الشبكات الدلالية العربية

ودورها في إثراء المحتوى العربي على الإنترنت

Arab semantic networks and their role in enriching Arabic content on the Internet

د. حنان أحمد فرج

الأستاذ المساعد بكلية الآداب جامعة المنصورة

الأستاذ المشارك بكلية الآداب جامعة الإمام عبد الرحمن بن فيصل

مستخلص:

تهدف الدراسة إلى إعطاء صورة لواقع الشبكات العربية الدلالية ودورها في إثراء المحتوى العربي على الإنترنت كما تهدف إلى التعريف بمفهوم الأنطولوجيا وعلاقتها بالويب الدلالي، كما ناقشت الدراسة الدوافع لبناء شبكات عربية دلالية والصعوبات التي تواجه بناء تلك الشبكات الدلالية مع شرح لواقع الشبكات الدلالية العربية والتعريف بنموذجين من الشبكات الدلالية العربية (Arabic WorldNet -Arabic Ontology).وأسفرت الدراسة على أن هناك ضعف شديد في بناء أنطولوجيا باللغة العربية، وتواجه بناء أنطولوجيا اللغة العربية تحديات نظراً لصعوبة بنائها والاتفاق على مكوناتها.

الكلمات الدالة: الأنطولوجيا، علم الوجود، الشبكات الدلالية العربية، المحتوى العربي .

Abstract:

The aim of this study to give a picture of the reality of the Arab semantic networks and their role in enriching the Arabic content on the Internet. It also aims to introduce the concept of ontology and its relation to the semantic web. The study also discussed the motives for building Arab networks, the difficulties encountered in constructing these semantic networks, explaining the reality of semantic networks. The study revealed that there is a serious weakness in building the Arabic language ontology, and the construction of the Arabic language ontology is facing challenges to the difficulty of building it and agreeing on its components.

مشكلة الدراســـة:

تعتبر شبكة الويب أغني المصادر المعلوماتية بما تحويه من مستندات ومعلومات ومصادر متنوعة يمكن الوصول إليها عن طريق محركات البحث التقليدية. غير أن تنظيم هذه المعلومات والمستندات بصورة تسهل عملية البحث فيها والوصول إليها، يعتبر أمراً غاية في الصعوبة يضاف إلى ذلك، أنه في ظل التزايد المستمر في حجم المعلومات المنشورة في شبكة الويب أصبح من الصعوبة بمكان قيام محركات البحث بإيجاد المعلومات المناسبة.

ومن هذه المشكلــة ظهرت فكـرة "الويب ذات الدلالات والمعاني اللفظيــة"، أو ما يطـلق عليـه بالإنجليزية مصطلح "Semantic Web " والتي هي امتداد للويب الحالية ولكن تختلف عنها بأنها تتفهم مدلولات الألفاظ والمعاني البشرية. ويمكن تمثيل الويب الدلالية بنسيج مترابط من المستندات التي تحتوي على معلومات محولة إلى معطيات يمكن للآلات قراءتها وفهم محتواها ومن ثم تقييمها.

وتعتبر الشبكات الدلالية (الأنطولوجيا) أحد أهم مكونات الويب الدلالية (Semantic Web) وحجر الأساس لعمله، وتعرف على أنها طريقة لتمثيل المفاهيم من حولنا وذلك عن طريق الربط بينها بعلاقات ذات معنى، مما يساعد على فهم أوسع للمفاهيم المختلفة، بهذه الطريقة نستطيع جعل الحاسب الآلي يصل لمستوى فهم وإدراك للمعاني قريب من فهم وإدراك الإنسان. (بامفلح، 2010) وعلى الرغم من صعوبة بناء الأنطولوجيات إلا أنها اصبحت من أهم التحديات التي تواجه بناء الويب الدلالي، وذلك لأن توفرها يعتبر بمثابة البنية الأساسية للويب الدلالي.

وتحظى الشبكات الدلالية بمكانـة مهمة في بيـئة المعلومات لكونـها من أكـثر الأدوات فاعلية في تنظيم المعرفـة، وبوصفها تمثيل رسمي للمفاهيـم والعلاقـات الموجـودة ضمن مجال معين (أحمد هـ.، 2013).

ولقد نمت الحاجة مؤخراً لوجود أنطولوجيا للغة العربية في كثير من التطبيقات والمجالات التي تتعامل مع مصطلحات اللغة العربية كمحركات البحث والترجمة الآلية وجعلها أكثر دقة وشمولية مما يستدعي الحاجة لوجود وسيط يقوم بتعريف المصطلحات العربية بمفاهيم دلالية تتيح المجال لفهم ومشاركة وتبادل البيانات بصورة واضحة ودون أي

غموض، حيث تحوي الأنطولوجيا على تعريف دقيق للمعنى الدلالي للبيانات المراد تبادلها، والمساعدة على الاستغلال الأمثل للمحتوى العربي المتوفر على الإنترنت والمساهمة في إثراء هذا المحتوى عليها.

ولتعزيز المحتوى الرقمي العربي على الإنترنت إنتاجاً واستخداما، ولتحسين فرص وصول جميع المواطنين العرب، يدفع إلى تطوير أنظمة وأدوات حتى نصبح قادرين على التواصل وتبادل المعلومات على الإنترنت الذي يعتبر حجر الزاوية لبناء نهضة وثقافة وحضارة الأمم. فوجود أنطولوجيا عربية من أجل التبادل السليم والفعال للبيانات يساهم بشكل مباشر في تعزيز المحتوى العربي على الإنترنت.

أهمية الشبكات الدلالية:

الشبكات الدلالية (الأنطولوجيا) تمثل الأساس لعملية تمثيل المعرفة في أي نظام، وتعمل على هيكلة المعرفة في المجال، وبدونها لا يمكن الخروج بمفردات لتمثيل المعرفة، وعليه فإن المفروض عمله أولاً هو وضع نظام فعال لتمثيل المعرفة تقدم مفرداته تحليل انطولوجي فعال لمجالها، والأنطولوجيا تسهل عملية تبادل المعرفة، ومشاركة الأنطولوجيا تسمح ببناء قواعد محددة للمعرفة الخاصة بحالات معينة وتزيد من إمكانية إعادة استخدام المعرفة.

ويذكر (الهادي، 2010) أن تنظيم المعلومات شهد العديد من التطورات الكثيرة خلال العقدين الأخيرين ارتبطت بتزايد واضح مستمر لمصادر المعلومات الإلكترونية المتاحة على الويب، مما أدى إلى ضرورة تطوير أو ابتكار أدوات ووسائل وأساليب جديدة لتنظيم المصادر في بيئة الويب. ويشير (النشرتي، 2012) إلى أن مصطلح التنقيب على شبكة الويب Web Mining له القدرة على اكتشاف المعرفة من واقع البيانات المتاحة على شبكة الويب مرتكزة في ذلك على ثلاثة قطاعات رئيسية:

- التنقيب عن محتوى الويب Web content mining.
- التنقيب اعتمادا على بنية الروابط المتاحة على الويب Web structure mining.
- التنقيب من واقع أنماط الافادة والاستخدام على الويب Web usage mining.

وتلعب الشبكات الدلالية (الانطولوجيا Ontology) دورا رئيسيا في كل من القطاعات الثلاث إذ تعمل على تمثيل المحتوى وفقا للمفاهيم والدلالات المراد التعبير عنها معتمدة في ذلك على أنماط المنطق والاستدلال الرياضي، كما تعمل على تحديد العلاقات والروابط بين البيانات المختلفة على الويب، وتعمل على تصنيف المحتوى لفئات وفقا لموضوعية استخدامه.

ويضيف (Jarra, 2011) ان الانترنت والاتصال السهل خلق بين الأنظمة حاجة ماسة ليس الى تبادل البيانات فقط، بل أيضا الى اتفاق حول معاني هذه البيانات Data Semantics ولذلك تعتبر الانطولوجيا الحجر الأساس للتبادل السليم والفعال للبيانات، حيث تحتوي على تعريف دقيق للمعنى الدلالي للبيانات المراد تبادلها. حيث تكتب هذه التعريفات بلغة المنطق Formal Logic كي يستطيع أي نظام فهمها وحسابها، بل والاستنتاج منها وبذلك فهي تعزز قابلية التشغيل بين الأنظمة في مجالات المعرفة المختلفة.

أهداف الدراسة:

تهدف هذه الدراسة إلى إعطاء صورة لواقع الشبكات العربية الدلالية وأثرها في إثراء المحتوى العربي على الإنترنت. وينبثق من هذا الهدف الرئيسي بعض الأهداف الفرعية المتمثلة في:

- إلقاء الضوء على مفهوم الأنطولوجي وأهميتها وعلاقتها بالويب الدلالي.
- التعرف على أهم الدوافع لبناء شبكات عربية دلالية .
- الصعوبات التي تواجه بناء الشبكات الدلالية العربية.
- واقع الشبكات الدلالية العربية .
- تعريف نماذج من الشبكات الدلالية العربية (Arabic WorldNet -Arabic Ontology)
- أثر الشبكات الدلالية في إثراء المحتوى العربي على الإنترنت.

أسئلة الدراسة:

س: ما هي أهم الدوافع لبناء شبكات دلالية عربية ؟

س: ما أهم الصعوبات التي تواجه بناء شكات عربية دلالية ؟

س: ما دور الشبكات الدلالية في إثراء المحتوى العربي على الإنترنت ؟

المنهج المستخدم:

لتحقيق الأهداف المرسومة سلفاً لهذا البحث، فقد تم استخدام المنهج الوصفي المعتمد على التحليل لكونه الأنسب لمعالجة الظاهرة قيد البحث، حيث يساعد على وصف واقع الشبكات العربية الدلالية والصعوبات التي تواجه بناء تلك الشبكات، وما لها من أثر في إثراء المحتوى العربي على الإنترنت .

