استخدام برمجيات الحاسوب في رقمنة البنية الهيكلية والمعرفية للموسوعات

Using computer software to digitize the structural and cognitive structure of encyclopedias

فارس شاشة [1]*

[1] جامعة محمد لمين دباغين. مخبر المجتمع الجزائري المعاصر، f.chacha@univ-setif 2.dz

تاريخ الاستلام: 2022/02/28 تاريخ القبول: 2022/06/06 تاريخ النشر: 2022/12/14

ملخص:

استخدمت برمجيات الإعلام الآلي في رقمنة الموسوعات وإتاحتها على الخط مما نتج عنه قيمة مضافة لها من حيث طريقة الإخراج والإصدار والتحديين وطريقة القراءة كما جعلت هذه البرمجيات المستفيد ينتقل بين مداخل الموسوعات بفعالية ويسر ويتمكن من ربط العناصر المتداخلة مع بعضها، وسوف نعمل في مداخلتنا هذه على دراسة أهم استخدامات تقنيات الإعلام الآلي في نمذجة الموسوعات سواء البنية الهيكلية والبنية المعرفية.

حيث سنعمل باستخدام المنهج الوصفي لدراسة التطورات التي عرفتها الموسوعات وتحولها من الشكل الورقي إلى الرقمي وإتاحتها في شبكة الانترنت مما أدى إلى تحول في بنيتها الهيكلية والمعرفية سهل عملية إتاحة المعلومات للقارئ وربط مداخلها بروابط تشعبية حتى أصبحت هذه المداخل تمثل حلقة مرتبطة مع بعضها.

كلمات مفتاحية: *الرقمنة، الموسوعات، الموسوعات الرقمية، البنية الهيكلية، البنية المعرفية.*

Abstract:

Automated software has been used to digitize encyclopedias and make them available online, resulting in added value in terms of output, release, challenges and reading method, and has made this software the beneficiary move between the entries of encyclopedias effectively and easily and able to connect the overlapping elements with each other, and in this intervention we will study the most important uses of automated techniques in the modeling of encyclopedias, both structural and cognitive structure.

Using the descriptive approach to study the developments of encyclopedias and their transformation from paper to digital form and their availability on the Internet, which has led to changes in their structural and cognitive structure, it is easy to make information available to the reader and to link entries and information to hyperlinks until these entrances become a link.

***Keywords**: Digitization; encyclopedias; digital encyclopedias; structure, cognitive structure*

* المؤلف المراسل: فارس شاشة ، الإيميل : f.chacha@univ-setif 2.dz

1. مقدمة :

تعد الموسوعات من أهم مصادر المعلومات لما تحتويه من معلومات دقيقة ومختصرة حول العديد من الموضوعات لذا بدا الإنسان في كتابتها منذ القديم وان كانت لم تظهر في شكلها الحقيقي المعروف حاليا إلا في الحضارة العربية عند تطوير المعاجم اللغوية إلى معاجم موسوعية تشرح أهم الكلمات شرحا واسعا مع ذكر الشواهد والأمثلة.

وفي مرحلة لاحقة أصبحت الموسوعات عبارة عن تجميع شامل لجميع فروع المعرفة الإنسانية أو لفرع واحد بذكر العديد من المقالات تكون مرتبة وفق ترتيب معين، غالبا يكون ألفبائيا وتهدف إلى تقديم حوصلة وافية حول موضوع معين.

لكن مع ظهور الحاسوب وتطوراته نتج عنه تحولت أشكال الموسوعات من ورقية إلى الكترونية رقمية وهذا لتوفير طرق سهلة لبحث واسترجاع المعلومات منها وتقديم طرق عرض مختلفة لعرض المعلومات والإبحار فيها باستخدام الروابط التشعبية، فبرمجيات الحاسوب وتقنيات الانترنت عملت على رقمنة الموسوعات وإنشاء منصات رقمية لإنشاء موسوعات رقمية تتاح عبر الوسائط الحديثة للتكنولوجيا وعبر شبكة الانترنت ويتجلى ذلك مثل موسوعة ويكبيديا، britanica .

نتيجة هذا التحول الرقمي في الموسوعات عملنا في بحثنا هذا على دراسة إشكالية تغير البنية الهيكلية للموسوعات وكيفية استثمار برمجيات الحاسوب في إنشاء موسوعات رقمية ذات بنية معرفية مترابطة

ولدراسة هذه الإشكالية قمنا بطرح التساؤلات التالية:

-كيف تعمل برمجيات الحاسوب على نمذجة البنية الهيكلية للموسوعات؟

-كيف تعمل برمجيات الحاسوب على نمذجة البنية المعرفية للموسوعات؟

أهداف الدراسة:

نعمل في دراستنا هذه إلى تحقيق مجموعة من الأهداف تتمثل في:

-دراسة تطور الموسوعات من الناحية الورقية إلى الموسوعات الرقمية المتاحة على شبكة الانترنت.

