقمية والميتاداتا، وذلك في أفضل الأحوال...، كل ذلك والعالم مشغول بالبحث عن إجاباتٍ مقنعة للسؤال: ماذا بعد الصورة الرقمية؟

فقد تبين للباحثين والمثقفين والمؤرخين محدودية الصورة الرقمية في مقابل الإتاحة الحرة

الإلكترونية لكافة مفردات النصوص القديمة، فضلاً عن قدرات البحث والوصول والربط والتحديد المكاني والزمني للنصوص ومفرداتها، وكل ما سبق يُعد من ركائز التحقيق الإلكتروني في البيئة الرقمية.

١ــ ٣ أهداف البحث:

هناك هدف عام لهذا البحث، وهو: الوصول للإطار العام للتصميم المنطقي لمنظومة "المحقق الإلكتروني" للنصوص العربية القديمة.

ولتحقيق هذا الهدف كان لا بد من وضع عدد من الأهداف الفرعية تتمثل في الآتي:

- تتبع تطور المعالجة الإلكترونية للنصوص في البيئة الرقمية.
- استكشاف تجارب وممارسات الغرب في هذا المجال في ما يتصل بتحقيق النصوص والمخطوطات اليونانية واللاتينية أو غيرها من اللغات الأجنبية القديمة والحديثة في البيئة الرقمية.
- حصر المبادرات والمعايير العالمية المتعلقة بتكويد وترميز النصوص في البيئة الرقمية وإمكانات كل منها، ودرجة الصلة والارتباط بالمخطوطات والنصوص القديمة على وجه الخصوص.
- التعرف على أوجه الاستفادة لعملية التحقيق Critical editing من تكنولوجيا المعلومات والاتصالات؛ بغرض تحليل النصوص وتدقيق الإصدارات والنسخ المختلفة للمخطوط الواحد، سواء بلغة واحدة أو أكثر من لغة في حال المخطوطات المترجمة. لقد قدمت التكنولوجيا الأدوات والتقنيات على مراحل متفاوتة وتم استخدامها في تطبيقاتٍ منفردة، ولم يتم تجميع كل ما سبق في إطارٍ وظيفي واحد؛ ولهذا السبب كان هذا البحث!

١ــ ٤ تساؤلات البحث:

هناك تساؤل عام لهذا البحث، وهو: ما مكونات أو عناصر الإطار العام للتصميم المنطقي المقترح لمنظومة "المحقق الإلكتروني" للنصوص العربية القديمة؟

وللإجابة عن هذا التساؤل العام، لا بد من الإجابة عن التساؤلات الفرعية الآتية:

- ما مراحل تطور المعالجة الإلكترونية للنصوص في البيئة الرقمية؟
- ما أبرز التجارب والممارسات في مجال تحقيق النصوص والمخطوطات اليونانية واللاتينية أو غيرها من اللغات الأجنبية القديمة والحديثة في البيئة الرقمية؟

- ما أبرز المبادرات والمعايير المتعلقة بتكويد وترميز النصوص في البيئة الرقمية وإمكانات كل منها؟
- ما درجة الصلة والارتباط بين المبادرات والمعايير العالمية لتكويد النصوص والمخطوطات والمصادر النصية القديمة على وجه الخصوص؟
- كيف يمكن لعملية التحقيق Critical editing أن تستفيد من تكنولوجيا المعلومات والاتصالات في هذا المجال؟

١ـ ٥ حدود البحث وأبعاده:

يهتم البحث بوضع مكوناتٍ أو عناصر الإطار العام للتصميم المنطقي المقترح لمنظومة "المحقق الإلكتروني" للنصوص العربية القديمة، وهو مجال علمي فرعي من أبرز مجالات الاهتمام تحت مظلة المجال المعرفي العريض "العلوم الإنسانية أو الإنسانيات الرقمية Digital Humanities". ولأغراض البحث تم استعراض الأدبيات والتجارب والمبادرات والمشروعات العالمية (دون أية حدودٍ لغوية أو مكانية أو زمنية) ذات العلاقة بالنصوص القديمة: تسجيلاً، وتحليلاً، ومقارنة، وغيرها من العمليات المطلوبة لمهمة تحقيق النصوص.

١ـ ٦ منهج البحث المستخدم وأدواته:

اعتمد البحث على المنهج الوصفي التحليلي الاستكشافي الذي استهدف كل من: الأدبيات المنشورة، والمشروعات والمبادرات العالمية والعربية (إن وجدت). البحث الاستكشافي يُسمى أحيانًا بالدراسة أو البحث الاستطلاعي الذي يهدف إلى تكوين رؤية سريعة أولية حول مشكلة معينة. ولا يحتاج هذا النوع من البحوث إلى استخدام عيناتٍ احتمالية كبيرة واستمارات رسمية. كما لا تتطلب نتيجة البحث عمقًا تحليليًّا، بل غالبًا ما تستهدف هذه الأنواع من البحوث تحقيق المزيد من التوضيح لموضوع أو ظاهرة أو مشكلة ما؛ بغرض تكوين خلفية معرفية تيسر على الباحثين الألفة بالموضوع وتطوراته.

وتهدف البحوث الاستكشافية إلى تحقيق ثلاثة أغراض تتمثل في: إشباع فضول الباحث ورغبته في الوصول إلى فهم أعمق للمشكلة أو الظاهرة محل البحث، وتطوير وسائل وطرق يمكن استخدامها في أية دراسات لاحقة. وأخيرًا، تحديد مدى جدوى القيام بأية دراساتٍ إضافية أخرى في المجال نفسه.

١ـ ٧ الأبحاث السابقة:

من أهم مصادر البحث المجلة العلمية الخاصة بمبادرة ترميز النصوص (TEI)، والتي تحرص على نشر الأبحاث العلمية في عدة تطبيقات ومجالات مختلفة تشمل فهرسة المخطوطات العربية المرقمنة. (TEI, The Journal of the Text Encoding Initiative, 2011 -)، ومن بين الأبحاث المنشورة في المجلة بحث تناول مبادرة TEI وفهرسة المخطوطات العربية الرقمية، والذي أكد أهمية الرقمنة في الكشف عن حقائق غير معلنة عن المخطوطات، شريطة عدم الاكتفاء بالصور الرقمية فقط. (Soualah & Hassoun, A TEI P5 Manuscript Description Adaptation for Cataloguing Digitized Arabic Manuscripts, 2012). كما تناول بحثٌ آخر تطبيق يستند إلى TEI لتحرير أوصاف المخطوطات. (Vertan & Reimers, 2012). وأخيرًا نشرت المجلة نفسها بحثًا عن سجل المخطوطات السلوفينية الحديثة المبكرة، يناقش البحث تفاصيل المخطوطات الحديثة المبكرة، وتشرح بنية السجل وترميزه (لا سيما تشفير البيانات الزمنية والجغرافية)، كما يُعرف البحث بالبوابة المبنية باستخدام برمجيات المستودعات الرقمية الشهيرة (فيدورا) التي تتيح للمستخدم تصفح المخطوطات والبحث عنها وتصدير بياناتها وفق تنسيق TEI، والتي تمكن من حصاد البيانات الوصفية. (Ogrin, Javoršek, & Erjavec, 2013)

كما كان هناك بحث علمي عن محرك بحث المخطوطات العربية القديمة يعتمد على الميتاداتا والشروح المُقدَّمة من لغة الترميز بـ XML. أكد الباحثون على أن النسخ التلقائي أو اليدوي Automatic or Manual Transcription للمخطوطات العربية القديمة أمر حتمي للتكشيف ونشر محتوياتها. فالطبيعة الخطية للكتابة العربية تُعد عائقًا أمام برامج التعرف الضوئي على الحروف (OCR). وعلى الجانب الآخر يبدو أن النسخ ببرنامج معالجة النصوص (مثلاً برنامج WORD)، أو من خلال التنسيق لغة تجهيز النصوص الإلكترونية HTML، ليس الحل المثالي؛ نظرًا لحقيقة عدم تنظيم أو وجود بنية للمحتوى. وفي هذا البحث، تم اقتراح محرك بحث للمخطوطات العربية القديمة، استنادًا إلى الميتاداتا وشروح XML، مما يسمح بإجراء عمليات البحث في قاعدة البيانات للوثائق المكتوبة بخط اليد والصور المكشفة استجابة لاستعلامات المستفيدين. (El Makhfi, El Bannay, Benslimane, & Rais, 2011)

ولفريق البحث نفسه بحثٌ علمي آخر عن نظام تكشيف وشرح، وبحث المخطوطات العربية القديمة. يوفر النظام المقترح القدرة على وضع المخطوطات العربية بتنسيق TEI XML، كما يشمل النموذج جميع البدائل المتعلقة بالتوصيف المعياري الدولي TEI (مبادرة تشفير النصوص)، مع ضمان التوافق مع الأدوات الحالية. ويوفر النظام قاعدة بيانات المنصة مفتوحة

المصدر المعروفة بـ(SDX (XML Documentary System، والتي تسمح بالبحث والتصفح وعرض هذه الوثائق وفق تنسيق الويب. هذا وقد تم اختبار النظام ونجاحه باستخدام بعض المخطوطات العربية القديمة. (El Makhfi, El Bannay, Benslimane, & Rais, System of indexing, annotation and search in the old Arabic manuscripts, 2011)

وفي ٢٠١٧م، قام مجموعة من الباحثين بدراسة تستهدف تحقيق المخطوطات العربية في العهد الجديد من خلال قاعدة ترميز النصوص وفق معيار TEI؛ بهدف تعزيز قراءتها والتعرف على مضمونها عن قرب من خلال مأتى جديد يُعرف بالعلوم الإنسانية الرقمية. (Clivaz, Schulthess, & Sankar, 2017)

تم تطبيق تقنية التصوير المستخدمة للمسح ثلاثي الأبعاد للمخ على مخطوطات العهد القديم القديمة، مما ينبئ باكتشاف أداة قيمة في الحفاظ على المخطوطات الثمينة وفتحها. (Mceachen, 2016)

وعلى جانب التصوير الرقمي للمخطوطات - في السنوات الأخيرة – تحول العاملون في مجال صيانة وحفظ المخطوطات إلى الأدوات الرقمية للحفاظ عليها. يتسم التصوير الرقمي بالعديد من المزايا، حيث تساعد على الحفاظ على الأصول الحساسة للتراث الوثائقي وإيجاد البديل الرقمي لها، حيث تسمح رقمنة الكتب القديمة لعددٍ أكبر من الناس بقراءتها أكثر مما لو تُركِت على حالتها المادية الأصلية. (Lewis, 2016)

وتُعد جامعة كوبنهاجن في الدنمارك موطنًا لمجموعة فريدة من مخطوطات البردي المصرية القديمة (يتجاوز عمرها ٣٥٠٠ سنة)، حيث يتم الاحتفاظ بالمجموعة الكبيرة من الوثائق من مصر القديمة في كوبنهاجن التي تكشف عن تفاصيل جديدة في العلوم الطبية في العصور القديمة، ويعكف فريقٌ دولي من الباحثين على ترجمة النصوص غير المستكشفة سابقًا، والتي - وفقًا لتصريح أحد الباحثين - تحتوي على رؤى جديدة ومثيرة في مصر القديمة. (Brix, 2018)

١ــ ٨ مصطلحات البحث:

تحقيق النصوص Text Critical editing: وتبدأ باختيار النص المناسب للتحقيق والدراسة، ثم توثيق العنوان والمؤلف والنسبة بينهما. ويُنصح بعد ذلك بالتأكد من مدى الحاجة إلى تحقيقه، وهل نُشر أو لا؟ مطبوعًا أو إلكترونيًّا؟ ثم تأتي مرحلة جمع النُّسخ الخطيَّة واستيفائها، ثم ترتيبها حسب منازلها بالمعايير المنهجية المنضبطة. وهنا تتجلى مهمة المحقق وعملية التحقيق في مرحلة تُعرف بالنَّسْخ والمقابلة، حيث يجتهد المحقق في نسْخ النسخة المعتمدة ومقابلة بقية النُّسخ

عليها وإثبات الفروق بينها، وذِكْر الترجيحات والتعليل بين القراءات في حال منهج التلفيق بين النُّسخ. لينتقل بعدها إلى مرحلة ضبط النَّص بنيةً ونحوًا وعروضًا وإملاءً، وضبط ما يشكل منه، وتقسيم النص إلى فقراتٍ وجملٍ ووضع علامات الترقيم وغير ذلك. وينتقل بعدها إلى مرحلة التعليق على النص فيما يحتاج إلى تعليق من قضايا ومسائل حسب رؤية المحقق وطبيعة النَّص ومراد المؤلَّف والفئة المستفيدة. وأخيرًا، يصل المحقق إلى مرحلة التكشيف وإعداد الكشافات بغرض تيسير الوصول للمحتوى.

