منهجية مقترحة لبناء وتطوير أنطولوجيا عربية "مجال اللسانيات الحاسوبية أنموذجا"

A proposed methodology for building and developing an Arabic ontology The field of computational linguistics as a model.

أمينة أدغيري*
طالبة باحثة بسلك الماجستير، تخصص لسانيات ومعجمية عربية
قسم لغة عربية
معهد الدوحة للدراسات العليا
amounaoudghiri1@gmail.com

تاريخ الوصول.....تاريخ القبول ..... تاريخ النشر.....

ملخص

عرف القرن العشرون بزوغ فرع جديد من اللسانيات التطبيقية، ألا وهو اللسانيات الحاسوبية؛إذ شكلت هذه الأخيرة منعطفا قويا في كيفية دراسة اللغة، باعتبارها أحد العلوم البينية التي تحاول دراسة اللغات الطبيعية في علاقتها مع أساليب الذكاء الاصطناعي، وبتعبير آخر فهي مرحلة توصيف اللغة للحاسوب، ومرحلة التوصيف هاته هي عبارة عن وضع القواعد والقوانين اللغوية والحاسوبية لنمذجة اللغة انطلاقا من مجموعة من الخوارزميات التي من الممكن أن يتعامل معها الحاسوب، ومن ثم محاولة توصيف هذه النمذجة للتّحقق من صحة تلك الظواهر اللغوية، فتأتي بعد ذلك المعالجة الآلية للغة الطبيعية والتي تستمد إطارها النظري من اللسانيات الحاسوبية وتقنيات الذكاء الاصطناعي لتقوم بدورها بتصميم و إنجاز أنظمة آلية لمعالجة المعطيات اللغوية تحليلا وتركيبا، فتكون بذلك أساليب الذكاء الاصطناعي وهندسة المعرفة حلا لمعضلة اللغة، إلا أن تطبيقات هذا الذكاء الاصطناعي وعملية الحوسبة اللغوية بحاجة ملحة إلى مجموعة من الموارد اللغوية، تأتي على رأسها **الأنطولوجيا**.

الكلمات المفاتيح

أنطولوجيا، مدونة، اللسانيات الحاسوبية، أداة بروتيجي،أداة نوج، شبكة كلمات، الشبكة الدلالية، التصنيف الفئوي، العلاقات الدلالية، الأصناف، الفئات، الأفراد، السمات، معجم لغوي، رقمنة المعاجم، مورد لغوي، معالجة اللغات الطبيعية.

Abstract :

The twentieth century marked the emergence of a new branch of applied linguistics, which is computational linguistics. The latter constituted a strong turning point in how language is studied, as it is one of the interdisciplinary sciences that attempts to study natural languages in relation to artificial intelligence methods.In other words, it is the stage of characterizing the language of the computer.This stage of characterization is about setting linguistic, and computer rules, and laws to model the language from a set of algorithms that

* أمينة أدغيري

the computer can deal with.Then try to characterize this modeling to verify the validity of those linguistic phenomena.Then comes the automatic processing of natural language, which derives its theoretical framework from computational linguistics, and artificial intelligence techniques, to design and implement automated systems for processing linguistic data analysis and synthesis.So that artificial intelligence and knowledge engineering methods are a solution to the language dilemma.But the applications of this artificial intelligence and the process of linguistic computing an urgent need for a set of linguistic resources, on top of which is **ontology**.

Keywords

Ontology, Corpus, computational linguistics, Protege, Nooj, WordNet, Semantic Net, Subsumption, Semantic relationships, Entities, Classes, Individuals, Attributes, Lexicon, Digitizing dictionaries, Linguistic resource, NLP.

أهداف الدراسة

يقوم البناء الأنطولوجي بالأساس على الشبكات الدلالية التي تعمل كأساس لبناء خرائط ذهنية ومفاهيمية، حيث تقوم على تكشيف الكيانات في رسم بياني لصورة شبكة مؤلفة من عقد لهذه المفاهيم ترتبط فيما بينها بواسطة أقواس تمثل العلاقات، لتكون بذلك مستودعا معرفيا ومعجميا ودلاليا ضخما، مادته الخام هي شبكة الكلمات بكل ما تحويه من بيانات معجمية ودلالية لمصطلحات لغوية دالة، ترتبط فيما بينها بشبكة دلالية ذات معنى. وعليه نتوخى من هذه الدراسة ما يلي:

- إنشاء أنطولوجيا حاسوبية متخصصة قائمة على مدونة لغوية؛
- تصميم منهجية محكمة لبناء أنطواوجيا عربية؛
- نقل هذه المنهجية من حيزها النظري إلى الشق التطبيقي؛
- العمل على إنشاء مدونة نصية خاصة بمصطلحات اللسانيات الحاسوبية، والتي من شأنها أن تكون المادة الخام التي ستقف عليها أنطولوجيتنا إبان عملية تحريرها؛
- استيقاء مادة المدونة من كتاب "مقدمة في حوسبة اللغة العربية"[1]؛
- الاعتماد على أحد الأدوات الحاسوبية الخاصة بتحليل المدونة؛
- الاعتماد على أحد الأدوات الحاسوبية الخاصة بتحرير الأنطولوجيا؛
- محاولة تقييم وإبراز مكامن القوة والضعف لهذه الأداة.