-تحليل مكونات الموسوعات من الناحية الهيكلية الشكلية والناحية المعرفية

-دراسة تغير بنية الموسوعات من حيث طريقة عرض المعلومات وارتباطات المداخل مع بعضها نتيجة استخدام برمجيات الحاسوب وانتقالها إلى موسوعات متاحة على شبكة الانترنت.

المنهج:

تعتمد هذه الدراسة على المنهج الوصفي باعتباره الأنسب للموضوع، وهو استقصاء يقوم على ظاهرة من الظواهر كما هي قائمة في الوقت الحاضر، بهدف تشخيصها وكشف جوانبها، وتحديد العلاقات القائمة بين عناصرها والعلاقات بينها وبين ظواهر أخرى، ويشار أيضاً أن الأسلوب الوصفي في البحث هو أحد أشكال التحليل والتفسير العلمي المنظم الذي يصف ظاهرة أو مشكلة محددة ويقوم الباحث العلمي من خلال الأسلوب الوصفي بتحليل الظاهرة تحليلا دقيقاً.

2.مفهوم الموسوعات:

الموسوعة كلمة يونانية الأصل معناها حلقة كاملة من المعرفة وهي عبارة عن تجميع شامل لجميع فروع المعرفة الإنسانية أو لفرع واحد منها، من خلال عدد كبير من المقالات والدراسات المكتوبة من قبل متخصصين وإعلام في الموضوع.

وتعطى الموسوعات معلومات وافية وشاملة عن الموضوع لذا يمكن اعتمادها كنقطة البداية في البحث عن فكرة شاملة عن إي موضوع قيد الدراسة والتقصي. وزودت العديد منها بالكشافات الموضوعية لزيادة أهميتها وقيمتها المرجعية، وتنتهي أغلب المقالات بقائمة ببليوغرافية مختارة وقيمة بأهم المراجع والمصادر في الموضوع.[1]

وتعرف كذلك بأنها عمل يحتوى على معلومات عن جميع الموضوعات أو قد يكون قاصرا على ميدان أو موضوع متخصص، مرتب ترتيبا منهجيا(هجائيا في العادة)،لكن يمكن ترتيبها كذلك ترتيبا موسوعيا أو هرميا حيث ترتب المواد طبقا للمفاهيم العريضة أولا ثم بالموضوعات المتعلقة بها والتي تتفرع منها[2]

ففي الترتيب الهجائي هناك أسلوبان متميزان يهدف الأول إلى تقديم عدد كبير جدا من المقالات القصيرة لتغطية المواضيع الصغيرة وهنا تبرز الحاجة إلى العديد من الإحالات وإلى الكشافات والأدلة المرشدة للربط بين المواضيع المختلفة وتوحيد الأقسام المتعددة للمواضيع في

حين يهدف الأسلوب الثاني إلى المقالات المطولة التي تحوي بداخلها مواضيع أصغر وفي هذه الحالة تبرز الحالة إلى الكشافات التحليلية .

أما الترتيب المصنف حسب الموضوعات فان معظم دوائر المعارف التي ظهرت قبل بداية الطباعة كانت ترتب بهذه الطريقة وتقوم فكرتها على أساس تقسيم المعرفة البشرية إلى قطاعات معينة في العلوم والفنون وترتيبها تبعا لأهميتها أو العلاقات المتبادلة بينها سواء في الإطار العام للقطاعات أو في الترتيب الداخلي لتفريعات كل قطاع وهذا النظام المصنف ما يزال يستعمل في كثير من دوائر المعارف وخاصة دوائر المعارف المتخصص

وتقسم الموسوعات إلى نوعين:

-**العامة:** وهي التي تعالج مختلف مجالات المعرفة الإنسانية بشمولية وموسوعية.

-**المتخصصة:** وهي التي تشمل موضوعا واحدا أو موضوعات مترابطة ومتداخلة

3.بنية الموسوعات

3. 1. المداخل:

أن الموسوعات ليست مجرد قائمة لمفردات ومصطلحات وشرحها بل هي كيان هيكل شبكي ذو بنية داخلية معقدة يرتبط مع خارجه بعلاقات متعددة ويموج داخله بشبكة كثيفة من العلاقات المتداخلة، وتسمى المصطلحات والمفردات التي تقوم بشرحها وتناولها الموسوعات بالمداخل.