وقد تم جمع المعلومات المطلوبة من خلال الاستعانة بأكثر من أداة، وذلك على النحو التالي:

- مراجعة الإنتاج الفكري المتعلق بالموضوع باللغة العربية والإنجليزية.
- جمع البيانات والمعلومات عن مشروعي تناولتهم الدراسة كنماذج للشبكات العربية الدلالية وهما شبكة الكلمات العربية Arabic WorldNet و الأنطولوجيا العربية Arabic ontology من خلال موقعي :

http://wordnet.princeton.edu

http://sina.birzeit.edu/ArabicOntology/index.html

الدارسات السابقة:

قد تم جمع المعلومات المطلوبة من خلال الاستعانة بأكثر من أداة، وذلك على النحو التالي

مراجعة الإنتاج الفكري المتعلق بالموضوع باللغة العربية والإنجليزية من خلال شبكة الإنترنت وعدد من قواعد البيانات .

في إطار تتبع أدبيات الإنتاج الفكري برزت العديد من الدراسات التي تناولت الشبكات الدلالية أو ما يطلق عليه الأنطولوجيا من بينها دراسة (Nouma N., 2013) التي ساهمت ف إبراز دور أنطولوجيا الويب الدلالي في تعظيم الخدمات التفاعلية وذلك من خلال بناء تطبيقات متقدمة في مجال التعليم العالي لتوفير بنية هيكلية ذات مغزى دلالي للمحتوى

تغطي الجوانب الرئيسية للموضوع بما في ذلك الطبقة العليا والتسلسل الهرمي للفئة الفرعية.

وتناقش دراسة (Calzada, 2007) القضايا والاتجاهات الرئيسية والمشاكل المتعلقة بأبحاث الأنطولوجيا، وتحديد المساهمات الممكنة من المتخصصين في مجال المكتبات والمعلومات للقيام بتطوير الأنطولوجيا الخاصة بالشبكة الدلالية، وذلك من خلال تقديم مفردات أكثر ثراءً يمكن أن تدعم الاستدلالات المنطقية كوسيلة لتحسين نظم استرجاع المعلومات .

وعرضت دراسة (Alkhalifa, 2006) التحديات والرؤى المستقبلية في هذا المجال موضحة أن التحديات اللغوية البحثية تكمن في تحديد النماذج اللغوية الدلالية للغة العربية ووضع التصنيفات اللغوية المختلفة للأسماء والأفعال وإجراء دراسات إحصائية لمختلف المكونات الدلالية لتحديد خصائصها في حين رأت الدراسة أن التحديات اللغوية التطبيقية تتمثل في التحليل النحوي ووجود محتوى لغوي كافي لاختبار هذه التطبيقات واستخدامها.

وتناولت دراسة (Nahli, 2004) مقارنة إرشادية لتحسين استخدام اللغة العربية في Arabic WordNet، تم تطبيق هذا النهج على WordNet العربية باستخدام قاموس AraMorph كمصدر ثنائي اللغة. وتلك الإرشادات عامة تمامًا بحيث يمكن تطبيقها بفعالية على اللغات الأخرى. ويتم التحقق من الكلمات الإنجليزية المستخرجة من المورد ثنائي اللغة مقابل Princeton WordNet من أجل تحديد تغطيتها وتحديد الكلمات التي تشترك في نفس مجموعة من synsets. وهذا يقلل بشدة من عدد الكلمات العربية المقابلة. ثم يتم فحص هذه الأخيرة مع المقابل في WordNet العربية لجعل الكلمات الجديدة تظهر وإضافة المرادفات الجديدة وإثراء اللغة العربية.

الدراسات العربية:

تناولت دراسة أحمد هندي (أحمد هـ، 2013) إلى المبادئ الأساسية المتعلقة ببناء الأنطولوجيا ومكوناتها وأنواعها، الدور الذي يمكن أن تقوم به في تنظيم المعلومات وتأثيرها في البحث والاسترجاع. وتوصلت الدراسة إلى ان الأنطولوجيا تساعد في التغلب على مشكلات تشتت وعشوائية المعلومات عن طريق معالجة النصوص بحسب دلالتها، وأن

تكون ادوات البحث مبنية على التمثيل المعرفي الدلالي والوصول لقواعد عامة يمكن تطبيقها في بناء أنطولوجيا المجال الواحد.

تتناول دراسة (نشرتي، 2011) بعنوان نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في إطار التنقيب عن البيانات مفهوم وواقع التكامل المعرفي علي الويب، وعملية التنقيب عن البيانات، ويعرض بعد ذلك الباحث لمفهوم الأنطولوجيات والبنية البنائية لها بالإضافة إلي لغات أنطولوجيا الويب ومستوياتها ووظائف الأنطولوجيات في التنقيب علي الويب، وتوصل الباحث إلى أن الأنطولوجيات توفر القدرة للتطبيقات والبرامج علي إدارة ودمج وتكامل المعرفة والمحتوي المتاح علي الويب من خلال استخدام المنطق وقواعد الاستدلال ويقترح الباحث محركات بحث دلالية عربية تتعاظم فيها قدرات البحث والاسترجاع وفقاً للمفاهيم والدلالات وكذلك إثراء المحتوي العربي بالأنطولوجيات العامة والمتخصص.

وتقدم الدراسة الخاصة (الهادي، 2010) بعنوان الأنطولوجيات ودورها في تنظيم المعلومات في بيئة الويب 2.0 مفهوم الأنطولوجيا وأوجهه الشبه والاختلاف بين الأنطولوجيا والوسائل الأخرى لتنظيم المعلومات مثل قوائم رؤوس المعلومات والمكانز وكذلك استخدامات الأنطولوجيا في علم المعلومات والويب الدلالي ومكوناتها وخطوات بناء الأنطولوجيات وأنواعها المختلفة ثم يعرض مجموعة من نماذج الأنطولوجيات المتكاملة مثل Bibliographic Ontology الأنطولوجيات الببليوجرافية، وهذه الأنطولوجية تقدم المفاهيم والصفات التي تصف الاستشهادات والإرشادات الببليوجرافية علي الويب الدلالي.

تتناول دراسة (جرار ج.، 2010) منهجية جديدة مقترحه في بناء الأنطولوجيات والتدليل على المعني يعتمد على صفات الأشياء وليس على أجناسها، وهذه الطريقة كما يذكر الباحثان أسهل على المستخدم، عند بناء أنطولوجيا بالمقارنة مع طريقة التصنيف، هذا بالإضافة إلى أن المستخدم لا يحتاج ضمن المنهجية المقترحة لاستعمال صفات جوهرية (وهي التحديد والتوصل إليها) حيث يكفي استعمال الصفات العرضية والجوهرية. كل ما عليه أن يفعله هو الإشارة إلى مجموعة الصفات التي تشير عند وجودها مجتمعه بشكل ضروري إلى شيء معين للتأكيد نحن لا نتحدث عن صفات جوهرية بل عن صفات مميزة والتي باجتماعها تشير إلى شيء دون غيره بالإضافة إلى سهولة المنهجية المقترحة، فإنها لا تؤدي إلى عدم اتساق .

وتتطرق دراسة (بكار ع.، 2010) إلى أهم الدوافع لبناء شبكات دلالية للعربية وتعريف الشبكات الدلالية وتناول البحث واقع بعض الشبكات الدلالية العربية، وأيضا تناول منهجية جديدة لبناء الشبكات الدلالية العربية باستخراج المفاهيم والعلاقات آليا او شبه ألياً من النصوص ومن مداخل المعاجم وقواعد البيانات المعجمية، والمنهجية المقترحة تعتمد على قواعد البيانات المعجمية وتمكن من اختزال الجهد والوقت، عدم التقيد بمجال معين .وتوصلت الدراسة الى بعض النتائج منها اللجوء غلى وسائل آلية او شبه ألية ضروري ويمثل تحديا. وأوضح أن البناء اليدوي للشبكات الدلالية باهظ الثمن ويستغرق وقتا كبيرا الأمر الذي يتطلب ضرورة اللجوء إلى وسائل آلية أو شبه آلية لأن أغلب المنهجيات تعتمد على التعلم الآلي من النصوص.

تتناول دراسة (بدر، 2002) بعنوان الأنطولوجيا وعلاقتها بعلم المعلومات والمكتبات تعريفات مصطلح الأنطولوجيا في الفلسفة وعلم المعلومات ثم تناولت الدراسة بناء الانطولوجيات وبعض استخداماتها، ثم عرض نماذج من الأنطولوجيات المتكاملة وتطبيقاتها مثل مشروع cyc ومشروع ميكروكوزموس Mikrokosmos وغيرها من المشرعات التي تهدف لبناء الأنطولوجيات، وتناولت الدراسة الأنطولوجيات الحديثة ومدى إفادتها من إنتاجية علم المعلومات والمكتبات، وتوصي هذه الدراسة بتطوير الأنطولوجيات وتطوير الأدوات والمنهجيات اللازمة لإنشاء وتصميم وتقييم الأنطولوجيات .

تعريف الأنطولوجيا Ontology:

هناك العديد من التعريفات التي تناولت الشبكات الدلالية أو ما يسمى الأنطولوجيا Ontology في علم المعلومات عبارة عن كلمة أعجمية مأخوذة من أنطولوجيا اليونانية ويتكون من كلمتين Onto وتعني الوجود، و Logei وتعني العلم؛ وعلم الوجود هو أحد مباحث الفلسفة وهو العلم الذي يدرس الوجود بذاته، الوجود بما هو موجود، مستقلاً عن أشكاله الخاصة، وكان أرسطو في القرن الرابع ق.م، أول من أدخل مفهوماً عن مثل هذه النظرية التي عنى بها العلم حول أعم قوانين الوجود «علم الوجود بما هو موجود» (الموسوعة العربية، 2018).