تحليل النص Text Analysis: استخدام الحاسب وبرمجياته في استخراج الرؤى والقيمة من قواعد النصوص الكاملة، سواء أكانت من نوع المهيكلة أو البنائية أم لا. كما تعتمد منصات تحليل النص عادةً على منهجٍ واحدٍ من المنهجين الأساسيين التاليين: المنهج القائم على الرياضيات (يركز على الذكاء الاصطناعى، والتعلم الآلي)، والمنهج الثاني القائم على اللغة (يركز على التحليل الدلالي، والمحتوى، والفهم المشابه للإنسان).

المحتوى غير المهيكل أو غير البنائي Unstructured Content: محتوى مخصص للاستخدام الإنساني البشري، يُقدَّم عادة على هيئة جمل وعبارات وفقرات. ويُعد المحتوى غير المهيكل أكثر صعوبة للحاسبات عند تطبيق برمجيات الفهم والتحليل مقارنة بالمحتوى المهيكل أو البنائي، حيث إن اللغة البشرية الطبيعية يكتنفها الغموض ونقص الوضوح والتكرار وعدم الدقة؛ ولذلك ينصح بتطبيق معالج اللغة الطبيعية (NLP) على المحتوى غير المهيكل؛ بغرض التحليل الآلي له.

المجموعة الكاملة/ مدونة النصوص Corpus: هو مجموعة من الوثائق أو النصوص ذات الصلة، ومن الأمثلة على ذلك ويكيبيديا وأعمال شكسبير التي تم جمعها. وهو أيضًا مجموعة إلكترونية تضم ملايين الكلمات التي يمكن دراستها؛ للتعرف على اتجاهات لغة الكتابة والنشر.

مبادرة تشفير النصوص (TEI): معيار لتمثيل المواد النصية في شكلٍ رقمي، من خلال وسائل ترميز النصوص. هذا المعيار هو المنتج التعاوني لمجتمعٍ من العلماء، في مجالات العلوم الإنسانية والعلوم الاجتماعية واللغويات، يعملون تحت مظلة تحالف TEI.

القسم الثاني: التحقيق الإلكتروني للنصوص القديمة: المفاهيم والتطبيقات

٢-١ تكويد النصوص Text Encoding:

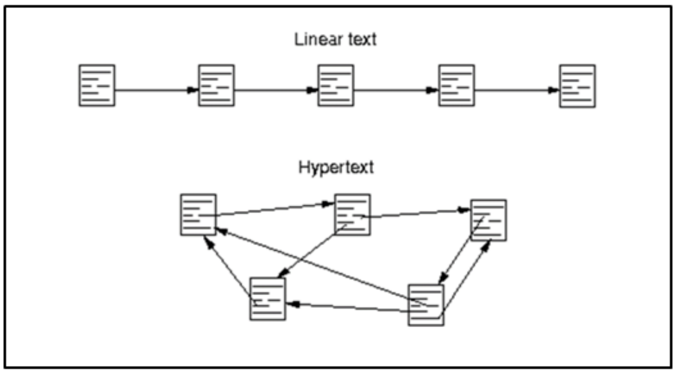
إن عملية تكويد أو تجهيز أو إعداد Markup أو ترميز أو تشفير الوثيقة النصية المتضمنة معلومات لن تمس أو تتصل أو حتى تؤثر على مضمون أو محتوى الوثيقة نفسها. وبمعنى آخر لا

يمثل **"التكويد"** جزءًا رئيسًا ضمن محتوى أو مضمون الوثيقة التي يتم تكويده، وإنما هو بمثابة هيكل بناء الوثيقة ويمتد أحيانًا ليشمل البنية النحوية للكلمات والجمل داخل الوثيقة، وقد يحدد في أحيانٍ أخرى طريقة عرض أجزاء الوثيقة على شاشة العرض المرئي أو في الصفحة المطبوعة. وتتضمن معلومات التكويد أو الترميز أو التجهيز مجموعة من التيجان Tags ومجموعة أخرى من أكواد الشكل المادي Format code المستخدمة داخل النص. وللاطلاع على التيجان المستخدمة في نص إلكتروني ما، يمكنك اختيار View sourceلأي صفحة من الصفحات الإلكترونية المعروفة على الإنترنت.

٢-٢ لغة تجهيز النص الالكتروني Markup language:

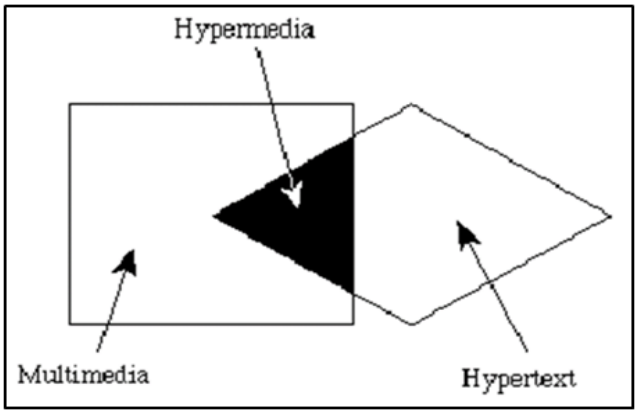
مجموعة من الواصفات Descriptors سابقة التحديد أو المحددة تحديدًا مسبقًا، ويُقصد بالواصفات مجموعة من الرموز والتيجان (Symbols & Tags). كما أن لغة تجهيز النص الإلكتروني بمثابة طريقة لتعريف أو تحديد الواصفات المستخدمة ضمن المعلومات الوصفية الخارجية أو المخبأة أو غير الظاهرة في الوثيقة النصية الإلكترونية، والتي تُستخدَم عادة في تحديد شكل وإخراج الوثيقة، وكذلك لعمليات التحليل الإحصائي للنصوص الإلكترونية. صممت لغات تجهيز النصوص الالكترونية للاستخدام مع برامج خاصة. وفي عام ١٩٨٦ تم ابتكار أو اختراع لغة معيارية مقننة، وهي لغة SGMLوهي اختصار لـ Standard Generalized Markup Language التي تم اعتمادها كمعيارٍ عالمي للغات لتجهيز النصوص الإلكترونية. كما تم اشتقاق لغة HTML من هذه اللغة المعيارية العامة لإنشاء أو خَلْق نصوص أو صفحات إلكترونية على شبكة الإنترنت.

وفي عام ١٩٩٨، قدم تحالف World Wide Web Consortium (W3C) نسخة مختصرة ومبسطة من اللغة المعيارية العامة لتجهيز النصوص SGML لغة أخرى عُرفت بـ XML (Xtensible Markup Language)، وهي لغة تتسم بالمرونة وعدم التقيد بأنظمة معينة للتشغيل، وشاع استخدامها في مجال المكتبات والمعلومات. وأهم ما يميز هذه اللغة: مرونة التعامل مع عناصر البيانات في النص الإلكتروني، وتحويل النص إلى ما يقابله في قواعد البيانات Databases. ومن هنا انطلقت مصطلحات وأوصاف، مثل: النص المترابط - النص الفائق – النص التشعبي – Hyper text. ويوضح الشكل(١) الاختلاف بين النص الخطي التقليدي والنص الفائق التشعبي في البيئة الإلكترونية.



شكل(١) النص الخطي التقليدي والنص الفائق التشعبي في البيئة الإلكترونية (htt)

كما يستخدم تكويد النصوص كطريقة لتمثيل المعلومات الرقمية، تسمح بالربط أو الوصل بين الملفات وعناصر البيانات، من خلال مجموعة من الروابط أو الوصلات الإلكترونية، كما توفر إمكانية استعراض وتصفح النص بطريقة غير خطية Non liner أو غير تتابعية تُعرف بعشوائية التصفح، حيث توضع الروابط أو الوصلات الكامنة داخل النص الفائق وكذلك الـ Iconsفي الوثيقة المكتوبة بلغة تجهيز النص الفائق بطريقة تيسر على العقل البشري تصفح المعلومات بالطريقة أو بالشكل الذي يرغب فيه القارئ، حيث يختار القارئ، من خلال لوحة المفاتيح أو الفأرة Mouse - Keyboardأو حتى بلمس الشاشة Touch، المعلومة التي يرغب في قراءتها بصرف النظر عن موقع تلك المعلومة. وهو ما يمكن استثماره للربط بين مفردات بعينها وردت في متن مخطوطة ما بما يقابلها من معانى داخل معجم للمصطلحات أو معجم البلدان للمواقع والمدن الجغرافية، أو ما يقابلها من أسماء للشخصيات في كتب التراجم أو غيرها. كما أن تقنية الربط تعدت فكرة الارتباط بالنوع الواحد، أي النصوص فقط Text only، وامتدت وتخطت الأنواع المختلفة للوسائط، وشاع مصطلح **الوسائط الفائقة Hypermedia**؛ للتعبير عن إمكانية الربط باستخدام الروابط الفائقة إلكترونيًا بين مجموعة متنوعة من المصادر منها المخطوطة والخريطة المصورة والمجسمات ثلاثية الأبعاد والتسجيلة السمعية ومقاطع الفيديو ... وغيرها. ويوضح الشكل (٢) مفهوم تكامل الوسائط بالرغم من اختلاف أنواعها لخدمة القارئ.



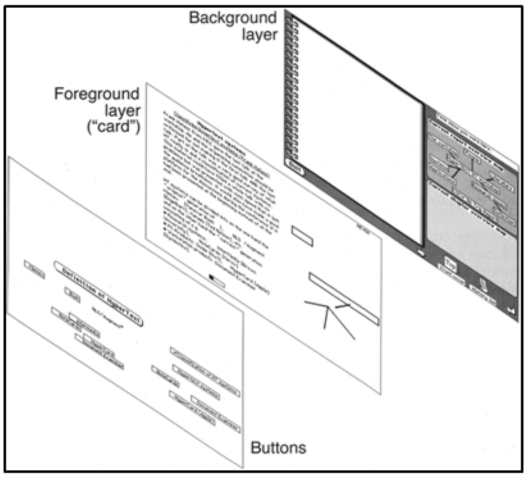
شكل(٢) تكامل الوسائط بالرغم من اختلاف أنواعها لخدمة القارئ (htt1)

ويكمن وراء هذه التطورات تقنية الروابط الفائقة Hyperlinks، والتي تتبع عدة تقسيمات وفق أسس مختلفة ومتنوعة، نذكر منها: (htt5) (htt4)(htt3)(htt2).

- **التقسيم وفق بيئة عمل الرابطة:** تنقسم إلى الروابط الداخلية، والروابط الخارجية.
- **التقسيم وفق اتجاه الرابطة:** تنقسم إلى الروابط الصاعدة، والروابط الهابطة.
- **التقسيم وفق نوع الوسيط:** تنقسم إلى روابط الصور، وروابط الفيديو.