حيث تكون هذه المداخل عبارة عن الواصفات التي تشخص الحالة الموضوعية للمفاهيم وتنقسم إلى:[3]

-المصطلحات اللاتي تدل على أو تشير إلى المفاهيم.

-المصطلحات التي تدل على كيانات فردية وتسمى هذه أيضا المصطلحات أيضا بأسماء الأعلام أو الهويات وتتكون من:

-أسماء مشروعات.

-أسماء جغرافية

-أسماء هيئات ومؤسسات وأفراد

-اختصارات واستهلاليات.

-أسماء تجارية

-أحداث تاريخية

والمداخل المكونة للموسوعات تكون ما يعرف بهياكل البيانات والتي تقسم إلى ثلاث أنواع كبرى هي:

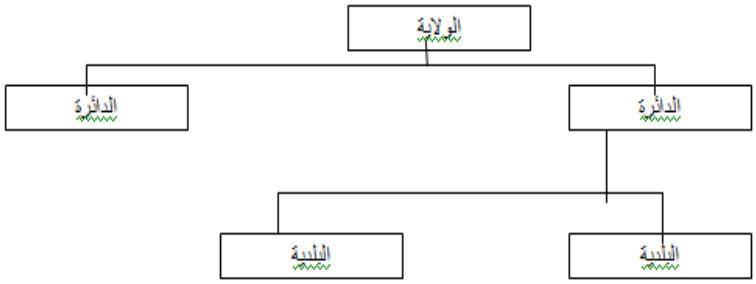
- سلاسل البيانات ذات الطابع المسطح:

وهي تلك التي تتساوى فيها جميع عناصر السلسلة إي لا يسود أي عنصر فيها على عناصر أخرى داخل السلسلة من أمثلتها :المداخل الكبرى في نفس المستوى وهي تعرف كذلك بالمصطلحات العامة في علم المكانز، حيث أن الموسوعات تكون محتوياتها مرتبة سواء هجائيا أو موضوعيا وفق مداخل تحوي مجموعة كبيرة من المعارف ويشترط فيها أن تكون مقيدة ومضبوطة.

- سلاسل البيانات ذات الطابع الهرمي(السلمي):

التي يسود فيها كل عنصر العناصر التي تدنوه في السلمية مثال ذلك سلمية الإداري في الجزائر.

الشكل رقم 01 يمثل عرض البيانات بطريقة سلمية

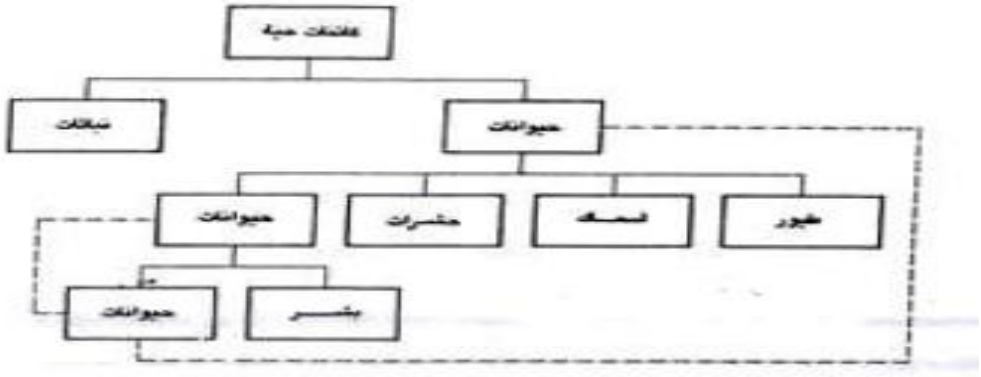


المصدر: من إعداد الباحث

- شبكات البيانات ذات الطابع الحلقي:

تزخر الموسوعات بالهياكل الحلقية، والهياكل السلمية الحلقية وذلك على مستوى المفردات وتعريفاتها ومسارات اشتقاقها، وخاصة على مستوى الاشتراك اللفظي كما هو مبين في المثال التالي[4]:

الشكل رقم 02 يبين طريقة عرض المعلومات بطريقة حلقية



المصدر:من إعداد الباحث

2.3 العلاقات:

ترتبط مداخل الموسوعات بعلاقات دلالية ويمكن نمذجتها في:

- **<u>علاقات التساوي:</u>**

عندما ننظر للمصطلحات على أنها متساوية (نفس المعنى أو نفس المعنى تقريبا) فانه يمكنها إن تضم في فئات تساوى، ومن ثم تحدد المصطلحات المتساوية لنفس المفهوم الواحد.[5] وتنتج هذه الظاهرة كثيرة عن الترادف في اللغة أو الكلمات الاستهلالية والمختصرات حيث أن للمداخل عدة كلمات تعبر عنها أو تكون مقاربة لها في المعنى مثل التلفون، الهاتف، المسرة فنقوم هنا باختيار كلمة واحدة كمدخل لتناولها في الموسوعة وباقي المصطلحات نستخدم الإحالات بين المصطلحات.