وأُعيد إنتاج هذا المفهوم في علم الكمبيوتر حديثاً بشكل تبوأ صدارة الأبحاث. فبعد انتشار الإنترنت انتشاراً واسعاً واستخداما في العديد من المجالات، نتجت حاجة ملحة لتوحيد

الأنظمة والبيانات الموجودة، وذلك لكي تتمكن هذه الأنظمة من التعامل فيما بينها للقيام بمهمة ما .

إن هذا النوع من التبادل البيني (Interoperability) للبيانات بين الأنظمة يحتاج إلى حل إشكالات كثيره لإنجازه، ليس فقط فيما يتعلق بالسرعة والسرية بل أيضا إلى اتفاق على طريقة التناقل وتركيب البيانات. والأهم من ذلك هو الاتفاق على المعنى الدلالي للبيانات المتبادلة ولحل إشكالية التعامل مع المعاني الدلالية للبيانات اقترح العلماء الأنطولوجيا كمرجع تعرف من خلاله معاني الأشياء المراد وصفها. بمعنى آخر، إن الأنطولوجيا هي تعريفات الأشياء المراد تداولها. (جرار م.، 2017).

ومنذ التسعينات أصبحت الأنطولوجيا من أهم المباحث في علم المعلومات، حيث أصبح لها مفهوم تفصيلي يضع معنى أو دلالة لكل مفهوم ومصلح ولفظ في ميدان معين، وتوضح هذه المفاهيم في نموذج بياني من المصطلحات والعلاقات الدلالية بين هذه المصطلحات (ويكيبيديا، 2018). وتعرف الشبكات الدلالية (الأنطولوجيا) بأنها توضيح مخصص وأكيد لمفهوم مشترك حيث يكون هذا التوضيح بواسطة مجموعة من المصطلحات مرتبة شجرياً في مجال معين، أي أن الأنطولوجيا تعتبر تطوراً في تمثيل المفاهيم. وهي تمثل علاقات دلالية بين المفاهيم، وغالباً ما تستخدم كطريقة لتمثيل المعرفة وتعتبر الشبكة مخطط موجهاً أو غير موجه مؤلفاً من عقد، التي تمثل المفاهيم بالإضافة إلى الخطوط.

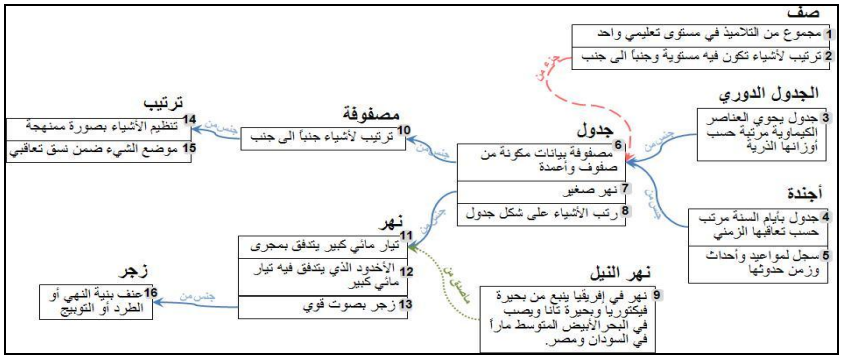
أي أنها عبارة عن " بيان منهجي أو نسقي للكيانات وعلاقاتها الموجودة في ميدان معين للدلالة على كتل البناء التي تستخدم لمساعدة الحاسبات والبشر في المشاركة في المعرفة " (الهادي، 2009) وهي تشبه قوائم رؤوس الموضوعات والمكانز في أنها تقوم على أساس تنظيم الكلمات في علاقات ترادفيه، ولكنها تختلف في تحليلها للفئات. وأطلق عليها البعض اسم خرائط المفاهيم أو خرائط المعلومات وقد أطلق عليها البعض الخرائط الذهنية.

أهم الدوافع لبناء شبكات عربية دلالية:

اللغة العربية هي اللغة الرسمية لمئات الملايين من الناس في عشرين دولة في الشرق الأوسط وشمال أفريقيا، وهي اللغة الدينية لجميع المسلمين من مختلف العرقيات في جميع أنحاء

العالم. والمثير للدهشة أن القليل قد تم في مجال الأنطولوجيا العربية. ومن ثم، فإنه يحفز على تطوير مورد لغوي عربي يكتشف ثراء اللغة العربية.

إن الأنطولوجيا العربية هي شجرة لمعاني المصطلحات العربية (انظر الرسم البياني أدناه) وهي تمثيل رسمي للمفاهيم التي ينقلها المصطلح العربي لكل مصطلح في اللغة العربية، يتم تحديد مجموعة من معانيها (أي المفاهيم)، والعلاقات الدلالية (مثل نوع فرعي من وجزء من) بين جميع المفاهيم. (حول اللغة العربية، 2018). ولذلك ينبغي تعزيز سبل التواصل بين الجهات التي تقوم بتنفيذها منعاً للتكرار والازدواجية وتعميم استخدام المصطلحات العلمية التي وضعتها المؤسسات اللغوية والمعاجم المصطلحية.



الشكل رقم (1) نموذج لأنطولوجيا اللغة العربية

الشبكات الدلالية العربية تعمل على الإستغلال الأمثل للمحتوى العربي على الويب من خلال جعله أكثر تعبيراً ودلالة حيث يعتمد على تكشيف الكيانات في صورة شبكة ترتبط مع بعضها البعض من خلال العلاقات بين المفاهيم المعبرة عنها داخل قطاعات موضوعية Classes ثم تحديد الروابط بين هذه الكيانات وإعادة هيكلتها في صورة تسمح للبرمجيات والآليات التعامل معها من خلال قدرة هذه اللغة على التعامل بصورة منطقية.

وتسهم في تطوير نظم استرجاع المعلومات بحيث تكون مبنية على التمثيل المعرفي الدلالي للمفاهيم والعلاقات في مختلف مجالات المعرفة مما يعمل على تحسين نتائج نظم استرجاع المعلومات المعتمدة على باللغة العربية. (أحمد هـ.، 2013).

تستعمل الأنطولوجيات العربية في العديد من التطبيقات الحديثة مثل (Jarrar, 2006):

• البحث عن المعلومات والاسترجاع دلالياً، حيث تستطيع محركات البحث فهم الدلالة المقصودة واسترجاع بيانات أكثر قربا للمعنى، باستخدام الأنطولوجيا، بدلاً من الفهم الحرفي للاستعلام، كما هو دارج حالياً.

• الترجمة الآلية، حيث أن المفاهيم في الأنطولوجيا مرتبطة بمثيلاتها من اللغات الأخرى ربطاً مفاهيمياً دلالياً، تستخدم الأنطولوجيا في المترجمات الآلية لتحسين جودة الترجمة.

• التأويل وفك الغموض، التطبيقات التي تحتاج لفهم معاني الكلمات في نص ما وفك الغموض الناتج عن المشترك اللفظي، تستعمل المفاهيم في الأنطولوجيا كمرجع معياري يساعدها على فكّ هذا الغموض.

• تكامل البيانات وقابلية التشغيل البيني التي يمكن أن تستخدم فيها الأنطولوجيا العربية كمرجع دلالي للعديد من نظم المعلومات المستقلة، أي استعمال المفاهيم في الأنطولوجيا كمرجع دلاليّ معياريّ مشترك بين أنظمة المعلومات.

• الويب الدلاليّ، حيث يتم ربط الكينونات والمسمّيات الواردة في صفحات الإنترنت بمفاهيمها في الأنطولوجيا، وذلك كي تتمكن محركات البحث من فهم دلالة البيانات ومسمّياتها وإزالة الغموض المعاني المستخدمة في مواقع الويب؛ بين العديد من التطبيقات الأخرى.

وقد أفرزت التجارب والممارسات العملية ثلاث تحديات تواجه تطبيق الويب الدلالي في مجال اللغة العربية، تتمثل في ضرورة بناء هيكل واضح لتمثيل البيانات الوصفية، وآليات اختيار المفردات المستخدمة لتوصيف البيانات والربط فيما بينها بعلاقات محددة، وأخيراً توصيف محتوى كل صفحة من صفحات البوابة الرقمية. وللعمل على تخطي هذه التحديات تتضمن الأنطولوجيا نماذج بيانات Data model تشتمل على استخدام تقنيات متعددة لتمثيل وعرض البيانات ومنها لغة لتنسيق تبادل البيانات ومعيار هيكلة المحتوى XML ومخطط العلاقات ومن أبرز آلياتها مخطط إطار وصف المصدر RDF، ولغة أنطولوجيا الويب OWL (أحمد أ.، 2015).

مفهوم الويب الدلالي وعلاقته بالأنطولوجيا :

يعتبر الويب في هذه الأيام المصدر الأوسع للمعلومات في جميع المجالات والمواضيع، وهو في توسع دائم سواء في حجم المعلومات أو في حجم المستخدمين. إن هذا الحجم الكبير من المعلومات والتوسع السريع في الوب يجعلان إدارة المعلومات والوصول إليها أمراً في غاية التعقيد؛ وقد عانى معظمنا من العدد الكبير من الوثائق والروابط التي يعيدها أي محرك بحث رداً على أية عملية استعلام، هذا عدا عن عدم توافق نتائج البحث مع طلب المستخدم. ولهذا نحتاج إلى تعديل بنية الويب الحالي تعديلاً جذرياً للوصول إلى بنية أكثر ملاءمة للتعبير عن الوثائق والاستعلام عنها (بدر، 2002).

وطرح مفهوم الوب الدلالي كتوسعة للبنية الحالية للويب، تُعرّف الموارد فيه باستخدام لغة توصيفيه تسمح بالتعبير عن المحتوى الدلالي للمورد بطريقة يمكن للآلة تفسيرها. وقد جرى اقتراح عدد من اللغات والآليات للتعبير عن المحتوى الدلالي للمورد، واستخدامه لاحقاً في عمليات التصنيف والبحث، غير أن كل آلية من هذه الآليات تهتم بعدد من الجوانب التطبيقية وتغفل جوانب أخرى. لذا، جرى توحيد جميع هذه اللغات والآليات في بنية موحدة أكثر شمولية وديناميكية هي الأنطولوجي، والتي تتيح التعبير عن المحتوى الدلالي للموارد تعبيراً فعالاً يفهمه الإنسان (باستخدام اللغات الطبيعية) والآلة على حد سواء، إضافة إلى الإمكانات الواسعة التي تقدمها هذه البنية للاستعلام عن الموارد (Oscar, 2003).