٢-٣ الطبقات المكونة للنصوص الفائقة Hypertexts' Layers:

هناك مَن ينظر ويشرح النص الفائق على أنه يتألف من ثلاث طبقات. ويوضح الشكل (٣) طبقة المفاتيح الظاهرة للمستخدم، من خلال مجموعة من المفاتيح المرئية، والتي تدل على وجود رابطة Link لمحتوى آخر. وطبقة المقدمة المتضمنة للمحتوى المُقدَّم المعروض أمام المستخدم، وهو الأساس الذي من أجله تم تصميم الموقع أو الصفحات الإلكترونية لإتاحته. وأخيرًا نجد الطبقة الخلفية التي تتضمن المحتويات المكملة للعناصر الواردة في المحتوى الأساس، والتي تم التعبير عنها بمجموعة من المفاتيح المرئية أو الظاهرة للمستخدم. (htt7)(htt6)



شكل (٣) الطبقات الثلاث المكونة لبنية النص الفائق

٢-٤ مبادرات تكويد (ترميز) النصوص Text Encoding Initiatives:

التكويد عبارة عن تمثيل النص كما يظهر في الأصل على الصفحة، وبالتالي يعتمد على الكلمات المفتاحية للبحث. وتكمن أسباب اللجوء لتكويد النص في أهمية بحث النص بأجزاء منه، مثل: أوله، وآخره، أو بهامشه. إلا أن عملية التكويد تستلزم إلمام القائم بها بمحتوى النص المُستهدَف، فضلاً عن فهمه للإخراج والخطوط وبعض الأمور الأخرى.

٢-٤-١ الوصف الأرشيفي المرمز (المكود) EAD: Encoded Archival Description

هو معيارٌ قائمٌ على لغة تجهيز النصوص الإلكترونية XML يستهدف تشفير أدوات البحث عن الأرشيفات أو تكويدها، قام بالتصميم والتطوير اللجنة الفنية الفرعية لمعايير الأرشيف المشفر التابعة للجمعية الأمريكية للأرشيف، بالشراكة مع مكتبة الكونجرس. تم تطويره عام ١٩٩٣ في جامعة بيركلي؛ ليكون أداة مشتركة بين المجتمع الأرشيفي لنشر البيانات البحثية على الويب. (Congress, 2019)

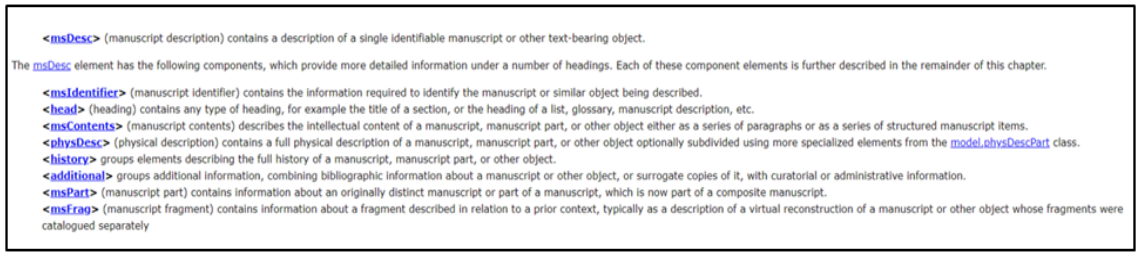
وهو معيارٌ يُستخدَم دوليًا في عددٍ متزايدٍ من الأرشيفات ومكتبات المخطوطات؛ لترميز البيانات التي تصف سجلات الهيئات والأوراق الشخصية. (Pitti, 1999)

٢-٤-٢ مبادرة ترميز النصوص TEI: Text Encoding Initiative

مشروع دولي بدأ عام ١٩٨٨؛ من أجل وضع مبادئ توجيهية لإعداد وتبادل النصوص الإلكترونية للأغراض البحثية، والوفاء باحتياجات ومتطلبات مجموعة واسعة من الاستخدامات من قِبل المهن اللغوية بشكلٍ عام، حيث يزداد الاحتياج إلى ممارسات الترميز الموحدة نتيجة زيادة استخدام كميات هائلة من النص الإلكتروني في مجالات البحث والصناعة، ولاسيما لمعالجة اللغة الطبيعية. (Ide) وفي يناير ١٩٩٤، قدم فريق Tel مجموعة من القواعد الإرشادية المتعلقة

بترميز أو تشفير النصوص القابلة للقراءة آليًّا أو تبادلها، والتي تُعد بمثابة مواثيق موحدة لتشفير مجموعة كبيرة من أنواع النصوص ومجموعة واسعة من التطبيقات الخاصة بها.

وتجدر الإشارة إلى أهم أقسام هذه المبادرة، وهو القسم العاشر المتعلق بوصف المخطوطات (الإصدار ٣.١.٠ المحدث في ديسمبر ٢٠١٦م، والذي خضع للتحديث مؤخرًا في ١٦ يوليو ٢٠١٩ م في إصدارة جديدة ٣.٦.٠) (TEI, 2019)، وتتألف محتويات الجزء الخاص بوصف المخطوطات من الأقسام الآتية: شكل (٤) محدد المخطوطة - رأس المخطوطة (البداية) ووصفها – المحتوى الفكري والأجزاء – الوصف المادي- التاريخ والنشأة والتطور– معلومات إضافية – أجزاء المخطوط – قطع المخطوطة.



شكل(٤) أقسام وصف المخطوطات ضمن مبادرة ترميز النصوص TEI: Text Encoding Initiative

وبالتالي يشمل الترميز أو التكويد الرقمي للمخطوطة البناء والمتن وعلاقاته الداخلية والخارجية والمسئولية والإيضاحات والأختام وغيرها. كما يشمل التكويد إلى جانب المتن، الهوامش أيضًا، وهو ما شجع الباحثين على التفكير في مبررات وضعها بهذه الطريقة، وهو ما يميزها عن المخطوطات المصورة رقميًّا فقط. ويمكن لهذه الطريقة إنشاء قاعدة بيانات نصية تكفل مرونة بحث النصوص. كما تسمح قواعد معيار التكويد TEI للمكودين بتحديد درجات أهمية لعناصر أو أجزاء الوثائق المستهدفة، ويتجلى أهمية التكويد عند التعامل مع النسخ المختلفة للمخطوطة، وهوامشها، والملاحظات المدوَّنة بها وغيرها من العناصر الأخرى.((Estill, 2016

ويرى بعض الباحثين أن المعيار TEI-ms يفتقر إلى القدرة على الوصف الشامل للمخطوطات العربية القديمة. ومن ثم تم اقتراح بعض التعديلات عليه؛ للتغلب على المشاكل المختلفة للفهرسة، ومن بين التعديلات المقترحة: تقديم الحروف العربية في المعيار، وتكامل أنماط الفهرسة المختلفة، وتقديم الحلول لوصف الأسماء العربية القديمة. ومن المؤكد أن مظاهر عدم كفاية المعيار لوصف المخطوطات العربية تمتد للمعيار الثاني EAD الذي تم إعداده خصيصًا للمجموعات الأرشيفية.

٢-٥ رقمنة (التحويل الرقمي) للنصوص القديمة: مخطوطات ووثائق وأرشيفات وغيرها

في سبتمبر ٢٠١٤، أصدر قسم الإفلا للكتب النادرة والمخطوطات "إرشادات للتخطيط لرقمنة الكتب النادرة والمخطوطات". (المكتبات، ٢٠١٤)

وفي يناير ٢٠١٥، تمت مُراجعة هذه الإرشادات لتغيير اسم الجهة المصدرة له من قسم الكتب النادرة والمخطوطات إلى قسم الكتب النادرة والمجموعات الخاصة. وقد أكدت وثيقة الإرشادات على أهمية الرقمنة ودورها في زيادة إمكانية اكتشاف واستخدام المقتنيات النادرة والخاصة مقارنة بالمجموعات العامة المتاحة للجمهور، ولولاها لظلت المجموعات النادرة والخاصة مُخبأة وغير مُستخدمة.

كما تهدف هذه الإرشادات إلى التنبؤ باحتياجات المستخدمين الذين يقومون الآن بالبحث رقميًّا، والبحث الذي يحتاج إلى تحليل النصوص الكاملة والبيانات الضخمة. وتشير الإرشادات إلى ميل رقمنة المجموعات النادرة إلى التعقيد؛ لأنه من المهم أخذ الجوانب المادية والفكرية والعلاقة بينها، بين أجزاء الكتاب ومجموعة المقتنيات ككل في الاعتبار. وعادةً ما يشمل الحفاظ على هذه العلاقة واصفات بيانات خاصة. وقد أكدت الإرشادات على أربعة أنواع للميتاداتا (الببليوجرافية/ الوصفية – الهيكلية – الصور/ الفنية – الإدارية)؛ من أجل إتاحة المجموعات/ المقتنيات المادية وتسهيل الوصول اليها. كما أشارت الوثيقة إلى قيام أمناء المكتبات بخلق نماذج ميتاداتا مُحدثة تقوم ليس فقط بإتاحة المواد الرقمية، ولكن أيضًا توفير المعلومات المطلوبة للحفظ طويل المدى، وكذلك لتسهيل الوصول إلى المواد من خلال أنظمة الشبكات الاستكشافية (المكتبات، ٢٠١٤). وهكذا يتبين أن فكرة تكويد النصوص الفائقة في مجال النصوص القديمة والمخطوطات والوثائق القديمة لا ذِكْر لها من قريبٍ أو بعيد.

٢-٥-١ المخطوطات الأصلية والبدائل الرقمية والمشتقات:

شاع في الماضي ولعدة سنوات إيجاد البدائل الرقمية Surrogates والمناسب من المشتقات Derivatives، وأحيانًا يُسمى بمسار الكيانات الرقمية أو كيانات المعلومات الرقمية Digital information Objects، ويُعرف الكيان المعلوماتي Information Object بأنه مفردة واحدة أو مجموعة مفردات من المعلومات الموجهة للإنسان، وتتم معالجتها إما من جانب البشر أو النظم كوحدة واحدة منفصلة ومستقلة بذاتها. وبصفة عامة، وبصرف النظر عن الشكل المادي أو الفكري الذي تتخذه، فإن هناك ثلاثة معالم أساسية تميزها، وهي:

- المضمون أو المحتوى Content: يشير إلى محتويات ومضمون الكيان.
- السياق Context: يشير إلي الإجابات المتصلة بالاستفسارات المتعلقة بمَن؟ وكيف؟

ولماذا؟ وأين؟ ومتى؟ عن كافة الجوانب المرتبطة بإنشاء الكيان المعلوماتي.

- البنية Structure: تتصلب العناصر والمجموعات الفرعية للكيان المعلوماتي وترابطها معًا، وربطها بعناصر أخرى تنتمي لكيانات معلوماتٍ مختلفة .

وقد ارتبطت معايير الميتاداتا Metadata (واصفات البيانات) بوصف وتحديد هوية وملامح الكيان المعلوماتي Information Objectالمتاح على الشبكة العنكبوتية.

كما ارتبط مسار رقمنة المخطوطات باختيار الوسيط الإلكتروني المناسب، حيث شهدت الوسائط المستخدمة لاختزان البيانات والمعلومات في البيئة الرقمية، على اختلاف أنواعها وأحجامها عبر نصف قرنٍ من الزمان، تطورًا ملحوظًا نتيجة للعوامل الآتية:

١. التقنية المستخدمة للتسجيل والقراءة (الثقوب – المغناطيسية – الليزر...).
٢. نوع المادة الخام المُستخدَمة لإنتاج الوسيط.
٣. السعة التخزينية للوسيط وسهولة الارتقاء بها.
٤. نوع المحتوى (أنواع البيانات والمعلومات المخزنة، من بياناتٍ ونصوصٍ وتسجيلاتٍ صوتية وصور وأفلام ومحتوى متعدد الأبعاد... إلخ).
٥. سهولة التداول والانتقال وتصغير حجم الوسيط (٢٤/٧) + التوافق مع الأجهزة المختلفة (حاسبات مكتبية ومحمولة ولوحية + أجهزة قراءة إلكترونية + هواتف ذكية... إلخ).
٦. القدرة على البقاء والتحمل وحفظ المضمون لأعمار طويلة.
٧. طبيعة الجمهور المُستهدَف بالمحتوى المنشور.
٨. البعد الاقتصادي للإنتاج والنشر والتوزيع.
٩. الوقت المستغرق للإنتاج والتوزيع.