وعلى الرغم من أن المترادفات الحقيقة نادرة في اللغات العامة، إلا أنها أكثر شيوعا في اللغة العلمية وهي من الأنواع الخمسة:

- ✓ الأسماء التجارية.
- ✓ الجذور المختلفة.
- ✓ المصطلحات التي بطل استخدامها
- ✓ الاستخدام المحلي
- ✓ الاستخدام الشائع

- **العلاقات الهرمية:**

وهي العلاقة التي تعبر عن علاقات العلوية (الوضع في مرتبة أعلى) superordination والتابعية (الوضع في مرتبة أدنى) subordination ومن أنواعها: علاقة الشمول (الجنس) generic relation وعلاقة الجزء/الكل[6] part/whole relation.

فعلاقة الشمول نجد فيها أن المفهوم الشامل العريض يمثل طائفة من المفاهيم ويكون أحد هذه المفاهيم مفهوما مخصصا وضيقا عنه حيث يكون عضو من أعضائه ويختلف المفهوم المخصص عن المفهوم العام في خاصية واحدة على الأقل.

مثال ذلك: البنوك/البنوك الإسلامية

أما علاقة الجزء/الكل فان المصطلح الكبير يمثل طائفة أشياء أو مفاهيم أما المفهوم الضيق فإنه يمثل أحد أجزائه.

مثال: الجزائر/تلمسان

- **علاقات الاقتران(الارتباط):**

وهي التي تستخدم لتغطية العلاقات الأخرى بين المفاهيم المتصلة يبعضها البعض اتصالا وثيقا، حيث أن علاقة الترابط المتبادلة تكون بصفة دائمة أي أن تكون الإحالة المدخل ا متصل ب بالمدخل ب والعكس صحيح ولازم.[7]

مثال ذلك: التدريس/التعليم

3.3.الإحالات:

تستخدم الموسوعات الإحالات للانتقال ما بين المداخل ولزيادة شرح المادة العلمية المقدمة ولتبيان العلاقات بين المداخل، ونجد في الموسوعات الإحالات التالية:

-إحالة انظر:

وهي إحالة استبدال تستخدم للإحالة من الصيغ والمداخل غير المستخدمة إلى الصيغ أو المداخل المستخدمة، وبهذا تساعد على تجاوز مشكلة الترادف وتعدد الصيغ المعبرة عن المفهوم الواحد، ومثالها: الأحوال الجوية انظر الأرصاد الجوية، الطقس.

وتستخدم هذه الإحالة خاصة في الترادف من المدخل غير المستخدم إلى المدخل المستخدم ومن هجاء غير مستخدم إلى هجاء مستخدم.

-إحالة انظر أيضا:

ترد هذه الإحالة للربط بين رأسين كلاهما مستخدم وبينهما صلة أو علاقة معينة، وتأتي هذه الإحالات في القوائم والفهارس وذلك لتجاوز ظاهرة التشتت الحاصلة بسبب الترتيب الهجائي في رؤوس الموضوعات، وتأتي هذه الإحالة على ثلاثة مستويات:[8]

أ- الإحالة الهابطة أو النازلة: وهي الإحالة من الموضوع العام إلى الموضوعات الفرعية له والمنبثقة منه

ب-الإحالة الصاعدة وهي الإحالة من موضوع خاص إلى موضوع عام.

ج-الإحالة من موضوع إلى موضوع أخر مساو له في الدرجة وبينهما صلة الانتماء للموضوع الأم.

4. رقمنة الموسوعات

نتيجة تطور برمجيات الحاسوب واستخدامه في العديد من المجالات، وخاصة في مجال البحث العلمي وإنتاج المعرفة نشأت عنه مصادر رقمية أي مكتوبة بلغة (0،1)، ومن أهم هذه المصادر الموسوعات الرقمية والتي يمكن أن تنشا رقميا أي لا توجد لها نسخة ورقية مثل ويكبيديا، أو يتم تحويلها من الناحية الورقية إلى رقمية مثل موسوعة Britanica.