وتعرف الويب الدلالية (أو ما يطلق عليها أحيانا "الويب ذات الدلالات اللفظية" أو "الويب ذات المعنى") (Semantic Web) على أنها "شبكة بيانات بالمعنى، أي أنه يمكن للبرامج الحاسوبية الخاصة أن تعرف ماذا تعني هذه البيانات". ويتطلب الوصول لهذه الطريقة من التفسير والفهم للبيانات الاستعانة بالأنطولوجي (Ontology)، والذي يعرف على أنه طريقة لتمثيل المفاهيم وذلك عن طريق الربط بينها بعلاقات ذات معنى، حتى تسهل ربط الأشياء الموجودة بعضها البعض ولفهم أوسع للمفاهيم المختلفة، وبهذا تعتبر الأنطولوجيا (ontology) أحد أهم مكونات الويب الدلالية وحجر الأساس لعمله، وتعرف على أنها طريقة لتمثيل المفاهيم وذلك عن طريق الربط بينها بعلاقات دلالية واضحة ذات معنى، مما يساعد على فهم أوسع للمفاهيم المختلفة. واعتباراً لما سبق يتبين إن حجر الأساس في

مفهوم الويب الدلالي يتطلب أن يكون للمفاهيم والمصطلحات معان واضحة، وعلاقات مفهومة وبهذه الطريقة نستطيع جعل الحاسب الآلي يصل لمستوى فهم وإدراك للمعاني قريب من فهم وإدراك الإنسان ولهذا تعتبر الأنطولوجيا العمود الفقري للويب الدلالي (بامفلح، 2010) .

تعتبر الأنطولوجيا تطوراً في تمثيل الكلمات والمفاهيم حيث تساعد في التغلب على مشكلات تشتت وعشوائية وفوضى المعلومات المتاحة في بيئة الويب وذلك عن طريق معالجة النصوص وتجزئتها بحسب دلالاتها بحيث يصبح النص مجموعة من العلاقات الدلالية الواضحة بين المفاهيم والمصطلحات الواردة بالنص (غنيم، 2013).

وقد اشار (عارف، 2015) إلى الويب الدلالي وعلاقته بالأنطولوجيا أن الويب الدلالي هو البيئة التي تعمل فيها الأنطولوجيا وتؤثر في قدراتها الدلالية بما تتيحه من امكانات فائقة في التعرف على المعاني والدلالات للألفاظ بطريقة تحاكي الفهم البشري.

الصعوبات التي تواجه بناء شبكات دلالية عربية :

تواجه بناء أنطولوجيا اللغة العربية تحديات نظراً لصعوبة بنائها والاتفاق على مكوناتها ونظراً لكون أدوات المعالجة اللغوية المتوافرة غير فعالة كفاية لمعالجة النصوص العربية التي كانت نادراُ ما تتضمن علامات الترقيم الصحيحة المساعدة على التحليل الصحيح للجمل.

وتشغل الدلالة ضمن المعالجة الحاسوبية للغة العربية موقعا مهما، فلا يمكن أن نتصور إنجاز معالجات عميقة للنصوص العربية، بدون معلومات كافية عن دلالة الألفاظ والعلاقات الدلالية بين الكلمات المكونة لتلك النصوص. وتتطلب المعالجة الدلالية الحاسوبية للغة العربية أن نأخذ أمرين بعين الاعتبار وهما: (داود، 2001).

أولا: ضرورة الاستناد إلى إطار لساني صوري قادر على توصيف مستويات اللغة الهدف.

ثانيا: ضرورة الاعتماد على مقدار كبير من المعلومات التي ترتب وتبوب أساسا انطلاقا من قواعد بيانات آلية.

إن بناء شبكات عربية دلالية يستوجب أخذ بعين الحسبان خصائص الكلمات ومدلولاتها، ومن ثم إيجاد حل ناجع للكثير من القضايا اللغوية التي تواجه المعالجة الحاسوبية للغات الطبيعية من قبيل اللبس الدلالي للمفردات والعبارات والنصوص، والترجمة الآلية، وبنوك المعطيات وغيرها من التطبيقات الآلية.

يجب إذن، ألا يغفل واضعو شجرة المفاهيم عامة قضايا مثل الترادف والمشترك اللفظي والتضاد، وغيرها من العلاقات الدلالية، لما لها من قيمة مضافة تتمثل في إزالة اللبس الدلالي والتركيبي، ومن ثم صياغة معمارية للأنطولوجي وفق بنية قوية ومحكمة مستمدة من المفاهيم اللسانية الخوارزمية، وبالتالي التمكن من حل مشاكل التحليل اللغوي الآلي بصفة عامة (داود، 2001).

ويذكر (جرار م.، 2017) أن غالبية قواميس اللغة العربية التي يعتمد عليها في جمع واستنباط التعريفات والمعاني لا يمكن استخدامها كونها تركز على تصريفات الكلمات، وغالباً ما تخلط بين التصريف اللغوي بالمعنى الدلالي، بل أن معظمها يدلل على المعنى بأمثلة إيمائية، ولا يحدد المعنى تصريحاً مباشراً.

تعتبر الأنطولوجي المعبرة عن مجال معين واحدة في جميع اللغات، ويكون الاختلاف في التعابير التي ترتبط بكل مفهوم، لذلك من الممكن أن نعرف مفهوماً C1´ مكافئاً للمفهوم C1 يرتبط بمجموعة التعابير المعبرة عن الوثيقة في اللغة الإنكليزية مثلاً؛ وبذلك يمكننا استخدام الأنطولوجي في تطبيقات أخرى مثل الترجمة الآلية (من حيث المبدأ).

أما بناء الأنطولوجي الخاصة بكل مجال فهو يجري إلى الآن يدوياً، ويستغرق الكثير من الجهد والوقت؛ وذلك لصعوبة الاتفاق على بنية موحدة لكل مجال، إضافةً إلى صعوبة تحديد حدود كل مجال من المجالات، وقد ظهر مؤخراً عدد من التطبيقات القادرة على استخراج الأغراض وفقاً لأنطولوجي معينة، من مجموعة من الوثائق الإلكترونية. ولعل من أهمها طريقة اقتُرحت لبناء الأنطولوجي آلياً هي دراسة (فاضل الحسن) وهي تقترح بناء الأنطولوجي من النصوص التي تصف مجالاً معيناً (الحسن، 2012).

ووفق الإطار العام للممارسات العملية لا تتوافر منهجية موحدة لبناء الأنطولوجيا، فقد تم تصور العديد من المنهجيات وفقاُ للمجال والبرمجيات المستخدمة والخبرات والمعارف

المتاحة، ومع ذلك يمكن استشراف ثلاثة أساليب لبناء الأنطولوجيا يمكن حصرها في الآتي (أحمد أ.، 2015):

- الأسلوب اليدوي: ويتألف من توسيع الأنطولوجيا القائمة بالفعل أو إنشاء أخرى جديدة، وذلك ينفذ في المستوى العلوي من الأنطولوجيا .
- الأسلوب الالي المعتمد على الأتمتة: ويستند إلى استخراج المعلومات بشكل آلي من مصدر المعلومات وتنظيمها وفق مخطط المفاهيم.
- الأسلوب المختلط: يعتمد على استخدام كلتا الطريقتين السابقتين، وهذا يعني بناء الأنطولوجيا اعتماداً على التقنية وتصور امتدادها يدوياً.

ويعتمد بناء وتنظيم الأنطولوجيا على الخبرة في المجال الموضوعي وتعبئة المفردات والمصطلحات والمترادفات والمفاهيم والعلاقات وذلك وفق مخطط المفاهيم للحصول على الحقائق باستخدام الاستدلالات والتعريف بالقواعد ثم اختيارها وأخيراً تنفيذ وتطبيق الأنطولوجيا (أحمد أ.، 2015).

نماذج من الشبكات العربية الدلالية:

تعتبر أنطولوجيا اللغة الإنجليزية (WordNet) بداية لإنشاء الأنطولوجيا اللغوية، وأخذت زخما كبيراً ليس فقط في التطبيقات، ولكن في مجالات الأبحاث، حيث لحق بها عدداً كبيراً من اللغات الأخرى مثل (الفرنسية والألمانية والإيطالية ....الخ) والجدير بالذكر أن هذه الأنطولوجيا ليست مجرد شجرات مفاهيمية منفصلة عن بعضها، بل لقد تم الربط بينها، وذلك من خلال ربط المفاهيم المتطابقة عبر اللغات المختلفة، أي نفس المفهوم يعطي رقما واحداً للتعبير عنه بغض النظر عن لغة المنشأ.

وبالنسبة للغة العربية، لم تتم أي جهود جادة لبناء أنطولوجيا لغوية حتى الآن، ما عدا مشروعان الأول (Arabic World Net) قامت به وكالة الاستخبارات الأمريكية (CIA) حيث اتبع الباحثين أسلوب ترجمة أنطولوجيا اللغة الإنجليزية، بما يقابلها باللغة العربية. ويذكر مصطفى جرار (Jarrar, 2011) إن منهجية الترجمة لا يمكن أن تقدم نموذجاً فعالاً، وذلك كون المفاهيم اللغوية تعبر عن أنساق فكرية وثقافية لا يمكن أن تتطابق في بنية إنتاجها، ودلالاتها بشكل حرفي، وذلك بالرغم من اشتراك اللغات في عدد كبير من المفاهيم. والمشروع

الثاني الأنطولوجيا العربية Arabic ontology وهو مشروع قامت به جامعة بيرزيت بفلسطين .