والراصد لهذا المسار يمكنه تبين الفجوة المعرفية بين خبراء التراث الثقافي والباحثين وخبراء تكنولوجيا المعلومات والحاسب الإلكتروني، وخصوصًا في البلدان النامية. وقد ضاعف التقدم السريع للوسائل والوسائط الإلكترونية من تعقيد العملية، حيث أصبح خبراء التراث الثقافي مطالبين بتحديث معارفهم؛ من أجل التعرف على الإنجازات الجديدة، لأجل إعداد خطة حفظ وتطوير سياسات أكثر فعالية لحفظ "**التراث الرقمي**". (Quintero, 2003)(Letellier, 2007)

وإلى جانب التطورات العالمية في مجال وسائط الحفظ الرقمي، شهد العالم تطورًا آخر لا يقل أهمية وهو ما يُعرف "**بالتقارب الرقمي والتكامل المعرفي**"، وهو ناتج تكامل مكونات التكنولوجيا الرقمية الأساسية وخصائصها، مثل: الصوت والنصوص والفيديو والصور والبث والعروض

وتدفق وسائل الإعلام والاتصال العالمي والخدمات الشخصية، حيث يتم تجميع مزيج من كل هذه الميزات والقدرات من عدة أنظمة إلكترونية في نظامٍ واحدٍ مبني على الحاسب الإلكتروني، ويتسم بالبساطة والتكامل؛ لتمكين الأفراد من التفاعل واللعب، والتواصل والتعاون وتبادل المعلومات في العديد من الطرق الجديدة والمختلفة. إلا أن ما يهمنا هنا في سياق الحفظ الرقمية وإتاحة المخطوطات يكمن في سهولة دمج كافة أشكال التعبير عن المعرفة (مصادر المعرفة)، على اختلاف لغة نشرها وأبعادها ومجالاتها، في مشروعٍ عالمي واحدٍ، وليكن المكتبة الرقمية العالمية World Digital Library، ويشمل ذلك الآتي: الوثائق – Documentsمعرفة الأفراد والجماعات – Individual & Group Knowledge الإدارات والمؤسسات Departments & Organizations. ويمكن لهذا المشروع أن يحقق الآتي: (١) حفظ الأصول وعدم المساس والعبث بها – (٢) النشر والإعلان والإحاطة العالمية بهذه الكنوز – (٣) الحد من الاجتهادات المكررة والمتضاربة أحيانًا والمشوهة أو الكاذبة أحيانًا أخرى. (شاهين، ٢٠٠٧)

والمتتبع لتطور قواعد ومعايير الوصف والإتاحة لمصادر المعلومات، على اختلاف أنواعها بما في ذلك الكيانات الرقمية أو كيانات المعلومات الرقمية، يعلم تمامًا أنها تتجه في طريقٍ واحدٍ وهو"الروابط والسياقات Links and Contexts"، وهو ما يعني - بكل وضوح – أهمية المحتوى أو المضمون وتكامله مع غيره.

أسس معهد الاستشراق التابع لجامعة زيورخ مدرسة فريدة من نوعها، تهدف إلى تعليم قراءة البرديات والمخطوطات والوثائق العربية القديمة بواسطة برنامج آلي مُبتكر. اللافت في تلك المدرسة أنها تساعد مراكز البحث والتوثيق المعنية من جميع أنحاء العالم على قراءة البرديات والمخطوطات القديمة، دون مقابل. ويقول البروفيسور كابلوني: إن العمل في هذا المشروع استغرق عامين كاملين بالتعاون مع "الجامعة الدولية للبرديات العربية" في برينستون الأمريكية، ومعهد متخصص في دراسة البرديات اليونانية بجامعة هايدلبرغ في ألمانيا، بالإضافة إلى متحف في برلين وآخر في الولايات المتحدة الأمريكية. وتعتمد المدرسة في طريقة تعليمها على دراسة أسلوب الكتابة في ٧ برديات تاريخية، ثلاث منها تعود إلى العصر الأموي، وهي عبارة عن قراراتٍ من والي مصر، والرابعة من العصر الفاطمي، أما الخامسة فهي رسالة من القرن العاشر بين شخصين عاديين، بينما السادسة والسابعة تعاقدات للشراء بين شخصين عاديين من نفس الفترة. وما أن يتم التعرف على شكل الخط وأسلوب الكتابة، يتمكن الباحث من مقارنة أشكال الكلمات مع تلك المكتوبة بالخط العادي المتداول الآن من خلال الحاسب الآلي، الذي يعطيه أقرب الاحتمالات والربط بين المصطلحات والمضمون لتلافي اللبس بين الكلمات، ويساعد البروفيسور كابلوني في توضيح أي لبسٍ قد يختلط على الدارس أو الباحث. إذ يمكن أن تكون هذه المدرسة

أولى الخطوات لتوثيق البرديات والمخطوطات العربية المنتشرة حول العالم. (العينين، ٢٠٠٤) (Rudolph, 2019)

٢-٥-٢ المسارات المختلفة لرقمنة نصوص المخطوطات:

بعد تتبع المسارات المختلفة لإتاحة بدائل التراث الوثائقي بصفة عامة، والمخطوطات على وجه الخصوص، يتبين وجود المسارات الآتية:

١. الإدخال الآلي للنص.

٢. ترميز النص الفائق.

٣. المسح الرقمي بمصاحبة التعرف البصري على الحروف[1] OCR، وهو المسار الذي يقع في منطقة وسط بين الرقمنة بنمط الصورة والرقمنة بنمط النص. وتجدر الإشارة إلى صعوبة التعرف البصري على الكتابة الخطية القديمة، فضلاً عن تدني نسب نجاحها. وبالتالي ينصح الباحثون بالقراءة المتأنية للصور الرقمية، بالرغم من صعوبة هذه العملية وعدم وجود أية تطبيقاتٍ لأدوات مساعدة للتعامل مع النصوص الإلكترونية. (Estill، ٢٠١٦)

٤. المسح الرقمي للمصغرات الفيليمة (الميكروفيلم – الميكروفيش)، وقد تبين معانة المنتج النهائي من الآتي: انخفاض درجة الوضوح – عدم انضباط درجات الإضاءة في الصفحة الواحدة وعلى مستوى كافة الصفحات – وضوح بصمات الأصابع، وأحيانًا الأصابع نفسها – الغياب التام للألوان وكافة أشكال التذهيب، وغيرها من وسائل وأشكال التمييز والتوضيح في المتن.

٥. التصوير الرقمي للأصول بالكاميرات الرقمية.

٦. المسح الرقمي للأصول بأجهزة المسح الرقمي.

وخلاصة القول هو ما تم توضيحه في الشكل (٥) أن هناك نمطان لرقمنة المخطوطات وعدة مسارات، وهما: نمط النص Textual mode (الرقمنة والفهرسة معًا)، ونمط الصورة Image mode. وبالطبع لا يزال نمط الصورة المتربع على عرش مشروعات التحول الرقمي للتراث الوثائقي والمخطوطات.

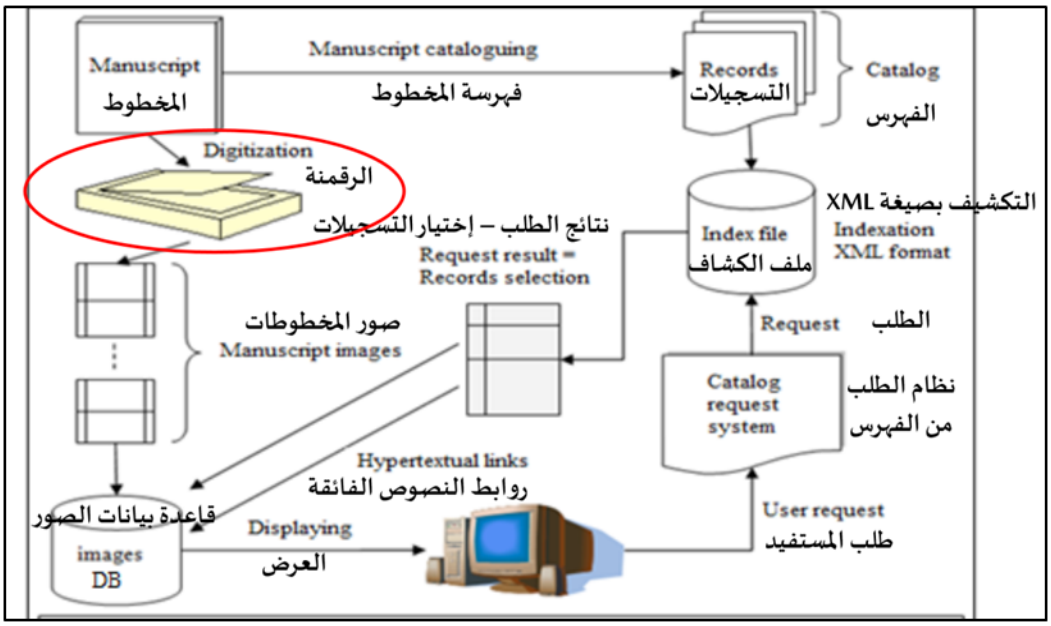
(1) OCR: Optical Character Recognition

شكل(٥) نمطان لرقمنة المخطوطات وعدة مسارات

إلا أن ما يؤخذ على نمط النص Textual mode أنه يستغرق وقتًا طويلاً، كما يعتمد على أكواد خاصة، وبالتالي برمجيات خاصة أيضًا للقراءة الإلكترونية. وعلى الجانب الآخر يُحسَب لنمط النص ما يقدمه من تيسيراتٍ في القراءة وتفسيرٍ للمحتوى وإمكانيات للربط بين مفردات المحتوى، هذا فضلاً عن إمكانية إنشاء قواعد بياناتٍ للنصوص الكاملة مع إمكاناتٍ للبحث تتسم بالشمول. أما نمط الصورة Image mode، فأهم ما يميزه سرعة الإعداد، إلا أنه يفتقد أساسيات دعم استرجاع المعلومات (المحتوى)، وخاصة بحث النص الكامل. ولا ننسى أن الصورة الرقمية Digital Image لا توصف أبدًا بأنها بنائية أو مهيكلة structured، ولكنها مرئية!، وللوصول إليها واسترجاعها لا بد من مصاحبة بياناتٍ لها تدعم استرجاعها. وهنا تكمن الحاجة الملحة لنظام فهرسة دقيق لوصف المخطوطة الرقمية، كما يمكن إثراء هذا الحل بخدماتٍ للشرح والتعليق والتصفح في قاعدة بيانات الكيانات الرقمية.

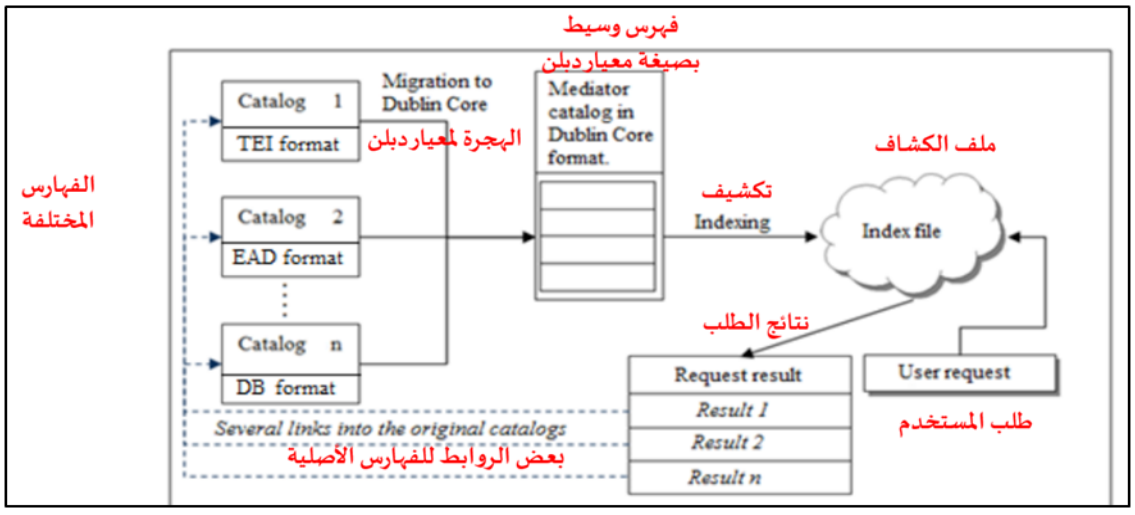
وتتناول الأبحاث المنشورة العديد من الجوانب المتعلقة بالنمطين؛ من أجل إظهار إما جوانب القوة أو مظاهر الضعف. وأي نمطٍ يحتاج إلى مصاحبة نظام فهرسة دقيق؟ مع تبرير الإجابة وتقديم الأمثلة. ويقترح الباحثون ما يُسمى بالحل الذكي للرقمنة Intelligent digitization solution، والمتضمن الجمع بين النمطين للحصول على أفضل النتائج، حيث سهلت تقنيات الربط Linking technology الانتقال بين نمطٍ وآخر. (Soualah & Hassoun, 2011)