1.4. استخدام لغة الروابط التشعبية لهيكلة بنية الموسوعات الرقمية:

الروابط التشعبية هي الطريقة التي يتم الربط بها مباشرة بين معلومات مختلفة سواء كانت نصية أو غيرها، وقد تكون هذه المعلومات متواجدة في صفحة واحدة أو صفحات مختلفة باعتماد روابط تظهر في واجهة تقوم بعرض نص أساسي وتقنيات بصرية مثل اللون، الأيقونات تقوم بتحديد هذه الروابط، ونستفيد من هذه المعلومات بالضغط بالفأرة.[9]

في النص الورقي تعرض المعلومات في فقرات تتابعية والقارئ يعتمد هذا الترتيب لقراءة هذه الوثيقة بطريقة مستمرة لكن في النصوص ذات الروابط التشعبية تعرض المعلومات في فقرات منفردة ومنفصلة، وتعرض هذه الفقرات المرتبطة في عدة طرق: اختيارية، ارتباطيه، استمرارية، تنضيدية(stratification).

1 – الاختيارية: تربط مجموعة المعلومات المحتواة في الفقرات بفهرس اعتمادا عليه يحيلنا كل مدخل إلى شرح وافي ولا يرتبط هذا الشرح بأي وثيقة أخرى.

2 – الارتباطية: يختار القارئ المعلومات التي يحتاجها و تحيله إلى معلومات أخرى مرتبطة بها هذه الأخيرة ترتبط بمعلومات و فقرات أخرى ويستخدم كثيرا في الموسوعات الالكترونية .

3 – الاستمرارية: تعرض المعلومات بطريقة تتابعية مثل النص الورقي، فيقسم النص إلى مقاطع، وكل مقطع يحيل إلى المقطع الذي يليه.

4 - تنضيدية: تعرض المعلومات في فقرات تكون مرتبة على عدة مستويات، كل مستوى يحيل إلى مستوى تحته والعكس، وترتبط هذه المستويات مع بعضها البعض.[10]

وعرض النصوص على شكل فقرات مرتبطة تؤدي إلى تشكل ظاهرة التناص وهو عبارة عن بؤرة نصية مركزية تتقاطع من خلالها عدد كبير من النصوص المتزامنة أو السابقة حيث أن التناص هو تقاطع نصوص ووحدات من نصوص في نص أو نصوص أخرى.

ويبدو النص المترابط على هيأة ملف يتضمن بعض نقط الإنشاء والترسيخ بواسطة أزرار أو بواسطة مستويات معلومة تمكن من الوصول بفضل نقرة على الماوس وبفضل تحول الرابط إلى عقد (جمع عقدة) والتي هي عبارة عن وحدات للمعلومات ذات طبيعة نصية، مرئية وسمعية أو سمعية/بصرية وعندما تتآلف وتتآزر العديد من الوسائط في بنية ملف أو وثيقة فإننا في هذه الحالة نتحدث عن (تعدد وسائطية)[11].

وتعد لغة HTML من أهم لغات الروابط التشعبية المستخدمة في الانترنت وفي تنظيم المصادر الرقمية، لذا استخدمت في الربط ونمذجة البنية الهيكلية للموسوعات الرقمية والانتقال ما بين المداخل كما تظهر بكثرة في موسوعة ويكيبيديا، حيث أن هذه الروابط هي عبارة عن شبكة معلومات مكونة من مجموعة من العقد (nœuds)، ومجموعة أخرى من الروابط (liens)، كل عقدة تحتوي إما نص أو صور...الخ. وتقوم الروابط بتوضيح العلاقات بين المعلومات المعروضة وهذه الروابط تكون ثنائية وموجهة، وقد تكون بين كل أو أجزاء من الوثيقة المعبر عنها بعقد.

وادي استخدام لغة Html وتطوراتها أن أصبحت الموسوعات عبارة عن شبكات مغيرة طريقة القراءة الكلاسيكية إذ أن القارئ حاليا يقوم بالانتقال بين المعلومات التي يحتاجها دون أن يقرا النص تتابعيا كما في السابق.

2.4.من استرجاع المعلومات إلى البحث الدلالي:

نتيجة استخدام برمجيات الإعلام الآلي في رقمنة مصادر المعلومات عامة والموسوعات خاصة مكنت الباحثين من سرعة استرجاع المعلومات والبحث فيها بطريقة بسيطة حيث مكنتهم من :

- إمكانية البحث من خلال كشاف الكلمات المفتاحية خارج السياق.
- إمكانية البحث بالهمزة وبدونها.
- إمكانية البحث بالهاء أو التاء المربوطة.
- إمكانية البحث بال التعريف أو بدونها.
- إمكانية البحث بالمنطق البولياني وأدواته.
- إمكانية بتر مصطلح البحث.
- إمكانية البحث باللاوصفات.