الأنطولوجيا العربية Arabic Ontology :

قامت جامعة بيرزيت بفلسطين بإعداد الأنطولوجيا العربية وهو مشروع يهدف الى تصنيف معاني الكلمات العربية وإعادة تعريفها وفق قواعد جديدة، ويستطيع الأنسان وكذلك الحاسوب استعمالها. وتعتبر خطوة هامة في تاريخ اللغة العربية فهي تؤسس لطريقة جديدة لتعريف معاني ودلالات الكلمات، بمعنى آخر، الأنطولوجيا هي شجرة معاني الكلمات العربية. أي أنها تصنيف لمفاهيم الكلمات وليست تصنيف للكلمات، حيث تم تحديد الصفات الجوهرية والمميزة لكل مفهوم وتصنيف هذه المفاهيم بناءً على الصفات الجوهرية المميزة لكل مفهوم دون غيره، وهي بذلك تتيح بناء علاقات أوسع بين المفاهيم، لا تقتصر على المرادفات والمعاكسات كما درجت المعاجم التقليدية (بيرزت، 2011) .

بمعنى آخر تقدّم الأنطولوجيا (١) تعريفاً للمعاني وتصنيفها تصنيفاً شجرياً بالاعتماد على الصفات الجوهرية والمميزة للمفهوم دون غيره، كما أن تصنيف المعاني يعتمد على مفاهيم رياضية ومنطقية جديدة - المنطق الوصفي ونظرية المجموعات. (٢) تعريفاً بالمفاهيم وإعادة صياغة دلالتها بناءً على ما توصلت إليه العلوم الطبيعية والإنسانية الحديثة وليس اعتمادا على ما شاع بين الناس.

أما في المعاجم الكلاسيكية فيُعرّف المفهوم عادة بمقاربته بأقرب المعاني إليه، دون تحديد صفاته الجوهرية اللازمة لتمييزه عن غيره، أو بعبارة أخرى، دون وضع الحدود الفاصلة بين معنى وآخر، فضلا عن أن تلك المعاجم، تميل لإعطاء المعاني وفق شيوعها بين الناس، وهو ما يعني تقديم المعنى الاستعمالي الغامض للمفاهيم، على المعنى العلمي الدقيق (جرار م.، 2017).

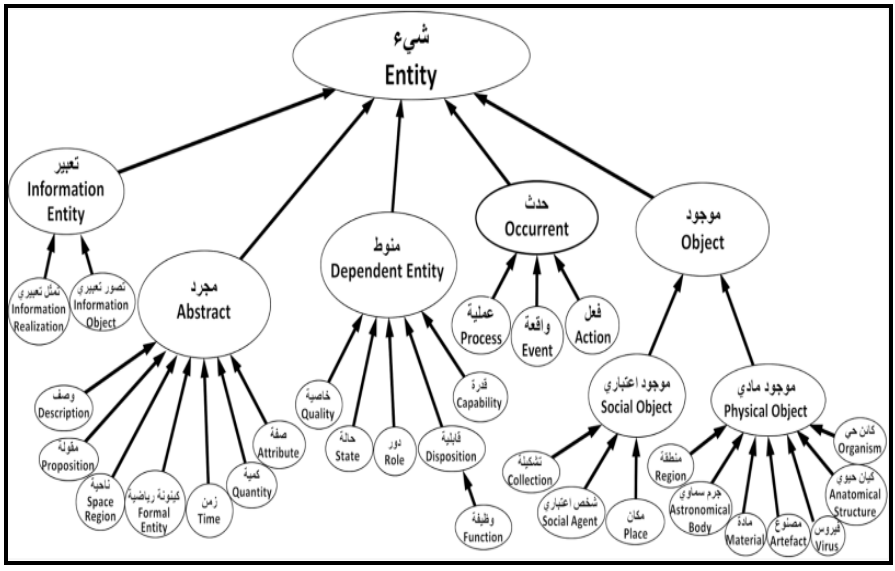
وقد تم جمع واستنباط تعريفات ومعاني من القواميس العربية المتاحة (ما يقرب من ثلاثين ألف مفهوم) وإعادة صياغتها وهندستها كتعريفات دلالية (جرار م.، 2017).

الأنطولوجيا تبدأ بمستويات عليا من المفاهيم حيث تشكل الانطولوجيا العليا معيارا لضبط الانطولوجيات الدنيا، والحكم على جودتها، وصولا لربط بعضها ببعض، ومن ثم

فالأنطولوجيا العليا. وهي أكثر المفاهيم تجريداً في اللغة العربية، حيث أنها تحمل في تعريفاتها جملة من الصفات الجامعة لكل المفاهيم اللغوية، وتنبثق المفاهيم في المستويات الأدنى تباعاً، بحيث (تُورِث) المستويات العليا صفاتها للمستويات الأدنى، التي تحمل بدورها صفات إضافية تضفي عليها تحديداً أكثر مما سبقها، وفي الوقت ذاته تميزها عن غيرها في المستوى ذاته.

ولقد تم ضمن مشروع الـ Arabic Ontology في جامعة بيرزيت بناء الشجرة العليا للمفاهيم العربية وذلك لتكون هذه الشجرة (وهي مكونة من عشر مستويات، أي ما يقرب من 400 مفهوم هي المستويات العليا للأنطولوجيا العربية. ولقد أعتمد في بناء هذه الشجرة العليا على أساس الاستفادة من المنهجيات المستخدمة في العديد من الأنطولوجيات اللغوية الاخرى، وخاصة سومو(SUMO)The Suggested upper merged Ontology، ودوتشيA Descriptive DOLCE Ontology for Linguistic and Cognitive Engineering، وبفو (BFO)، فضلا عن إسهامات الفلاسفة. (Jamal Daher, 2010) .

الشكل التالي يوضح المستويات الثلاث الأعلى في الشجرة.



شكل رقم (2) المستويات العليا في الأنطولوجيا العربية

فلسفياً، تفرّق الأنطولوجيا بين المفهوم (concept) ومصادقيه (instances)، حيث أن لكل مفهوم مصاديق، تصدق صفاته، ولا يوجد مفاهيم بدون مصاديق، وكل مفهوم يتم تعريفه إما بالحدّ او بالرسم، كما في الفلسفة والمنطق .

ما هي المعلومات الموجودة في الأنطولوجيا وكيف يتم تمثيلها؟

كل مفهوم في الأنطولوجيا يُعطى (١) رقماً فريداً يدلّ عليه، (٢) مجموعة الكلمات التي تُسمي هذا المفهوم، مترادفات Synonyms، (٣) تعريفاً مختصراً يحدد الصفات المميزة لهذا المفهوم دون غيره (Gloss)، أول كلمة في تعريف المفهوم تكون هي المفهوم الأعلى، فمثلاً يبدأ تعريف مفهوم "موجود مادي" بكلمة "موجود"، ويلي ذلك الصفات المميزة التي تخصص هذا المفهوم من المفهوم الأعلى. وهذا يعني أن المفهوم هو مجموعة الصفات المحددة في تعريفه بالإضافة إلى الصفات المحددة في المفهوم الاعلى منه، وهكذا. إضافة لذلك، يتم ربط المفاهيم ببعضها بعلاقات دلالية، اهمها علاقة النوع (Sub/Super Type Of) و علاقات دلالية أخرى مثل علاقة جزء-من، العلاقة السببية أو العلاقات الناتجة عن الأفعال .. وغيرها.

ما العلاقة بين الأنطولوجيا وشبكة المفردات (WordNet)؟

وعلى الرغم من أن شبكات المفردات تشبه الأنطولوجيا في الشكل، إلا أنها مختلفة، فشبكة المفردات هي عبارة عن مجموعات مترادفات بينها علاقات فقط، فهي لا تعرّف أو تصف المفاهيم بناءً على المعارف العلمية المثبتة او المتفق عليها كما هو موجود في الأنطولوجيات، وإنما تعكس ما شاع استعماله لدى العامة، وفي ذلك غموض دلالي كبير. كما أن شبكات المفردات لا تعرّف المفاهيم بصفاتها الجوهرية/المميزة، وبالتالي لا يمكن توريث صفات للمفاهيم الأدنى كما هو الحال في الأنطولوجيات. إضافة إلى ذلك فإن المستويات العليا في الأنطولوجيا العربية تم تأسيسها علمياً وفلسفياً ومنطقياً وبدقةٍ عالية، وتم تحقيقها عبر عدة سنوات، لكي يتم اشتقاق المعاني منها، وهذا يضمّن صحة التصنيف في المستويات الأدنى.

كل مفهوم تم إضافته الى الانطولوجيا، تم ربطه بمفهوم يقابلـه في شبكة المفردات WordNet، وتم بناء خوارزمية تعتمد مبدأ (Graph Algorithm)، هذه الخوارزمية تعمل على ترجمة الكلمة إلى مرادفاتها في الانجليزية ثم يتم ترجمة الكلمات الانجليزية إلى العربية.

مشروع Arabic WorldNet: –

فقد تم بناء المشروع بالاعتماد على نفس هيكل ومحتويات شبكة الكلمات الإنجليزية Princeton WordNet ((PWN، وشبكة الكلمات الدلالية الأوروبية Euro WordNet. ويسهل عملية البحث والتصفح داخل الشبكة وعرض الترجمات العربية للمفاهيم مع مقابلاتها من خلال البحث والملاحة عبر (PNW/ EWN)، ويعمل ترميز المفاهيم بالأرقام والروابط على سهولة التعرف على تلك المفاهيم التي تم ترجمتها بالفعل إلى العربية، فضلاً عن تلك التي تم تعيينها لترجمتها خلال مشروع(AWN) .وهي قابلة للتطبيق على PNW2 مما يسهل عملية الترجمة إلى اللغة الإنجليزية وعشرات من اللغات الأخرى المتوافرة في PNW2. وتضم AWN 11269 مفهوم و23481 كلمة (Arabic WorldNet, 2018) .