ويوضح الشكل (٦) منظومة مقترحة من فريق باحثين لرقمنة المخطوطات العربية، وإتاحة صورها الرقمية، وتيسير البحث والوصول إليها من خلال تسجيلات الفهرسة.(Soualah & Hassoun, 2011)



شكل(٦) مقترح نظام لرقمنة المخطوطات العربية وإتاحة صورها

بينما يصف الشكل(٧) هيكل استخدام معيار دبلن للميتاداتا Dublin Core لدعم احتياجات التشغيل المتداخل لفهارس المخطوطات العربية الرقمية، معتمدين على بروتوكول "بروتوكول مبادرة الأرشيفات المفتوحة لجمع البيانات الوصفية Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). تكمن الفكرة في عمل معيار دبلن كجسرٍ بين الفهارس المختلفة بتنسيقاتها المتباينة. (Soualah & Hassoun, 2011)



شكل(٧) الحل المقترح لدعم احتياجات التشغيل المتداخل لفهارس المخطوطات العربية الرقمية

تتمثل الاستراتيجية في إنشاء تسجيلاتٍ ببليوجرافية، باستخدام مفردات أو عناصر Simple Dublin Core، بدءًا من الفهارس المعتادة والتي قد يتم ترميزها بتنسيقات مختلفة، بينما يتم إنشاء رابطة تلقائيًّا بين التسجيلة المعتمدة على معيار دبلن Dublin Core والتسجيلة الأصلية، وذلك من خلال العنصر<dc: relation>.

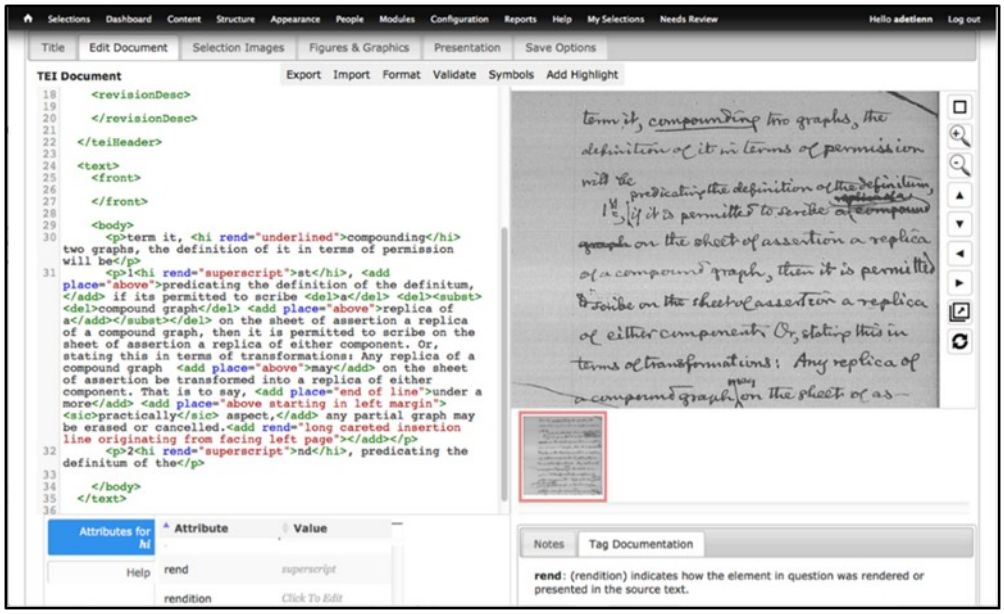
وبهذه الطريقة يتم الانتفاع من هيكل المعيارDublin Core في الوصول إلى التسجيلات الببليوجرافية بطريقة أكثر انضباطا وفق الاسم أو العنوان أو الموضوع. وعلاوة على ذلك، يتم توفير الوصول/ الإتاحة من خلال الكلمات الدالة، باستخدام فهرس معيار Dublin Core باستخدام العنصر<dc: description>.

حيث صمم هذا العنصر ليشتمل على المحتويات الفكرية لمصدر المعلومات الرقمي (كيان المعلومات الرقمي)، كما يمكنه أن يوفر فرصة عظيمة لتنفيذ نظام التكشيف الحر للنصوص Free text indexing system. (Soualah & Hassoun, 2011)

ومن المؤكد أن الرقمنة وسيلة فعَّالة للكشف عن المخطوطات وغيرها من الكنوز المخبأة منذ فترة طويلة في مكتباتنا. ومع ذلك، فإن مجرد تقديم صور رقمية للمخطوطة لا يكفي؛ لأن الصور لا تنقل العديد من الجوانب المهمة في المخطوط، وتشمل: الوصف الكوديولوجي (مجموعة المخطوطات، والتجليد، إلخ) – الوصف الباليوجرافي (الكتابة اليدوية، إلخ) – تاريخ انتقال المخطوطات. (Soualah & Hassoun, A TEI P5 Manuscript Description Adaptation for Cataloguing Digitized Arabic Manuscripts, 2012)

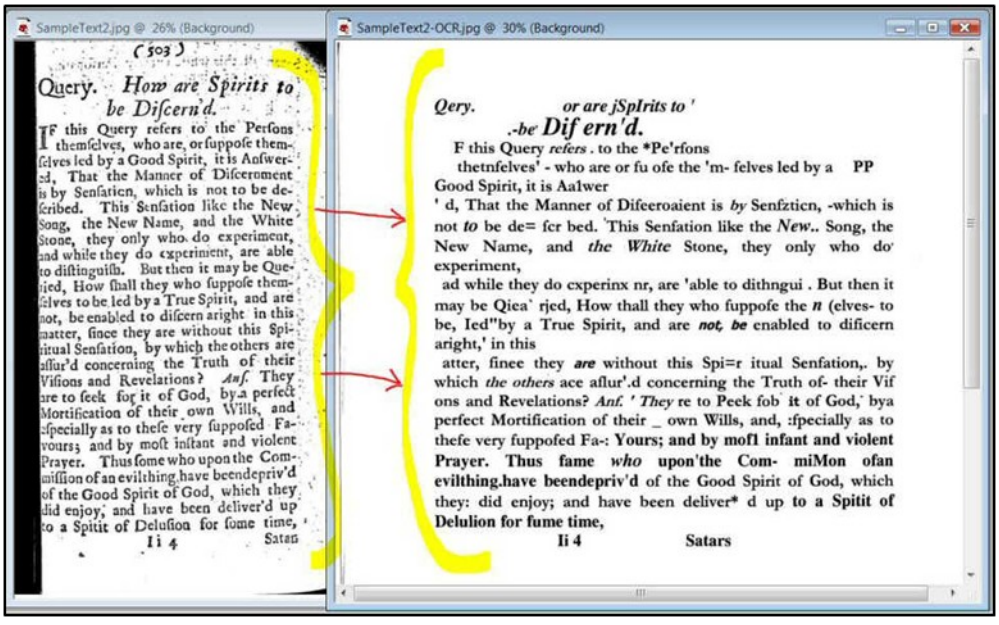
٢-٥-٣ العروض المتنوعة للنصوص الإلكترونية للمخطوطات:

شهد عاما ٢٠١٣ و٢٠١٤، للمرة الأولى، ظهور معايير مفتوحة لمشاركة الصور وعرض الكيانات الرقمية المعقدة في شاشات العرض. وقد تصدر المشهد ما يُعرَف بالإطار العام الدولي للعروض التوافقية للصور الرقمية International Image Interoperability Framework (IIIF). وقد التزمت بهذا الإطار العام للعرض العديد من المكتبات الوطنية والجامعات والمتاحف وغيرها من مؤسسات التراث الثقافي.(Digirati, 2019). كما واكبت تقنيات العرض التطور العالمي في تكويد النصوص القديمة، ومنها المخطوطات، ووفرت إمكانية عرض الصورة الرقمية للنص الأصلي مقابل النص المكود بـ TEI شكل(٨).



شكل(٨) عرض الصورة الرقمية للنص الأصلي مقابل النص المكود ب -TEI (htt8)

كما يمكن عرض الصورة الرقمية للنص الأصلي مقابل النص الناتج عن تقنية التعرف البصري على الحروف شكل(٩).



شكل(٩) عرض الصورة الرقمية للنص الأصلي مقابل النص الناتج عن تقنية التعرف البصري على الحروف (htt9)

٢-٦ مشروعات رقمنة النصوص القديمة بمبادرة TEI: الفهرست نموذجًا

خصص موقع مبادرة ترميز النصوص TEI صفحة تضمنت حصرًا بالمشروعات العالمية المُستخدَمة للمبادرة. (TEI, TEI: Projects Using the TEI, 2019) ومن بين عشرات المشروعات كان مشروع **الفهرست:** الفهرست الموحد للمخطوطات الإسلامية وغيرها من المخطوطات الشرق أوسطية. وكانت النشأة كمشروعٍ مدعم مشترك بين أكسفورد وكمبريدج عُرف بـ Oxford-Cambridge Islamic Manuscript Catalogue on-line (OCIMCO) ، يهدف إلى إتاحة المجموعات التاريخية من المخطوطات الإسلامية في المملكة المتحدة قابلة للبحث على الإنترنت دون أية قيودٍ للوصول إليها. كما كان هناك هدف آخر للمشروع هو إنشاء ممارسة عملية لمخطط Schema مصمم خصيصًا لوصف مخطوطات الشرق الأوسط. ولهذا الغرض توسع المشروع وانضم إليه شركاء مساهمين من هيئات أخرى. (Bodleian Library, 2019)

القسم الثالث: التصميم المنطقي لمنظومة "المحقق الإلكتروني" للنصوص العربية القديمة

٣-١ الإنسانيات الرقمية وتحليل النصوص:

عرفت الإنسانية تطبيقات الحاسب الآلي في معالجة اللغة وتحليل النصوص ضمن إطار معرفي واسع يُعرف بالإنسانيات الرقمية Digital Humanities يحظى باهتمام العالم، ويتسم بتعدد الارتباطات، ويستهدف استثمار الإمكانات الفائقة لتكنولوجيا المعلومات والاتصالات؛ من أجل الارتقاء بأساليب ومناهج تدريس العلوم الإنسانية، والبحث العلمي أيضًا (California, 2013). ومن أبرز تطبيقات الإنسانيات الرقمية: مبادرة تكويد النصوص Text Encoding Initiative – TEI"، والتحرير الإلكتروني Electronic editing، والاتصال العلمي والنشرScholarly communication and publishing، والأدب (النتاج الفكري) الإلكتروني E-literature، وتحليل النصوص وتنقيبها Textual analysis and text mining وغيرها. والمتخصص في موضوع "تحليل النصوص الرقمية وأدواتها Textual analysis tools"، خاصة المتاح منها على الشبكة العنكبوتية (الويب)، يهتم بالإجراءات والعمليات السابقة للتحليل، وتشمل عمليات تكويدها أو ترميزها Text Coding، واللغات المستخدمة لهذا الغرض، والمعايير العالمية التي وضِعت من أجلها. وتجدر الإشارة إلى جهود جماعة بحث اللغات التابعة لجامعة ليدز بإنجلترا في إنشاء مستودع لمفردات القرآن الكريم؛ بغرض تيسير المستويات الثلاثة للتحليل اللغوي: التحليل الصرفي، وشبكات الإعراب، وأنطولوجيا الدلالات (توصيف المعاني).