لكن طريقة البحث الكلاسيكية السابقة لم تعالج الجانب الدلالي في المداخل والارتباطات التي تنشأ بين معلومات الموسوعات فتم إنشاء كشاف التباديل الذي يعد أداة مساعدة للوصول إلى المصطلحات في القسم الهجائي من المكنز حيث يستطيع المستفيد عن طريق هذا الكشاف الوصول إلى المصطلح الذي يبحث عنه في القسم الهجائي ولم يتمكن من الوصول إليه.

وتعتمد فكرة كشاف التباديل على استخدام كل كلمة مهمة من كلمات المصطلحات التي تتكون من كلمتين أو أكثر كمدخل في الكشاف وتدرج تحت هذه الكلمة جميع المصطلحات التي وردت فيها هذه الكلمة، وترتب المصطلحات ترتيبا هجائيا وبذلك يستطيع المستفيد معرفة صيغة المصطلح الذي يبحث عنه ويتمكن من الوصول إلى المصطلح في القسم الهجائي من المكنز.

مثال:

الشعر:

أدوات الشعر

أغراض الشعر

بحور الشعر

وبعد تطور برمجيات معالجة اللغة وخاصة الناحية الدلالية ظهرت طريقة جديدة للبحث في مصادر المعلومات الرقمية هي البحث الدلالي حيث يمثل هذا الأخير أحد خوارزميات البحث التي تأخذ في الاعتبار معاني الكلمات والمعنى السياقي للمصطلحات وليس فقط النَّمط المماثل للحروف، وعلى الرغم من النجاح الحالي لمحركات بحث الويب الموجودة الآن إلا أن عملية استرجاع المعلومات بصورته الحالية ما زالت تعاني قصورا يتمثل في غياب فهم كلمات البحث ومعناها في السياق. ومن المتوقع أن يقوم البحث الدلالي بتعويض هذا النقص من خلال استخدام خوارزميات البحث التي تأخذ في الاعتبار معاني الكلمات والمعنى السياقي للمصطلحات مما يبشر بفرصة أكبر لزيادة دقة نتائج البحث والحصول على المزيد من النتائج ذات الصلة.

فمن خلال فهم معنى الكلمات الموجودة في مصادر البحث، يتوقع أن تكون النتائج التي تنتج عن عملية البحث متصلة بصورة أكبر بكلمات البحث وأن المصادر التي لم يكن في الإمكان الحصول عليها في نتائج البحث لعدم احتوائها بصورة مباشرة على كلمات البحث - بالرغم من أنها ذات علاقة بها- سوف تظهر في المعلومات التي تم استرجاعها[12]

واحد أهم التقنيات التي يستخدمها البحث الدلالي هي الأنطولوجيا والتي تعرف بأنها خطة أو قائمة تتضمن استخدام الفئات الدلالية للمفاهيم المهمة في ميدان معين مع تعريف أو تبصرة توضيحية لكل مفهوم واليات لعرض علاقاته بالمفاهيم الأخرى.

وقدمت وثيقة لغة قائمة مفاهيم الويب تعريفا للأنطولوجيات علي أساس الاستخدام وذلك بأنها تقوم بتعريف المصطلحات المستخدمة في وصف وتمثيل أحد المجالات المعرفية، وعادة ما يستخدم قائمة المفاهيم الأفراد وقواعد البيانات والتطبيقات التي تحتاج إلي تقاسم أو مشاركة المعلومات فيما بينها، وتشمل قائمة المفاهيم التعريفات الحاسوبية للمفاهيم الأساسية في المجال المعني، مع تحديد طبيعة العلاقات بينها كما أنها تقوم بترميز المعرفة في مجال ما، وكذلك المعرفة المشتركة بين المجالات وهي بهذه الطريقة تجعل المعرفة قابلة لإعادة الاستخدام وفي هذا السياق[13]

فالأنطولوجيات هي وسيلة تحديد ورسم العلاقات الدلالية بين المفاهيم والمصطلحات حتى تكون قابلة للقراءة والمعالجة آليا وقابلة للفهم ليس فقط بالنسبة للبشر ولكن من قبل

الآلات أيضا مما يسهل على البرمجيات الحاسوبية فهمها وسهولة معالجة وتنظيم واسترجاع مصادر المعلومات التي وضعت قائمة المفاهيم من أجلها.

واستخدام الأنطولوجيات أدى إلي تحسين التكشيف التقليدي وحوله إلى تكشيفا دلاليا والذي يعتمد على إضافة كلمات توضح معنى كلمات جملة المساءلة باستخدام الأنطولوجيات ليتم التعبير عن الوثائق حيث يعتمد على الفرضية التالية "أن معاني المعلومات النصية يفهم اعتبارا من الروابط الدلالية والمفاهيمية الموجودة في التخصص العلمي الذي تتكلم عنه محتويات الوثائق.