الهيكل والتنظيم:

نظراً لأنه ينبغي أن تكون شبكة الكلمات الدلالية العربية (AWN) قابلة للبحث والملاحة عبر شبكة الكلمات الإنجليزية PNW وشبكة الكلمات الأوروبية EWN، إما بشكل مباشر أو غير مباشر من خلال الفهرس البيني ILI(Interlingua Index) أو الأنطولوجيا. فقد تم تصميم قاعدة بيانات وواجهة مستخدم بلغات متعددة. ((Arabic WorldNet, 2018.

وتتكون قاعدة البيانات من أربعة أنواع من الكيانات الرئيسية التالية (البند – الكلمة – النموذج – الرابط) بما في ذلك المعاني، وطبقات الأنطولوجيا مع معلومات فريدة ومميزة مثل التعليق أو الوصف على المفاهيم. أشكال الكلمات المستخدمة هي الجذور أو جمع التكسير. (Arabic WorldNet, 2018)

تحديث قاعدة البيانات:

يتم تحديث قاعدة البيانات تلقائياً بطريقة آلية أو شبه آلية بالاعتماد على ترجمة المفاهيم الأساسية الواردة في شبكة الكلمات الإنجليزية PNW وشبكة الكلمات الأوروبية EWN، ويتم

التركيز على الترميز اليدوي للمفاهيم والعلاقات الأكثر تعقيداً وأهمية. ويتم تكوين مجموعة من المفاهيم الأساسية المشتركة Common basic concepts (CBC)بين 12 لغة بشبكة الكلمات الإنجليزية والأوروبية مع أهم المترادفات، وهي جزء لا يتجزأ من الإطار الدلالي للشبكة .

ويتم توسيع WordNet الأساسية شبه تلقائيا من أعلى إلى أسفل، بالنظر للمفاهيم الأساسية المشتركة (Common basic concepts (CBC؛ ويتم ترجمة المفاهيم الخاصة باللغات الأخرى إلى أقرب معنى باللغة العربية.

وواجهة التحرير تعمل على التواصل مع خادم قاعدة البيانات باستخدام SOAP (Simple Object access Protocol) والسماح بتكوين قاعدة بيانات مشتركة من المفاهيم بالاعتماد على موارد معجمية مختلفة. (Jarrar, 2011).

وقد تم تمييز الأنواع التالية من المفاهيم الأساسية:

1.المفاهيم الأساسية المشتركة: Common Base Concepts (CBC)

المفاهيم التي تعمل كمفاهيم قاعدة في لغتين على الأقل.

2- مفاهيم القاعدة المحلية: Local Base Concepts (LBC)

المفاهيم التي تعمل كمفاهيم أساسية بلغة واحدة فقط؛

3- مفاهيم القاعدة العالمية: Global Base Concepts (GBC)

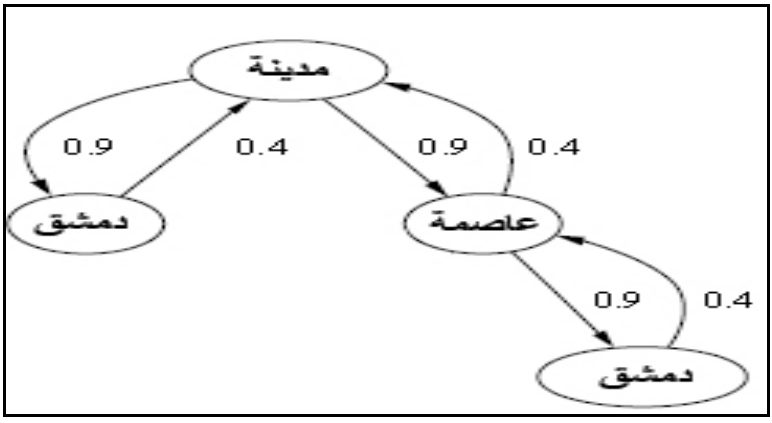
المفاهيم التي تعمل كمفاهيم أساسية في جميع لغات العالم .

الأنطولوجيـــا:

يتم الربط بين المفاهيم بروابط أنطولوجية (دلاليـة) مضبوطة في الشكل والمفهوم. ويتم بناء المستويات العليا من الأنطولوجيا بالاعتماد على الأنطولوجيا الأجنبية SUMO(The suggested upper merged Ontology وتحتوي على 1000مصطلح و 4000مفهوم، وعملية الربط تتم خارجيا بعد عملية التصميم النهائي ليس كما هو الحال مع الأنطولوجيا العربية Arabic 0ntology.

القواعد الأساسية لاختيار المفاهيم:

1- العلاقات: المفاهيم مبنية على قدر كبير من الترابط، ويتم ذلك من خلال الاعتماد على سلاسل hyponymy، فعلاقة التخصيص في AWN تستخدم أنطولوجية باعتبار أن هيكلتها تعتمد على العلاقة Subclass of/ Instance of (IS-A) وتعتمد قيمة الموثوقية بين المفهومين طرفي العلاقة ضمن ANW على عدد الوصلات بين المفهومين (Oscar, 2003).



الشكل رقم (3) مقتطف من أنطولوجية AWN

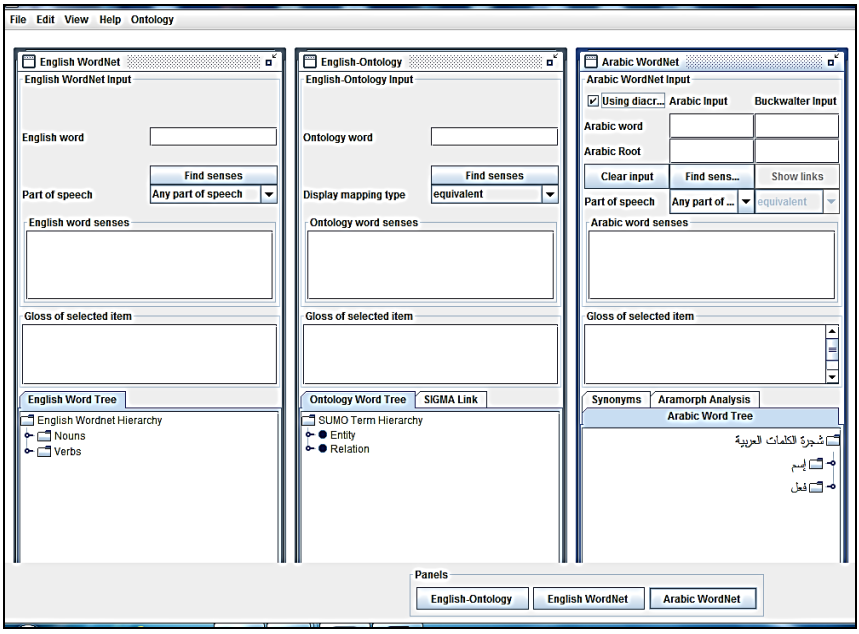
2- الصلة: إضافة المفاهيم الأساسية والأكثر استخداما،والبارزة باللغتين العربية والإنجليزية لها الأولوية. (علي، 2013)
3- العمومية: يبدأ بالجنس الأعلى للمفهوم. (علي، 2013)
4- إضافة الكلمات والمجلدات التي يمكن ربطها بسهولة إلى PNW مع علاقات التكافؤ (تلقائيا) المستمدة من القواميس ثنائية اللغة. (علي، 2013).

واجهة المستخـــدم:

الواجهة باللغتين العربية والإنجليزية بالإضافة إلى تسهيلات للبحث والتصفح تقدم للمستفيدين حيث تتوافر مجموعة من المكونات لكل منها مزاياه النسبية مثل التوافق مع described in black و Elkateb لأنهما يمكنهما التعامل مع النص العربي، ومع ذلك فقد

افترضت AWN نموذج بيانات مختلف تماماً، حيث الكلمات العربية مرتبطة بشكل مباشر مع المعاني في

PNW. (Arabic WorldNet, 2018)



الشكل رقم (4) واجهة مستخدم أنطولوجية AWN

عيوب الشبكة الدلالية العربية: (بكار ع.، 2010)

- التغطية المحدودة لشبكة الكلمات العربية مقارنة مع مثيلتها الشبكة الإنجليزية.
- الروابط والعلاقات بين المفاهيم أقل كثافة عنها في الشبكة الإنجليزية .
- بعض المفاهيم العربية غير موجودة في الإنجليزية والعكس .
- اختلاف في الجنس لبعض المفاهيم المترجمة بين العربية والإنجليزية .
- الاعتماد على الترجمة كمنهجية لبناء شبكة دلالية عربية لا يمكن أن تقدم نموذجاً فعالاً، وذلك كون المفاهيم اللغوية تعبر عن أنساق فكرية وثقافية لا يمكن أن تتطابق في بنية إنتاجها، ودلالاتها بشكل حرفي.

- العلاقات الدلالية يشوبها بعض الغموض الدلالي .
- يتم ربط المستويات العليا من الأنطولوجيا بالأنطولوجيات العامة مثل SUMO خارجيا بعد استكمال بنائها.

واقع الشبكات الدلاليـــة العربيـــة:

ظهرت في السنوات العشر الأخيرة الكثير من التطبيقات التي تعتبر فيها الأنطولوجيا بالغة الأهمية مثل الحكومات الإلكترونية والتجارة الإلكترونية، محركات البحث، المكتبات الرقمية ...وغيرها من التطبيقات، وذلك لكي تتمن هذه الأنظمة من التكامل فيما بينها، فمثل هذا النوع من التبادل (Interoperability) للبيانات بين الأنظمة يحتاج إلى الاتفاق على المعنى الدلالي للبيانات المتبادلة مما يتيح تبادل آلي للبيانات مبني على قواعد موحدة للفهم. (أحمد هـ.، 2013).