Leeds, 2019)

وفي سبيل البحث عن إجابات للسؤال الرئيس للبحث: كيف يمكن لعملية التحقيق Critical editing أن تستفيد من تكنولوجيا المعلومات والاتصالات في هذا المجال؟، وهو ما يشمل تحليل النصوص وتدقيق الإصدارات والنسخ المختلفة للمخطوط الواحد، سواء بلغة واحدة أو أكثر من لغة في حال المخطوطات المترجمة.

قدمت معامل الأبحاث في مجال الإنسانيات الرقمية نظمًا إلكترونية للشروح والروابط في ما بين مصادر المعلومات المختلفة في مجالات معرفية متنوعة، منها الفنون والتاريخ وغيرها. وفيما يلي نماذج فعلية لهذه الأنظمة:

- منظومة SALSAH. (Lausanne, 2019)
- مبادرة The Open Islamicate Texts Initiative (OpenITI) لإتاحة النصوص في الحضارة الإسلامية. (Aga Khan University (AKU), 2016)
- قاعدة بيانات رقمية للمصادر الكتابية العربية، وقاموس تحليلي للشعر العربي المبكر والكلاسيكي. (ESF), 2011))، ويشمل المشروع طرق البحث الجديدة في معاجم اللغة العربية، وتأسيس مجموعة تعاون دولية لمشروعٍ جديدٍ على قاعدة بياناتٍ تحليلية للكتاب والشعر العربي، وتنظيم ورش العمل بين كبار العلماء الأوربيين الذين تناولوا دراسة اللغة العربية في السياق المناسب لمشروع قاعدة البيانات. واستهدف المشروع على وجه الخصوص: (١) النقوش العربية، (٢) البرديات، (٣) نص القرآن، (٤) الشعر العربي المبكر.
- قاعدة بيانات تحليلية للشعر العربي. (University, 2013)
- رقمنة أوائل المطبوعات من الكتب العربية. (University, 2013)

كما تجدر الإشارة إلى أنه، وفي سياق البحث الميداني لمشروعات مماثلة أو قريبة لمشروع "المحقق الإلكتروني للنصوص القديمة"، توصل الباحث إلى العديد من المشروعات التي استثمرت تقنيات التحليل الإلكتروني للنصوص القديمة تحت مظلة ما يُعرف بالفيلولوجيا أو (علم النصوص القديمة) أو دراسة النصوص اللغوية دراسة تاريخية مقارنة؛ لفهمها والاستعانة بها في دراسة الفروع الأخرى التي يبحث فيها علم اللغة. وتشير الكتابات إلى أنه في القرن التاسع عشر لم يميز بين هذا المصطلح ومصطلح علم اللغة؛ وذلك لارتباط البحث اللغوي بالنصوص القديمة أيضًا. ولذلك نرى أن مجال الفيلولوجيا يتحدد في قسمين:

١. قسم اختص بفك الرموز القديمة والاهتمام بالآثار .

٢. قسم اهتم بتحقيق النصوص والمخطوطات؛ بغية نشرها.

والفيلولوجيا تعني أيضًا: معرفة الأدب الجميل ودراسة نصوصه – دراسة لغة معينة بالتحليل النقدي لنصوصها – الدراسة الشكلية للنصوص في المخطوطات.

ويرجع مصطلح الفيلولوجيا إلى عام ١٩٦٠، وكان الهدف من الفيلولوجيا في دراسة النصوص إعادة تشكيل أو تجديد اللغة المنقرضة أو الميتة، ولاسيما بعد اكتشاف اللغة السنسكريتية، حيث اعتمدت الفيلولوجيا على القِدم والمقارنة من أجل هذا التجديد. ومفهوم الفيلولوجيا عند الغربيين ينصرف إلى دراسة اللغتين: الإغريقية واللاتينية، وبالتالي لا يجدون حرجًا للرجوع إلى هاتين اللغتين لإيجاد مصطلحاتهم العلمية، مثل (الجيولوجيا -الفيزيولجيا.... إلخ). وقد نبه المستشرق (غويدي) إلى الفرق بين فقه اللغة والفيلولوجيا بقوله: إن كلمة الفيلولوجيا تصعب ترجمتها بالعربية، وهي لا تدل على مصطلح فقه اللغة؛ فبينهما فروق واختلافات. (الحمدو، ٢٠٠٦)

ومن أبرز المشروعات ذات الصلة بهذا المجال الآتي: مشروع مخطوطات القرآن والفيلولوجيا الرقمية. وقد أكد الباحث الرئيس للمشروع على أن "الفيلولوجيا" والأدوات/ المناهج الرقمية بمثابة ثورةً في مجال العلوم الإنسانية، أعادت النظر بشكلٍ جديدٍ نحو المخطوطات القرآنية وتحقيقها؛ بهدف إنتاج أرشيف ثري يعرض جميع الشواهد المخطوطية، ويفسر عملية إنتاج نصوص المخطوطات والسياق الذي كُتِبَت فيه. (فيديلي، ٢٠١٦)

٣-٢ تحقيق المخطوطات الماضي ... الحاضر ... المستقبل:

التحقيق هو بذل الجهد واستقصاء البحث؛ بغية الوصول إلى حقيقة ما قاله مؤلف النص، أو هو: عملية مركبة تقتضي إخراج نص مضبوط يكون على الصورة التي قاله عليها صاحبه، أو أقرب ما يكون إلى ذلك على الأقل. ويقتضي هذا أن يطرح من يريد تحقيق نص ما عدة تساؤلات، منها:

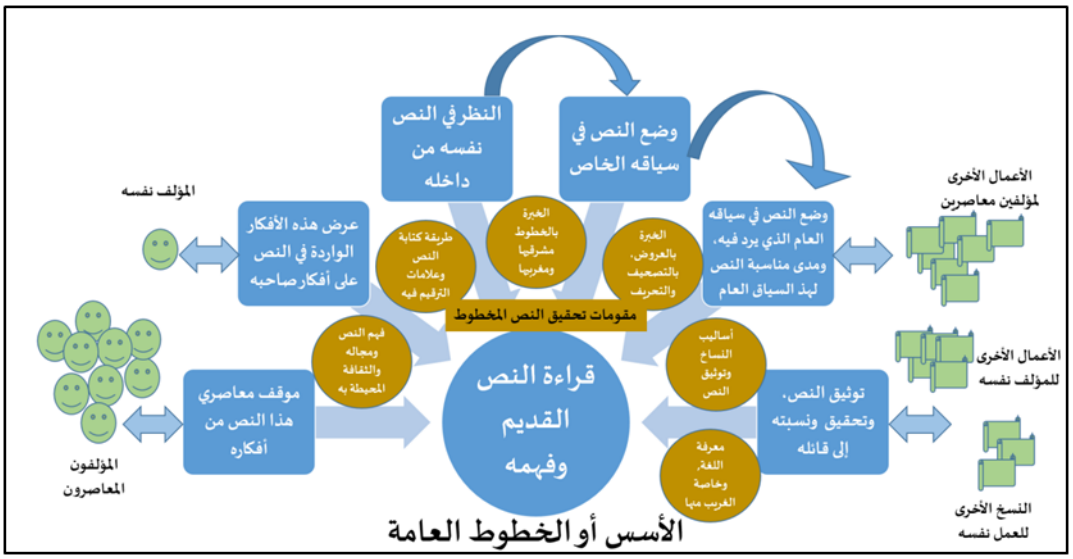
- هل نسبة النص إلى مؤلفه صحيحة؟ وإذا لم تكن صحيحة، فهل النص منسوب خطأ إلى غير صاحبه، أو أنه نص منتحل بأكمله؟
- هل النص نقي كامل خالٍ من التشويه أو النقص أو الزيادة؟
- ما تاريخ النص؟

وبذلك، يكون الكتاب المحقق هو الذي صحَّ عنوانه، واسم مؤلفه، ونسبة الكتاب إليه، وكان متنُه أقرب ما يكون إلى الصورة التي تركها مؤلفه. (فلوسي، ٢٠٠٧-٢٠٠٨)

ومن أبرز المشاكل التي تواجه النص التراثي الآتي: (اللطيف، ٢٠٠٩)

- قضية الانتحال في الشعر العربي القديم.
- نسبة بعض الكتب الى غير أصحابها.
- فهم بعض النصوص القديمة على غير وجهها.
- بقاء بعض النصوص القديمة جامدة ثابتة دون تطويرٍ أو استثمار.
- أخذ النصوص من أطرافها، دون تمحيص.
- التفسير على فهمٍ خاص يريده الدراس.
- إنقاص النص بحذف بعضه أو بتره؛ خدمة لغرض لا يؤدي إليه النص، وإنما إلى ما يريده الباحث.
- الفهم السقيم ناقص العدة. الفهم الخاطئ بسبب عدم فهم السياق الخاص والعام للنص، أو عدم فهم أفكاره الخاصة، أو عدم النظر إلى مدى مواءمة هذه الأفكار مع آراء أصحابها الأخرى. إلى غير ذلك من دواعي النظر الخاطئ والفهم المنقوص.

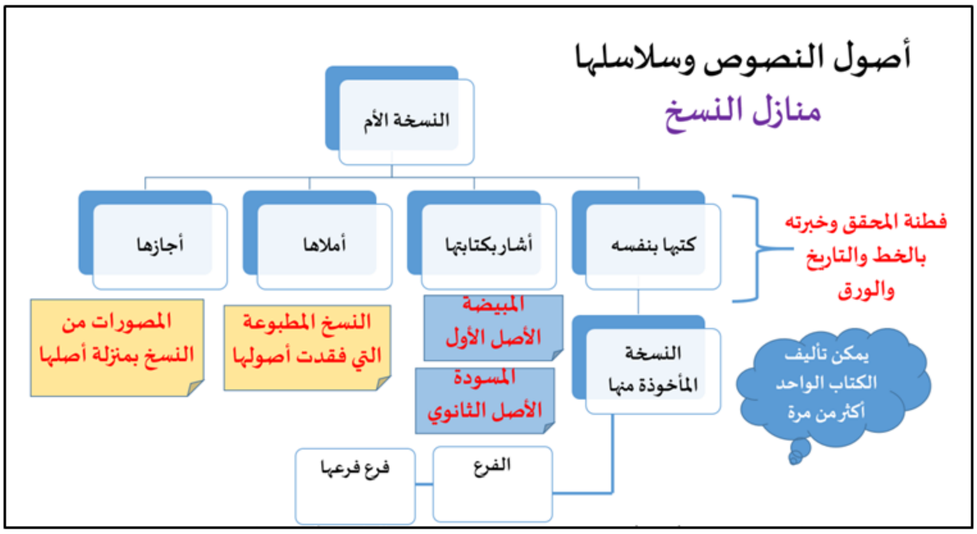
ومعظم ما سبق مشكلات يمكن لمنظومة المحقق الإلكتروني مواجهتها وتقديم الحلول لها. كما اتفق معظم الكُتاب والمحققين على خطوات أساسية يمكن تسميتها بالأسس أو الخطوط العامة للقراءة والتحقيق للنصوص القديمة كما يوضحها الشكل(١٠). (اللطيف، ٢٠٠٩)