3.4. نمذجة محتويات الموسوعات:

حتى نتمكن من الانتقال من الموسوعة ذات الكتلة الواحدة أي التي تعرض معلوماتها بطريقة تتابعية وترتبط محتوياتها بروابط تشعبية بسيطة، عمل الباحثون على إنشاء موسوعات ذات كتل متعددة وتعالج محتوياتها دلاليا حتى تمكن الباحثين من الاستفادة القصوى من معلومتها وهو ما يظهر في موسوعة دلامبير، لذا استخدمت لغة XML.

تنتمي لغة التحديد القابلة للامتداد eXtensible Mark up Language (XML) إلى عائلة لغات التحديد والتي يمكن تعريفها بأنها لغات لهيكلة النصوص وترميزها بشكل يسهل التعامل معها ومعالجتها آليا وهى لغات ترميز أكثر منها لغات برمجة بالمعنى المفهوم والمتعارف عليه لكلمة برمجة .

حيث نجد قاموس علم المكتبات والمعلومات المتاح على الخط المباشر ODLIS يعرفها كالآتي:
"مجموعة محددة مسبقا من الواصفات سواء كانت (رموز و تيجان) أو طريقة لتعيين وتحديد تلك الواصفات والتي تستخدم في إدماج وتضمين إيه معلومات خارجية داخل وثيقة نصية إلكترونية و عادة ما تشير إلى تحديد شكل تلك المعلومات أو لتيسير إجراءات التحليل التي تتم عليه".[14]

ولقد صممت لغات التحديد لتعمل مع برامج محددة إلا انه قد تم اعتماد لغة التحديد المعيارية العامة SGML كمعيار عالمي في عام 1986 وبعد ذلك خرجت منها لغة تحديد النص الفائق التي تستخدم في إنشاء صفحات الانترنت، وفى عام 1998 قامت رابطة الشبكة العنكبوتية بالتوصية باستخدام النموذج المصغر من لغة التحديد المعيارية العامة SGML

والذي عرف فيما بعد بلغة التحديد القابلة للامتداد XML وتختلف معظم لغات التحديد عن قواعد البيانات في قدرتها على تحديد العناصر المتضمنة داخل النص دون تمييز عناصر البيانات الهيكلة إلا أن لغة التحديد القابلة للامتداد لديها القدرة على تحويل أي نص إلى قاعدة بيانات متكافئة[15]

ذ أن مهمة لغات التحديد تتركز في ترميز المعلومات وصياغتها في بناء هيكلي موحد يسهل التعامل معه بواسطة كافة الأنظمة والتطبيقات فهي تعنى وتركز على شكل المعلومات وتحديد الإجراءات التي ستتم عليها.

وتعمل لغة الامتداد الموسع في هيكلة المعلومات وتسهيل تمثيلها واسترجاعها دلاليا وهو ما ينقلنا من الاسترجاع البسيط للمعلومات إلى الاسترجاع الذكي، وحفظ المعلومات الرقمية على المدى البعيد.

واستخدام لغة Xml في نمذجة الموسوعات أدى إلى تحويل الموسوعات إلى شبكة دلالية وهي آلية تحاول مشابهة التمثيل الإنساني للذاكرة لتمثل نموذج يبين كيف تمثل المعلومة في الذاكرة و كيفية الولوج ألبها، تمثل المعارف بواسطة عقد (nœud)ويربط بينها بواسطة روابط دلالية كما تعرف على أنها عبارة عن شبكة تمثل العلاقات الدلالية بين المفاهيم، غالباً ما تستخدم كطريقة لتمثيل المعرفة. تعتبر مخططاً أو بياناً مؤلفاً من عقد يمكن أن يكون موجهاً ويمكن أن يكون غير موجه، الفرق بين الشبكة الدلالية والمكنز أن الأخير يجمع مصطلحات وكلمات في ميادين متعددة، بينما في الشبكة الدلالية فيكون للمصطلح مفهوم واحد، حيث كل مفهوم يحدد بقيود (attribues) دلالية تكون خاصة بالعقد و وروابط تربطه بمختلف العقد .[16]

5. خاتمة:

من خلال تناولنا للبنية الداخلية للموسوعات وإسهامات برمجيات الحاسوب في نمذجتها، وجدنا أن هذه الأخيرة عملت على نقل الموسوعات الرقمية الأولى من موسوعات تشبه الموسوعات الورقية الى موسوعات رقمية ذات بنية متعددة تشمل النصوص والفيديوهات والصور مرتبطة فيما بينها باستخدام لغة الروابط التشعبية وتكون محتوياتها ذات بنية دلالية مما يجعل البحث عن المعلومات فيها يتم بكفاءة ودقة.