منهجية الدراسة

تحاول هذه الدراسة البحث عن خطة محكمة يستطاع من خلالها بناء أنطولوجيا عربية متطورة وعليه يمكن صياغة إشكالية الدراسة على النحو الآتي: **إلى أي حد يمكن تصميم منهجية ناجعة لبناء وتطوير أنطولوجيا عربية؟**، وعليه فرضت هذه الإشكالية الاعتماد على الأسلوب المسحي من خلال الاستعانة بمدونة لغوية مضمنة لمصطلحات الدراسة، بغية إخضاعها للوصف والتحليل،

وهذا من شأنه استحضار المنهج الكمي في عملية جرد هذه المصطلحات من المدونة وإحصائها، ليأتي في الأخير المنهج التطبيقي محاولين من خلاله تطبيق منهجية بناء الأنطولوجيا على حقل اللسانيات الحاسوبية عبر أدوات حاسوبية خاصة.

مقدمة

يعد مصطلح **أنطولوجيا** مصطلحا يوناني الأصل، ظهر في القرن الثامن عشر سنة 1613 مع أرسطو، حيث استعمل أول مرة في مجال الفلسفة التحليلية فكان من بين أحد مباحثها الرئيسية، إلى أن أصبح عنوانا لأول مرة لكتاب كريستيان فون وولفChristian Wolff. إن مفهوم الأنطولوجيا عند أرسطو يتمظهر في علم الوجود بما هو موجود[2] ؛ إذ يتضمن دراسة طبيعة الموجودات الكونية أو ما نفترض أنه موجود، وكذا البحث عن كينونة الكيانات والأشياء المختلفة، وكذا أنواعها، وبنيتها، وخصائصها، وواقعها وعملياتها، ثم العلاقات القائمة بينها، بهدف بلوغ الحقيقة واستكشاف أصول العالم وموجوداته بما فيها من فئاته، وأنواعه وعلاقاته.

مع مرور الزمن وتطور المصطلحات والتكنولوجيا، انتقل هذا المصطلح في مرحلة التسعينيات مع حقيبته التعريفية ليحيط بمجال اللغة، ويصبح من بين أحد العلوم الأساسية في اللسانيات الحاسوبية وخاصة ميدان التحليل الدلالي، كما نجده أيضا يغزو واقع العالم الافتراضي؛إذ أصبح من بين أهم التقنيات المستعملة لنقل المعرفة وتطور الشبكة العنكبوتية و تقنيات الذكاء الاصطناعي،مما جعله من بين أولى اهتمامات المعلوماتيين واللسانيين على حد السواء، خاصة في مجال علم الدلالة؛ كون الأنطولوجيا تقوم بالأساس على المعنى و دلالة المفاهيم، انطلاقا من علاقات دلالية محددة ترتبط فيما بينها بشكل هرمي للوصول إلى عمق المعرفة.

رغم اهتمام كل من اللسانيات وعلوم الحاسوب بالأنطولوجيا، إلا أن كلا منهما ينظر إليها من زاوية محددة؛ فاللسانين يرونها مكنزا لغويا ضخما، يعبّر عن خريطة معرفية لشبكة دلالية معقدة ومنتظمة هرميا، لبلوغ معرفة دقيقة للمجال المراد دراسته. أما في مجال الهندسة المعلوماتية، فيأخذون هذا المصطلح كأحد التقنيات المعرفية لتطوير المعرفة الحاسوبية وحل المشكلات الدلالية المتعلقة بأدوات البحث والبرمجيات الآلية، سواء في معالجة اللغات الطبيعية أو استرجاع المعلومات أو تطوير كفاءة التطبيقات الحاسوبية.

يقودنا الحديث عن الأنطولوجيا إلى استحضار المدونة، كونها عماد أي بناء أنطولوجي، وتأثيرها واضح على مستوى تقويم نتائج البحث اللغوي وتأكيد أو ترجيح الكثير من الفرضيات حول اللغة، ونفي فرضيات أخرى[3]، وعليه سيكون توظيفنا لها دعما لما نروم إليه من تحرير لأنطولوجيا، إذ من شأنها أن تزودنا بما يلزمنا من مصطلحات الدراسة مع الدراسة الإحصائية لها، وتحقيق هذه العملية

يتطلب الاعتماد على أداة تحليل مناسبة، وقد حددناها في أداة Nooj الكفيلة بتحليل مدونات ومعالجتها معالجة لسانية وإحصائية.