شكل(١٠) الخطوط العامة والإطار العام لمنظومة المحقق الإلكتروني

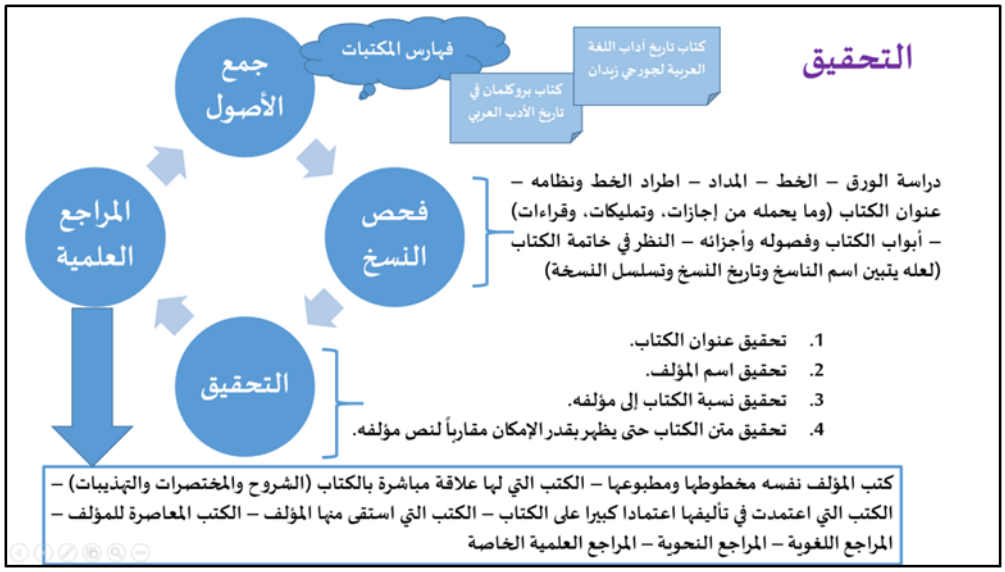
ومن أهم خطوات التحقيق: تحديد النسخة الأم للعمل والنسخ الأخرى ومنازلها (كتبها بنفسه أو أشار بكتابتها أو أملاها أو أجازها...)، وماذا عن المبيضات والمسودات للعمل نفسه؟ وهل هناك

نسخ محققة؟ وغيرها من التساؤلات التي يُفترَض في منظومة المحقق الإلكتروني الإجابة عنها، من خلال عملية بحث واحدة كما يوضحها الشكل(١١). (هارون، ١٩٩٨)



شكل(١١) تحقيق النصوص كما وردت في المؤلفات المنشورة

وإجمالاً، يوضح الشكل(١٢) مهمة أو صنعة التحقيق وجوانبها أو أبعادها المختلفة التي تتعلق بالمحتوى أو المتن، والوسيط المادي المُستخدَم وأدوات الكتابة ونوع الخط ... وغيرها. والواضح جدًّا أن التحقيق قائم على المقارنة والمضاهاة، وهو ما يعني أن المنظومة المقترحة للمحقق الإلكتروني يُفترَض بها الاعتماد على مستودع رقمي للنصوص الكاملة للمخطوطات والفهارس وكتب التراث والمراجع العلمية وكتب التحقيق وكافة الأعمال المطبوعة للمخطوطات المحققة، بصرف النظر عن أماكن نشرها وتواريخها، ويمكن أن نضيف إلى ما سبق الأعمال باللغات الأجنبية في المجال داخل الحدود نفسها الموضوعية والنوعية والجغرافية والزمنية.



شكل(١٢) تحقيق النصوص بأبعاده المختلفة

بينما يوضح الشكل (١٣) فئات المراجع اللغوية الواجب توافرها بنصوصها الكاملة في المستودع الرقمي الداعم للمحقق الإلكتروني، وتشمل: معاجم الألفاظ، والمعاني، والأسلوب، والمعربات، واللغات التي تمتّ بصلةٍ وثيقة للغة العربية والفارسية والأردية وغيرها.

شكل(١٣) المراجع اللغوية اللازمة لمحقق النصوص

٣-٣ التصميم المنطقي المقترح للمحقق الإلكتروني للنصوص العربية القديمة:

٣-٣-١ بناء أو تأسيس مستودع النصوص التراثية العربية والإسلامية:

وهو بمثابة قاعدة بيانات المراجع العلمية واللغوية الإلكترونية للنصوص الكاملة والميتاداتا المعيارية - متاحة على الخط المباشر على الإنترنت- تحصر وتصف النصوص التراثية العربية والإسلامية في المجالات المعرفية المختلفة، وتربط فيما بينها وبين ما تمت ترجمته لليونانية واللاتينية واللغات الأوربية الحية، هذا إلى جانب ما تم تحقيقه منها وتم نشره. ويمكن لهذا المشروع الأولي أن يحقق الاستفادة من معايير التنظيم الجديدة للوصف العالمي للكيانات المعلوماتية في البيئة الرقمية وتقنيات البيانات الرابطةLinked Data وتقنيات ترميز النصوص الفائقة، وغيرها. ويوضح الشكل (١٥) بنية ومكونات المستودع الرقمي المقترح.

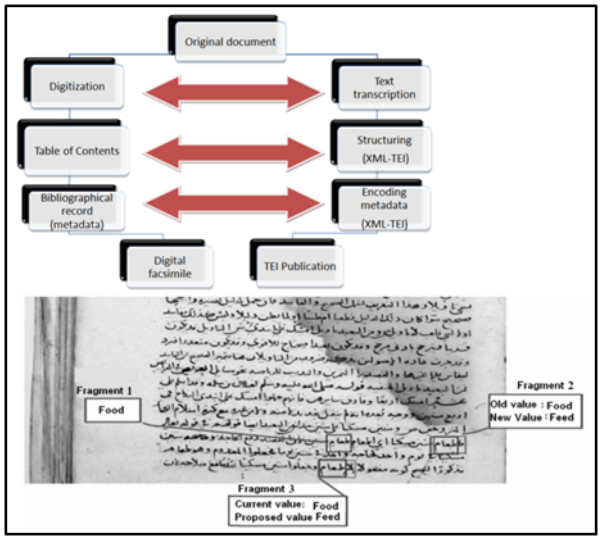
شكل(١٥) بنية مستودع النصوص التراثية العربية والإسلامية

٣-٣-٢تطوير وترقية منظومة عربية موحدة للتعرف الضوئي على الحرف المخطوط العربي وترميز النص الأصلي:

وهو تطور تقني بات لازمًا لأمورٍ تتعلق بالبحث العلمي وتنظيم التراث العربي، على اختلاف أشكاله وموضوعاته، ويستهدف تيسير استكشاف مضمون تلك الأصول التراثية والتعرف عليها إلكترونيًّا، هذا إلى جانب التقليل أو الحدّ من الاعتماد على عناصر الميتاداتا في التعرف على المحتوى، مقابل إتاحة البحث الحر على مستوى الكلمة داخل المتن على مستوى المصدر الواحد أو عدة مصادر. ويوضح الشكل(١٦) الاختيارات المختلفة أمام مشروع منظومة التعرف الضوئي وترميز النصوص. وتشمل:

- الرقمنة مقابل الكتابة أو الإدخال الإلكتروني للنص.

- قائمة المحتويات مقابل الترميز أو تكويد النص الكامل وفق معيار XML – TEI.
- التسجيلة الببليوجرافية (الميتاداتا) مقابل الميتاداتا المكودة وفق معيار XML – TEI.
- الصورة الرقمية مقابل نشر النص الكامل وفق المعيار TEI.



شكل(١٦) الاختيارات المختلفة أمام مشروع منظومة التعرف الضوئي وترميز النصوص

٣-٣-٣ المحقق الإلكتروني أداة كل محقق للتراث العربي:

وهو المنتج النهائي المترتب على ما سبق، ويشمل تجهيز منظومة إلكترونية متكاملة بعنوان «المحقق الإلكتروني»، أو هو نظام معلومات متكامل داعم لمعظم مراحل وخطوات عملية تحقيق التراث. ويمكن للمحقق الإلكتروني تحقيق أقصى فائدة من إمكانات تقنيات ترميز النصوص الإلكترونية، والربط فيما بينها لأغراضٍ مختلفة. كما يمكن لنظام المحقق الإلكتروني تقديم العون في الخطوات الأساسية في عملية التحقيق التي تبدأ بالتفكير في اختيار النص، مرورًا بتوثيقه، وآليات البحث عن نُسخه وترجماته، وقراءة النص وتحريره ونقده ودراسته وخدمته.

القسم الرابع: النتائج والتوصيات

٤-١ النتائج:

- بالرغم من تأكيد الباحثين على أن النسخ التلقائي أو اليدوي Automatic or Manual Transcriptionللمخطوطات العربية القديمة أمرٌ حتمي للتكشيف ونشر محتوياتها، نظرًا للطبيعة الخطية للكتابة العربية التي تُعد عائقًا أمام برامج التعرف الضوئي على الحروف (OCR)، فإن هذه الطريقة ليست الحل المثالي؛ نظرًا لحقيقة عدم تنظيم أو وجود بنية للمحتوى، ولهذا السبب تتجه الأنظار نحو تكويد النصوص الإلكترونية وبخاصة معيار TEI.
- تُعد مبادرة ترميز النصوص TEI: Text Encoding Initiativeمن أبرز المشروعات الدولية التي استهدفت وضع مبادئ توجيهية لإعداد وتبادل النصوص الإلكترونية للأغراض البحثية. وتجدر الإشارة إلى أهم أقسام هذه المبادرة، وهو القسم العاشر المتعلق بوصف المخطوطات والذي خضع للتحديث مؤخرًا في ١٦ يوليو ٢٠١٩ م في إصدارة جديدة ٣.٦.٠.
- غياب فكرة تكويد النصوص الفائقة في مجال النصوص القديمة والمخطوطات والوثائق القديمة في وثيقة الإفلا الصادرة عن قسم الإفلا للكتب النادرة والمخطوطات عام ٢٠١٤م بعنوان: "إرشادات للتخطيط لرقمنة الكتب النادرة والمخطوطات".
- الفجوة المعرفية بين خبراء التراث الثقافي والباحثين وخبراء تكنولوجيا المعلومات والحاسب الإلكتروني، وخصوصًا في البلدان النامية.
- هناك نمطان لرقمنة المخطوطات وعدة مسارات، وهما: نمط النص Textual mode (الرقمنة والفهرسة معًا)، ونمط الصورة Image mode. بينما تشمل المسارات: الإدخال الآلي للنص – ترميز النص الفائق – المسح الرقمي بمصاحبة التعرف البصري على الحروف OCR – المسح الرقمي للمصغرات الفيليمة (الميكروفيلم – الميكروفيش) – التصوير الرقمي للأصول بالكاميرات الرقمية – المسح الرقمي للأصول بأجهزة المسح الرقمي.
- الحل الذكي للرقمنة Intelligent digitization solution يعني الجمع بين النمطين (نمط النص ونمط الصورة)؛ للحصول على أفضل النتائج، حيث سهلت تقنيات الربط Linking technology الانتقال بين نمطٍ وآخر.
- واكبت تقنيات العرض التطور العالمي في تكويد النصوص القديمة، ومنها المخطوطات، ووفرت إمكانية عرض الصورة الرقمية للنص الأصلي مقابل النص المكود بـ TEI، كما وفرت أيضًا إمكانية عرض الصورة الرقمية للنص الأصلي مقابل النص الناتج عن تقنية التعرف البصري على الحروف OCR.

- عرفت الإنسانية تطبيقات الحاسب الآلي في معالجة اللغة وتحليل النصوص ضمن إطار معرفي واسع يُعرف بالإنسانيات الرقميةDigital Humanities يحظى باهتمام العالم، ويتسم بتعدد الارتباطات، ويستهدف استثمار الإمكانات الفائقة لتكنولوجيا المعلومات والاتصالات؛ من أجل الارتقاء بأساليب ومناهج تدريس العلوم الإنسانية، والبحث العلمي أيضًا.
- قدمت معامل الأبحاث في مجال الإنسانيات الرقمية نظمًا إلكترونية للشروح والروابط فيما بين مصادر المعلومات المختلفة في مجالاتٍ معرفية متنوعة، منها الفنون والتاريخ وغيرها.
- هناك مشروع قائم يستهدف مخطوطات القرآن باستثمار الفيلولوجيا الرقمية. وقد أكد الباحث الرئيس للمشروع على أن "الفيلولوجيا" والأدوات/ المناهج الرقمية بمثابة ثورة في مجال العلوم الإنسانية، أعادت النظر بشكلٍ جديدٍ نحو المخطوطات القرآنية وتحقيقها؛ بهدف إنتاج أرشيف ثري يعرض جميع الشواهد المخطوطية، ويفسر عملية إنتاج نصوص المخطوطات والسياق الذي كُتِبَت فيه.
- إن معظم ما يُسمى مشكلات تحقيق النصوص القديمة، يمكن لمنظومة المحقق الإلكتروني مواجهتها وتقديم الحلول لها.
- ينبغي على المنظومة المقترحة للمحقق الإلكتروني الاعتماد على مستودع رقمي للنصوص الكاملة للمخطوطات والفهارس وكتب التراث والمراجع العلمية وكتب التحقيق وكافة الأعمال المطبوعة للمخطوطات المحققة، بصرف النظر عن أماكن نشرها وتواريخها، كما يمكن أن نضيف إلى ما سبق الأعمال باللغات الأجنبية في المجال داخل الحدود نفسها الموضوعية والنوعية والجغرافية والزمنية.