وهذه البنية الجديدة للموسوعات حولتها إلى شبكات دلالية ترتبط المداخل المكونة للموسوعات والمعلومات فيها مع إضافة توضيحات وشروح تسهل عملية البحث والاسترجاع

كما تمكن المستفيدين من هذه الموسوعات استرجاع معلومات دقيقة ويتم الانتقال بين هذه المعلومات المسترجعة باستخدام الروابط التشعبية والتي كانت في البداية بسيطة لكن ظهور Xml أدى إلى تحسين نمذجة الموسوعات وتطوير بنيتها الهيكلية التي تربط بين مختلف المداخل والمعلومات المشكلة لها.

6. الإحالة والتهميش:

[1] -عامر قنديلجي.مصادر المعلومات من عصر المخطوط الى عصر الانترنت.الأردن: دار الفكر، 2000.ص78.

[2] -موسوعة الشامي.الترتيب الموسوعي.

[3] -محمد فتحي عبد الهادي.المكانز كأدوات للتكشيف واسترجاع المعلومات.القاهرة: مكتبة غريب، 1989.ص56.

[4] -نبيل علي.اللغة العربية والحاسوب.القاهرة: تعريب للفكر العربي ،1988.ص488

[5] - محمد فتحي عبد الهادي.المرجع نفسه. ص56

[6] - محمد فتحي عبد الهادي.المرجع نفسه.ص60

[7] ا محمد فتحي عبد الهادي. المرجع نفسه. ص62

[8] -موسوعة الشامي.موضوع الاحالة

[9] - Vandendooren, Christian. Du papyrus à l'hypertexte : Essai sur les mutations du texte et de la lecture. Paris : La Découverte.1999. p113

[10] - Vandendooren, Christian. Ibid. p117

[11] ليوناردو اولاتس،.تعريب عبده حقي،.موت المؤلف بين النص المترابط والنص الايبرميدي. متواجد على الخط

https://sirat-hayat.page.تمت الزيارة يوم2022/04/15.

[12] -محسن الرشوان. المعالجة الآلية للنصوص العربية.السعودية: مركز الملك عبد الله لخدمة اللغة العربية،2019.ص47

[13] W3.Owl: web ontology language overview.[en line]www.w3.org/TR/owl-features.(visite le 14-04-2022)

[14]-MARC-XML.[en line] http://www.abc-clio.com/ODLIS/odlis_A.aspx- visite le20/02/2022

[15] - Site officiel du MARC-XML : http://www.loc.gov/standards/marcxml/

[16] . Saadani, L., Bertrand-Gastaldy, S. Cartes conceptuelles et thésaurus : essai de comparaison entre deux modèles de représentation issus de différentes traditions

disciplinaires. [En line] :http://www.cais-acsi.ca/2007call_fr.htm.(visite le 14/04/2022)

7. قائمة المصادر والمراجع:

1-المراجع العربية:

1. عامر قنديلجي، مصادر المعلومات من عصر المخطوط الى عصر الانترنت، الاردن، دار الفكر، 2000.

2. ليوناردو اولاتس، تعريب عبده حقي، موت المؤلف بين النص المترابط والنص الايبرميدي. متواجد على الخط ://sirat-hayat.page.tl

3. محسن الرشوان، المعالجة الآلية للنصوص العربية، السعودية، مركز الملك عبد الله لخدمة اللغة العربية،2019.

4. محمد فتحي عبد الهادي، المكانز كأدوات للتكشيف واسترجاع المعلومات، القاهرة، مكتبة غريب،1989.

5. موسوعة الشامي.

6. نبيل علي، اللغة العربية والحاسوب، القاهرة، تعريب للفكر العربي ،1988.

2-المراجع باللغة الأجنبية:

1. MARC-XML.[en line] http://www.abc-clio.com/ODLIS/odlis_A.aspx- visite le 20/02/2022

2. Saadani, L., Bertrand-Gastaldy, S**. Cartes conceptuelles et thésaurus : essai de comparaison entre deux modèles de représentation issus de différentes traditions disciplinaires**. [En line] :http://www.cais-acsi.ca/2007call_fr.htm.(visite le14/04/2022)

3. Site officiel du MARC-XML : http://www.loc.gov/standards/marcxml/

4. Vandendooren, Christian. **Du papyrus à l'hypertexte : Essai sur les mutations du texte et de la lecture**. Paris : La Découverte.1999

5. Owl: web ontology language overview. [en line]www.w3.org/TR/owl-features. W3**.** (visite le 14-04-2022)