هناك العديد من تطبيقات الأنطولوجيا والتي تعتمد في في بنائها على أكثر من طريقة، فهي تعتمد على المجال والتطبيق والبرنامج المستخدم. فالكلمة في الأنطولوجيا التطبيقية لها معنى واحد في سياق التطبيق ولها هدف معين، كما يجب أن تبنى العلاقات حسب الأكثر احتمالية والأقرب في مجال اهتمامها. أما في الأنطولوجيا الدلالية يتم الربط بين معاني الكلمات والعلاقات فيها بين المعاني وليس بين الكلمات، أي الكلمات والمعاني تبنى في استقلالية عن تطبيق معين. (بدر، 2002).

وبالرغم من كون معظم تطبيقات الأنطولوجيا تركز على مجالات محددة كما سلف تبيانه، إلا أن هناك توجهات حديثة لبناء أنطولوجيا لغوية، خاصة بعد نجاح مشروع Arabic WordNet والذي يعتبر الآن أنطولوجيا شاملة للغة العربية، ولكن كونها تعتمد على الترجمة من شبكة الكلمات الدلالية WordNet كمنهجية لبناء شبكة دلالية عربية لايمكن أن تقدم نموذج فعالاً يتطابق في بنيته ودلالاته بشكل حرفي مع اللغة العربية. ولذلك جاء مشروع بناء أنطولوجيا للغة العربية في جامعة بيرزيت، والذي يحتاج إلى جهود متضافرة ومتعددة لتكون أنطولوجيا شاملة للغة العربية.

ويذكر بن حمادو وبكار أنه: يبقى الاهتمام بالشبكات الدلالية العربية محدود للأسباب التالية:

- إن اهتمام المخابر العربية بموضوع الشبكات الدلالية محدود وأهم الأعمال التي أنجزت تمت في مخابر أجنبية، الأعمال البحثية التي تتعلق بأنطولوجيا اللغوية العربية غير مكتملة ومنقولة من شبكات أجنبية عبر الترجمة (بكار ع.، 2010).
- تحتاج أغلب تطبيقات معالجة اللغة الطبيعية وخاصة اللغة العربية إلى التعامل مع المستوى الدلالي للغة. لهذا السبب ظهرت محاولات لبناء أدوات معالجة دلالية للنصوص من خلال بناء شبكات دلالية كالأنطولوجيا العربية والتي تحتاج سنوات من الجهد لإنجازها بشكل شمولي. (بكار ع.، 2010).
- وجدير بالذكر أن غالبية قواميس اللغة العربية التي يعتمد عليها في جمع واستنباط التعريفات والمعاني لبناء الشبكات الدلالية؛ لا يمكن استخدامها كونها تركز على تصريف الكلمات وغالباً ما تخلط التصريف اللغوي بالمعنى الدلالي، بل أن معظمها يدلل على المعنى بأمثلة إيمائية ولا يحدد المعنى تصريحاً مباشراً. (جرار م.، 2017).

ويتم اختيار القواميس وفقاً للمعايير التالية: (أحمد هـ.، 2013)

- أن يكون القاموس غير معتمد في بناءة علي تصريف الكلمات وذلك لأن الأنطولوجيا لا تهتم بتصريف الكلمات بل تهتم بالمعاني الدلالية وتعدد هذه المعاني لكل مصطلح.

- أن يكون المعني الذي يحدده القاموس واضحا غير مخلوط بمعاني أخري و أن يكون التعبير عن المعني فقط بكلمات توضح هذا المعني

- جودة التعريف وطريقة تركيبة وأن يكون مكتوب بطريقة واضحة وبكلمات صحيحة.

دور الشبكات الدلالية العربية في إثراء المحتوى العربي على الإنترنت:

يرجع اهتمام المواقع العالمية باللغة العربية لأنه يتكلم بها أكثر من 360 مليون متحدث حول العالم أي ما يعادل 7% من تعداد السكان العالمي وكذلك لأنها إحدى اللغات الست الرسمية للأمم المتحدة .كذلك للأهمية الاستراتيجية للشرق الأوسط وكذلك لازدياد الوزن

التجاري والديبلوماسي والإعلامي للدول العربية. فمنذ 2001 قامت الحكومة الأمريكية ومؤسسات أمريكية أخرى بإطلاق عديد من المواقع باللغة العربية لتعميق التواصل مع العالم العربي. كما أن ازدياد عدد مستخدمي الإنترنت في العالم العربي أدّى إلى زيادة هذا الاهتمام (علي، 2013).

وبداية من عام 2009 لوحظ ارتفاع في عدد المواقع العالمية التي تدعم اللغة العربية مثل موقع فيسبوك وآخر المواقع التي تم تدعيم اللغة العربية فيها هو اليوتيوب المملوك لعملاق البحث جوجل.وجود هذا الكم الهائل من الإنسان المستهدف من العالم العربي يدفع إلى تطوير أنظمة وأدوات حتى نصبح قادرين على التواصل على الإنترنت (ويكيبيديا، 2018).

ولتعزيز المحتوى العربي على الإنترنت إنتاجاً واستخداما لدعم التنمية المستدامة و ولتحسين فرص وصول جميع المواطنين العرب للمحتوى رقمي ذو جودة عالية وبأقل تكلفة، إضافة إلى دعم زيادة توافر واستخدام المحتوى الرقمي العربي وخلق بيئة مواتية للإنتاج الدائم على الإنترنت؛ كان لابد من توفر الأدوات التي تساهم بشكل مباشر في تسهيل عملية التبادل هذه. التحول إلى مجتمع المعرفة. (ويكيبيديا، 2018).

فوجود أنطولوجيا عربية قادرة على التبادل السليم والفعال للبيانات ينتج مزيداً من الأدوات التي تسهم بشكل مباشر في تعزيز المحتوى الرقمي على الإنترنت، وإغناء الإنترنت بمحتوى عربي يمكن المستفيدين من إيجاد المصطلح العربي الذي يناسب ويواكب متطلباتهم وبترجمات متعددة. مع الأخذ بعين الاعتبار إلى إشكالية صناعة المحتوى الرقمي من حيث خصوصية اللغة العربية وكتابة الحروف والتشكيل واتجاه الكتابة وقلة مشاكل محركات البحث واللغة العربية.

ولذلك ينبغي إجراء مسح شامل لجميع مشاريع المعالجة الآلية للغة العربية الجارية داخل الوطن العربي وجمعها في قاعدة معطيات واحدة، وتعزيز سبل التواصل بين الجهات التي تقوم بتنفيذها منعاً للتكرار والازدواجية وتعميم استخدام المصطلحات العلمية التي وضعتها المؤسسات اللغوية والمعاجم المصطلحية.

الشبكات الدلالية تعمل على إثراء محتوى الويب من خلال جعله أكثر تعبيراً ودلالة حيث يعتمد على تكشيف الكيانات في صورة شبكة ترتبط مع بعضها البعض من خلال

العلاقات بين المفاهيم المعبرة عنها داخل قطاعات موضوعية Classes ثم تحديد الروابط بين هذه الكيانات وإعادة هيكلتها في صورة تسمح للبرمجيات والآليات التعامل معها من خلال قدرة هذه اللغة على التعامل بصورة منطقية.

في ظل كثرة وفوضى المعلومات في بيئة الإنترنت والتي تؤدي إلى صعوبة في إيجاد المعلومات الهادفة والمفيدة؛ تساعد الشبكات الدلالية في التغلب على مشكلات تشتت وعشوائية وفوضى المعلومات المتاحة على الإنترنت، وتسهم في تطوير نظم استرجاع المعلومات بحيث تكون مبنية على التمثيل المعرفي الدلالي للمفاهيم والعلاقات في مختلف مجالات المعرفة مما يعمل على تحسين نتائج نظم استرجاع المعلومات. (أحمد هـ.، 2013).

يمكن استثمار الشبكات الدلالية في التنقيب على البيانات لتحقيق مهام ووظائف التنقيب حيث أظهرت تقنيات التنقيب ومنها زواحف الويب، ونماذج هيكلة البيانات ...الخ محدودية في معالجة البيانات نظراً لعجزها الكامل عن توفير الدلالات والمفاهيم المرجوة في معالجة كل من نص المحتوى والروابط بين كيانات الويب، ومن هنا تأتي أهمية الشبكات الدلالية في توفير أطر مفاهيمية ودلالية وعلائقية تعمل من خلالها على توصيف المعلومات وتحديد أوجه العلاقات بين كيانات المعلومات في سياق أو مجال موضوعي محدد مما يوفر فرص عظيمة لفهم وتبادل المعلومات على السواء بين البشر وبرمجيات الويب، ومن ثم عمل التنقيب على الويب على توظيفها لسد ما تتسم به تقنياته من عجز على صعيد الدلالات والمفاهيم. (نشرتي، 2011)

وفي سياق تحقيق التكامل المعرفي يتم استثمار الشبكات الدلالية في أطار التنقيب على الويب لتقديم الكثير من المزايا والفوائد هذه المزايا تتمثل في: (نشرتي، 2011)

1- التكامل المعرفي على صعيد محتوى الويب: إذ تعمل الانطولوجيات على تحويل شبكة الويب بمحتواها وروابطها وهيكلها الى مستودع منظم للبيانات يكفل القدرة على البحث واسترجاع المحتوى وفقا للمفاهيم والدلالات وتعمل الانطولوجيات على اكساب البرمجيات والتطبيقات القدرة على استنباط حقائق جديدة من واقع المصادر المتاحة من خلال استخدام المنطق وقواعد الاستدلال.
2- تحسين البحث على الويب Improved search to Web data: حيث تمكن الانطولوجيات من تحسين عمليات التكشيف في أدوات البحث وذلك في ظل ما

توفره من محددات للمحتوى ومحددات للروابط وتقديمها لدلالات ومفاهيم تعمل كقيمة مضافة لبيانات الويب، وتنقل عملية التكشيف داخل أدوات البحث من التكشيف اللفظي الى التكشيف المعتمد على المفاهيم والعلاقات بين الكيانات المختلفة.