٤-٢ التوصيات:

- تبني مؤسسات ذاكرة المجتمعات (المكتبات، والأرشيفات، والمتاحف) بالتعاون مع المؤسسات البحثية للتصميم المنطقي للمحقق الإلكتروني للنصوص العربية القديمة بما في ذلك المخطوطات. وهو ما يعني توزيع المهام والواجبات والتعاون في إنجازها وفق المعايير العالمية.
- توحيد الجهود من أجل الارتقاء بمعارف وخبرات خبراء التراث الثقافي؛ من أجل التعرف على الإنجازات الجديدة في تقنيات المعلومات والاتصالات، لأجل إعداد خطط حفظ وتطوير سياسات أكثر فعالية لحفظ "التراث الرقمي" تواكب العصر الجديد.
- توجيه الاهتمام الى مجال "الإنسانيات الرقمية" بمعناها الأكاديمي في جامعاتنا ومراكز أبحاثنا، ووضع الحوافز المناسبة لإجراء البحوث والدراسات في هذا المجال.
- تدريس مقررات في أقسام المكتبات والمعلومات تستهدف إكساب الدارس المعارف الأساسية والخبرات العملية المرتبطة بتكويد أو ترميز النصوص وتعلم أوجه الاستفادة القائمة على هذا التطبيق أو هذه الممارسة في الأبحاث العلمية.
- إدارج رقمنة التراث الوثائقي وحفظه بالمفهوم الواسع ضمن أولويات الاهتمام في أجندة المؤسسات الثقافية والتعليمية والبحثية، وخاصة مؤسسات ذاكرة المجتمعات. ويُقصَد بالمفهوم الواسع لرقمنة التراث الوثائقي: التكويد والترميز والحلول الذكية التي تجمع بين التصوير والمسح الرقمي وعمليات التكويد أو الترميز للمحتوى أو النصوص.
- تأسيس تجمعات أكاديمية تهتم بالإنسانيات الرقمية، وإصدار مجلاتٍ ودورياتٍ متخصصه فيه .

المصادر:

المصادر الأجنبية:

(n.d.). Retrieved from http://www.iicm.tugraz.at/Ressourcen/Archive/hgbook/images/img2.gif

(n.d.). Retrieved from http://www.nadin.ws/wp-content/uploads/2007/03/vising-fig1.gif

(n.d.). Retrieved from http://www.cyberartsweb.org/cpace/ht/htlecture/images/russell1.gif

(n.d.). Retrieved from http://literacybeat.files.wordpress.com/2011/07/figure-3-hypertext-i-have-a-dream.png

(n.d.). Retrieved from http://www.iicm.tugraz.at/Ressourcen/Archive/hgbook/images/img3.gif

(n.d.). Retrieved from http://refractory.unimelb.edu.au/wp-content/uploads/2012/12/fan2.gif

(n.d.). Retrieved from http://media.nngroup.com/media/editor/2012/10/31/structure-hypercard-objects.gif

(n.d.). Retrieved from http://www.cyberartsweb.org/cpace/ht/christanto/pictures/link_smpl.gif

(n.d.). Retrieved from http://www.peirce.iupui.edu/img/STEP.png

(n.d.). Retrieved from http://www.passtheword.org/jane-lead/FW/OCR-Comparison2.jpg

(ESF), T. E. (2011). Methods of Digital Philology for the Study of Early and Classical Arabic Language. Retrieved from Greek into Arabic - http://www.greekintoarabic.eu/indexd61c.html : http://www.greekintoarabic.eu/fileadmin/pdf/Berlino_ottobre2011_programma.pdf

Aga Khan University (AKU), U. L. (2016). The Open Islamicate Texts Initiative (OpenITI) Creating the digital infrastructure for the study of the premodern Islamicate world. Retrieved from https://iti-corpus.github.io/

Bodleian Library, U. o. (2019, August 5). FIHRIST: Union Catalogue of Manuscripts from the Islamicate World. Retrieved from https://www.fihrist.org.uk/

Brix, L. (2018, August 14). Unpublished Egyptian texts reveal new insights into ancient medicine: The large collection of documents from Ancient Egypt are held in Copenhagen and are revealing new details about medical sciences in antiquity. Retrieved from ScienceNordic: http://sciencenordic.com/unpublished-egyptian-texts-reveal-new-insights-ancient-medicine

California, U. o. (2013). GE Redesign: Humanistic Inquiry . Retrieved from https://strategic.usc.edu/2013/09/24/humanistic/

Clivaz, C., Schulthess, S., & Sankar, M. (2017). Editing New Testament Arabic Manuscripts in a TEI-base: fostering close reading in Digital Humanities. Retrieved from Journal of Data Mining and Digital Humanities: https://hal.archives-ouvertes.fr/hal-01280627/document

Congress, L. o. (2019, April 10). EAD. Retrieved from Library of Congress - Standards: https://www.loc.gov/ead/

Digirati. (2019, August 5). THE UNIVERSAL VIEWER: A leading web-based viewer for digitised artifacts using the IIIF standard. Retrieved from Digirati : https://digirati.com/work/galleries-libraries-archives-museums/products/universal-viewer/

El Makhfi, N., El Bannay, O., Benslimane, R., & Rais, N. (2011). Search engine of ancient

Arabic manuscripts based on metadata and XML annotations. Retrieved from IEEE Xplore : https://ieeexplore.ieee.org/document/6148588/authors#authors

El Makhfi, N., El Bannay, O., Benslimane, R., & Rais, N. (2011). System of indexing, annotation and search in the old Arabic manuscripts. Retrieved from IEEE Xplore Digital Library : https://ieeexplore.ieee.org/document/6148587/authors#authors

Estill, L. (2016). Encoding the Edge: Manuscript Marginalia and the TEI . Digital Literary Studies , https://journals.psu.edu/dls/article/view/59715/59912 .

Graham, S. I. (2013). Final Draft → Early Emergences: Father Busa, Humanities Computing, and the Emergence of the Digital Humanities. Retrieved from The Historian's Macroscope : http://www.themacroscope.org/?page_id=601

Ide, N. (n.d.). Encoding standards for large text resources: The Text Encoding Initiative . Retrieved from ACL Home Association for : https://www.aclweb.org/anthology/C94-1094

Lausanne, D. H. (2019). SALSAH. Retrieved from https://salsah.org/

Leeds, L. R.-U. (2019). The Quranic Arabic Corpus. Retrieved from The Quranic Arabic Corpus : http://corpus.quran.com/

Letellier, R. W. (2007). Recording, documentation, and information management for the conservation of heritage places: guiding principles. Los Angeles: Getty Conservation Institute.

Lewis, D. (2016, November 3). How Experts Are Digitizing Ancient Manuscripts: Digital preservation is more work than it might seem. Retrieved from SMITHSONIAN.COM : https://www.smithsonianmag.com/smart-news/how-experts-are-digitizing-ancient-manuscripts-180960685/

Mceachen, B. (2016, September 29). Advanced new tools for reading Old Testament scrolls: Sophisticated technique means ancient texts can be preserved and digitally enhanced. Retrieved from Eternity News: https://www.eternitynews.com.au/world/advanced-new-tools-for-reading-old-testament-scrolls/

Ogrin, M., Javoršek, J. J., & Erjavec, T. (2013). A Register of Early Modern Slovenian Manuscripts. Journal of the Text Encoding Initiative [Online], https://journals.openedition.org/jtei/715.

Pitti, D. V. (1999). Encoded Archival Description: An Introduction and Overview. D-Lib Magazine, http://www.dlib.org/dlib/november99/11pitti.html .

Quintero, S. (2003). The use of three-dimensional techniques of documentation and dissemination in studying built heritage. Leuven, Belgium: University of Leuven.

Rudolph, U. (2019, July 9). Welcome to Islamic Studies! Retrieved from University of Zurich, Institute of Asian and Oriental Studies: https://www.aoi.uzh.ch/en/islamwissenschaft.html

Soualah, M. O., & Hassoun, M. (2011). Which Metadata for Ancient Arabic Manuscripts Cataloguing? Proc. Int’l Conf. on Dublin Core and Metadata Applications 2011, (pp. 137-146).

Soualah, M. O., & Hassoun, M. (2012). A TEI P5 Manuscript Description Adaptation for Cata-

loguing Digitized Arabic Manuscripts. Journal of the Text Encoding Initiative [Online], http://journals.openedition.org/jtei/398 ; DOI : 10.4000/jtei.398.

TEI. (2011 -). The Journal of the Text Encoding Initiative. Retrieved from TEI: https://journals.openedition.org/jtei/

TEI. (2019, July 16). 10 Manuscript Description. Retrieved from P5: Guidelines for Electronic Text Encoding and Interchange: https://tei-c.org/release/doc/tei-p5-doc/en/html/MS.html

TEI. (2019, August 5). TEI: Projects Using the TEI. Retrieved from TEI: https://tei-c.org/activities/projects/

University, B. (2013, May 22). Analytical Database of Arabic Poetry. Retrieved from Digital Islamic Humanities Project : https://islamicdh.org/2013/05/22/analytical-database-of-arabic-poetry/

Vertan, C., & Reimers, S. (2012). A TEI-based Application for Editing Manuscript Descriptions. Journal of the Text Encoding Initiative, https://journals.openedition.org/jtei/392.

المصادر العربية:

الحمدو ,ع .ا .(٢٠٠٦) .شَبَكةُ الفَصِيحِ لِعُلُومِ اللُّغةِ العَرَبِيّةِ /http://www.alfaseeh.net/vb/ . Retrieved from archive/index.php/t-16581.html

العينين ,ت .ا ,٢٠٠٤) .يوليو .(٢معهد الاستشراق التابع لجامعة زيورخ" :سنقرأ لكم هذه المخطوطة العربية بتقنية حديثة Retrieved from ".!وحدات الإنتاج التابعة لهيئة الإذاعة والتلفزيون السويسرية الأولى =ز SWI swissinfo.ch SWI swissinfo.ch: shorturl.at/epzJ2

اللطيف ,م .ح .(٢٠٠٩) .كيف نقرأ النص التراثي؟ وبيان أثر العروض في ضبطه وتحقيقه .القاهرة :مكتبة الإمام البخاري للنشر والتوزيع.

المكتبات ,ا .ا .(٢٠١٤) .إرشادات للتخطيط لرقمنة الكتب النادرة والمخطوطات.IFLA .

شاهين ,ش .ك .(٢٠٠٧) .مجتمع المعرفة والمكتبة الرقمية العالمية :نموذج مقترح لمعايير اختيار المحتوى لضمان التكامل المعرفي .مؤتمر الاتحاد العربي للمكتبات والمعلومات.جدة، السعودية : الاتحاد العربي للمكتبات والمعلومات.

فلوسي ,م .(٢٠٠٨-٢٠٠٧) .محاضرات في تحقيق المخطوطات .جامعة الحاج لخضر باتنة -كلية العلوم الاجتماعية والعلوم الإسلامية.

فيديلي ,ا .(٢٠١٦) .مخطوطات القرآن والفيلولوجيا الرقمية Retrieved from International Qur'anic Studies . Association: https://iqsaweb.wordpress.com/2016/11/21/%d9%85%d8%ae%d8%b7%d9%88%d8%b7%d8%a7%d8%aa-%d8%a7%d9%84%d9%82%d8%b1%d8%a2%d9%86-%d9%88%d8%a7%d9%84%d9%81%d9%8a%d9%84%d9%88%d9%84%d9%88%d8%ac%d9%8a%d8%a7-%d8%a7%d9%84%d8%b1%d9%82%d9%85%d9%8a%d8%a9/

هارون ,ع .ا .(١٩٩٨) .تحقيق النصوص ونشرها. القاهرة : مكتبة الخانجي.