3- القدرة على الإبحار في محتوى بصورة أكثر تكاملية **Better browsing capabilities:** تكفل الانطولوجيات القدرة على الإبحار في المحتوى اعتمادا على المفاهيم والعلاقات المنطقية التي تربط بين الكيانات المختلفة بدلاً من الاعتماد على الروابط الفائقة وحسب في عملية الابحار والتصفح لمحتوى الويب.

4- إضفاء الطابع الشخصي في إتاحة البيانات **Personalization of Web data access:** من خلال محددات من البيانات تضاهي الملف الشخصي لاهتمامات مستفيد ما، ويتم ذلك الأمر من خلال تقيم توصيات بصفحات أو مواقع او محتوى يضاهي هذه الاهتمامات، أو عن طريق تصفية واستبعاد صفحات الويب او المحتوى والتي لا تضاهي احتياجات المستفيد.

نتائج الدراســـة:

في ضوء الأهداف التي تضمنتها الدراسة جاءت النتائج على النحو الآتي:

- الجهود لبناء أنطولوجيا باللغة العربية ضعيفة جداً بخلاف مشروعي Arab WorldNet Arab Ontology.
- الأنطولوجيا تعتبر من أحد أهم مكونات الويب الدلالي وحجر الأساس لعمله.
- من أهم دوافع بناء شبكات عربية دلالية:
  - أ- الإستغلال الأمثل للمحتوى العربي على شبكة الإنترنت وجعله أكثر دلالة .
  - ب- تحسين نتائج نظم إسترجاع المعتمدة على اللغة العربية .
  - ج- تستعمل الأنطولوجيا العربية في العديد من التطبيقات مثل الترجمة الآلية – الاسترجاع الآلي ....الخ .
- ندرة توافر قواعد ومبادئ أساسية موحدة لبناء الأنطولوجيا .

- يواجه بناء أنطولوجيا اللغة العربية تحديات نظراً لصعوبة بنائها والاتفاق على مكوناتها
- أدوات المعالجة اللغوية للغة العربية المتوافرة غير فعالة كفاية لمعالجة النصوص العربية غير المشكولة التي كانت نادراً ما تتضمن علامات الترقيم الصحيحة المساعدة على التحليل الصحيح للجمل .
- الشبكة الدلالية العربية Arab WorldNetكونها تعتمد على الترجمة لا يمكن أن تكون نموذج فعال يتطابق في بنيته ودلالاته مع اللغة العربية.
- يمكن استثمار الشبكات الدلالية في التنقيب على البيانات لتحقيق مهام ووظائف التنقيب والتكامل المعرفي على الويب.
- الشبكات الدلالية تعمل على إثراء محتوى الويب من خلال جعله أكثر تعبيراً ودلالة من خلال الربط بين المفاهيم المعبر عنها داخل قطاعات موضوعية والروابط بين هذه الكيانات بصورة تسمح للبرمجيات والآليات التعامل معها.

التوصيات:

- وضع قواعد ومبادئ موحدة وأساسية لبناء الأنطولوجيا.
- إطلاق المشاريع والأبحاث والمساهمات المستقبلية لبناء أنطولوجيا شاملة للغُة العربية .
- تطوير الأدوات والمنهجيات اللازمة لإنشاء وتصميم وبناء أنطولوجيا للغة العربية .
- بناء إطار دلالي للغة العربية يعمل على إغناء تطبيقات الحاسب التي تتعامل مع مصطلحات اللغة العربية كمرحات البحث والترجمة الآلية وجعلها أكثر دقة وشمولية.
- إجراء مسح شامل لجميع مشاريع المعالجة الآلية للغة العربية الجارية داخل الوطن العربي وجمعها في قاعدة معطيات واحدة، وتعزيز سبل التواصل بين الجهات التي تقوم بتنفيذها منعاً للتكرار والازدواجية.

المراجـــع:

1. Arabic WorldNet. (2018, فبراير). Retrieved from http://wordnet.princeton.edu
2. حول اللغة العربية (2018, 1 8). Retrieved from: http://sina.birzeit.edu/ArabicOntology/index.html
3. Calzada, P. .. (2007). Catologies and the semantic web: problems and perspectives for LIS professionals. Probiems and Perspectives for lis professionals, 303-311.
4. Jamal Daher, M. J. (2010). Towards A Methodology For Building Ontologies - Classify By Properties. (In Arabic). In proceedings of the 3rd Palestinian International Conference on Computer and Information Technology. Palestine.
5. Jarrar, M. (2006). Towards The Notion Of Gloss, And The Adoption Of Linguistic Resources In Formal Ontology Engineering. In proceedings of the 15th International World Wide Web Conference, (pp. 497-503).
6. Jarrar, M. (2011). Building A Formal Arabic Ontology (Invited Paper). In proceedings of the Experts Meeting On Arabic Ontologies And Semantic Networks. Alecso, Arab League. Tunis.
7. M, E. S. (2006).
8. M, E. S. (2006). Arabic WordNet and the Challenges of Arabic. Proceedings of Arabic NLP/MT.
9. Nahli, R. D. (2004). dictionary, Enhancing Arabic WordNet with the use of Princeton WordNet and a bilingual.
10. Nouma N., B. N. (2013). A stop towards Better Understanding and development of Universitu Ontology. Education Domain International Science congress Association, 2(10), 57-60.

11. Oscar, C. a.-L.-P. (2003). Methodologies, tools and languages for building ontologies. Where is their meeting point. Data & Knowledge Engineering, 46, 41–64.

12. أحمد ,أ .ف .(2015) .أنطلوجيا الويب الدلالي ودورها في تعزيز المحتوى الرقمي :دراسة في المفاهيم والبنية الهيكلية والخدمات التفاعلية في البوابات الدلالية للتعليم الالكتروني .مجلة المكتبات والمعلومات العربية.(4)25 ,

13. الحسن ,ف .(2012) .نواة نظام تعلّم الأُنتولوجي من النصوص العربية. ommunications of the Arab Computer Society, 5(1).

14. الموسوعة العربيةRetrieved from http://www.arab-ency.com .(2018) .

15. النشرتي ,م .(2012) .نحو التكامل المعرفي من واقع توظيف الإنطولوجيا في إطار التنقيب عن البيانات. المؤتمر الثالث والعشرون للإتحاد العربي للمكتبات والمعلومات: الحكومة والمجتمع والتكامل في بناء مجتمع المعرفة. وزارة الثقافة الفنون والآثار: قطر والإتحاد العربي للمكتبات والمعلومات (أعلم) .

16. الهادي ,م .ف .(2009) .الإتجاهات الحديثة في التحليل الموضوعي للمعلومات وموقف قطاع المعلومات العربي منها. المؤتمر العشرون للإتحاد العربي للمكتبات والمعلومات . المغرب.

17. الهادي ,م .ف .(2010) .الأنطولوجيا ودورها في تنظيم المعلومات في بيئة الويب .2الملتقى العربي الرابع لتكنولوجيا المكتبات والمعلومات :الجيل الثاني للويب وتأثيره على قطاع المكتبات والمعلومات .pp) ,ص ص .(13-1مصر.

18. بدر ,أ .(2002) .الأنطولوجيا وعلاقتها بعلم المعلومات والمكتبات .مجلة المكتبات والمعلومات العربية.(1)22 ,

19. بكار ,ع .ا .(2010) .الشبكات الدلالية والأنتولوجيا العربية الواقع والآفاق. دمشق: جامعة صفاقس.

20. بكار ,ع .ا .(2010) .الشبكات الدلالية والانطولوجيا العربية :الواقع والآفاق .ورشة عمل إثراء المحتوى الرقمي على الإنترنت. دمشق.

21. جامعة بيرزت. (2011). انطولوجيا اللغة العربية. تم الاسترداد من http://sina.birzeit.edu/ArabicOntology/index.html

22. جرار ,ج .ط .(2010) .نحو بناء منهجية لبناء هندسة الأنطولوجيا :التصنيف بالصفات. المؤتمر الفلسطيني الدولي الثالث للكمبيوتر وتكنولوجيا المعلومات.(6-1 .pp) ,
23. جرار ,م .(2017, 10 7). Retrieved from http://www.jarrar.info/publications/J11.pdf.
24. داود ,م .م .(2001) .العربية وعلم اللغة الحديث .القاهرة :دار غريب.
25. علي ذيب الاكلبي، و محمد جعفر عارف. (ابريل 2015). دور الانطولوجيا في دعم محركات البحث الدلالية في البيئة العربية. مكتبة الملك فهد الوطنية، 25(4).
26. علي ,ن .غ .(2013) .بناء نواة نظام مساعد على إنشاء أنطولوجية عربية انطلاقاً من النصوص .مجلة جامعة دمشق للعلوم الهندسية.(1)29 ,
27. غنيم ,م .ا .(2013) .بناء نواة نظام مساعد على إنشاء انطولوجيا عربية إنطلاقاً من النصوص .مجلة جامعة دمشق للعلوم الهندسية.(1)29 ,
28. فاتن سعيد بامفلح. (2010). محركات البحث الدلالي في ظل تطبيقات الويب الدلالي. المجلة العربية للأرشيف والتوثيق والمعلومات، 1(29).
29. نشرتي ,م .س ,2011) .ديسمبر .(الشبكة العنكبوتية الدلالية :هوية تبحث عن الوجود: دراسة تأصيلية تحليلية. .(27) Cybrarians Journal
30. هندي عبد الله هندي أحمد. (2013). .بناء أنطولوجيا علم المكتبات والمعلومات في بيئة الويب الدلالية :دراسة وصفية تحليلية. / أعمال مؤتمر الإتحاد العربي للمكتبات والمعلومات (أعلم) الرابع والعشرون / مهنة ودراسات المكتبات والمعلومات: الواقع والتوجهات المستقبلية، (الصفحات 463-485).
31. ويكيبيديا .(2018). .المحتوى_العربي_على_الإنترنت Retrieved from https://ar.wikipedia.org/wiki .