إن الحاجة إلى بناء أنطولوجيا بناء نصف آليا، تستدعي الاستعانة بأدوات برمجية تساعدنا على تحقيق ذلك، ولعل برنامج Protege من أكثر الأدوات شيوعا واستخداما في هذا المجال، وعليه سنجعل منه واجهة تمثيلنا لأنطولوجينا والتي حددناها في مجال "اللسانيات الحاسوبية"، لنشرح خطوة بخطوة منهجية البناء الأنطولوجي الشبه الآلي لهذا الحقل المعرفي، لنخرج في الأخير بباقة من التقييمات التي تخص فعالية الأنطولوجيا في خدمة اللغة العربية ومدى تحقيق أداة التحرير الآنفة ذكرها لهذه الخدمة.

I. الأنطولوجيا الحاسوبية العربية (دراسة تنظيرية)

1. ماهية الأنطولوجيا

يسوقنا الحديث عن الأنطولوجيا إلى تصور خريطة معرفية من مفاهيم حقل معرفي معين؛ حيث ترتبط هذه المفاهيم فيما بينها بعلاقات دلالية مختلفة.فتعرّف[4] على أنها نمذجة تسعى إلى التمثيل المفاهيمي للمجالات والقطاعات المعرفية، وما تعكسه من معرفة فرعية وما بداخلها من الكيانات وعلاقات تربط بين هذه الكيانات[5]، لتكون بذلك المفاهيم والعلاقات الدلالية أساس أي بناء أنطولوجي معرفي.

أضاف الفيلسوف ويلسون Wilson في هذا الصدد، أن الأنطولوجيا تمثل بشكل خاص مفردات غنية لغويا تعمل على وصف مجال يمكن أن يكون مفهوما من قبل أجهزة الحاسوب[6] . فبمحاذاة كل من المفاهيم والعلاقات، أتى ويلسون بخاصية الغنى، وهي خاصية تحيل بالأساس على المواصفات الدلالية التي تحملها المفاهيم؛ إذ تمثل هذه المواصفات خصائص مميزة تميز مفهوما عن آخر عبر منحه سمات تجعله معترفا به من طرف الكيان المنسوب إليه، فيكون قابلا للمعالجة آليا، ومن تم يتسنى فهمه من قبل البشر والبرمجيات الحاسوبية على حد السواء.

كثيرا ما نجد خلطا بين الأنطولوجيا وشبكة الكلمات، غير أن هناك فرقا بينهما، وهذه حقيقة لا مراء فيها، إذ تمثل الأولى مجال علوم الإدراك الذي يتعامل معها كخريطة معرفية معقدة تحوي مفاهيم مستقاة من المدونة، وتستمد هذه المفاهيم قوتها من السياق الذي ترد فيه. بيد أن شبكة الكلمات تعبّر عن علوم اللغة، فتستمد مادتها من المعاجم اللغوية، وتركز اهتمامها على ثنائية المبنى والمعنى وما يربط بينهما من علاقات. لنخلص من هذا الطرح إلى أن كل أنطولوجيا هي شبكة كلمات، ولكن ليست كل شبكة كلمات أنطولوجيا.

تستقي الأنطولوجيا بنيتها من الواقع اللغوي، فهي تمثيل لمجال معين يعكس واقع لسان بعينه، وهذا ما ذهب إلى إثباته السعيد بقوله على أنها "مجموعة من الكتل المعرفية التي تتلاحم عناصرها لتكون تمثيلا لمجموعات منظمة من المفاهيم المشتركة في حقول دلالية محددة، مترابطة بعلاقات دلالية، وتستقي مادتها من مختلف فروع المعرفة،شريطة أن تعكس واقع اللغة المعينة"[7]. تعبّر الأنطولوجيا من هذا المنطلق عن شبكة دلالية تربط كيانات العالم بعلاقات معينة خاضعة لمعايير دلالية محددة، إذ ميّز اللغوي أتول Authoul بين نوعين من هذه العلاقات، معبرا عنهما بقوله إن الأنطولوجيا هي البنية التي تنظم فيها المفاهيم والعلاقات بشكل هرمي عبر علاقة التصنيف الفئوي Subsumption والعلاقات الدلالية ضمن مجال معين[8]، أما التضمين الفئوي فهو عنده يعكس علاقة الجزء من الكل؛ كأن نقول "صنف من SubClass of" أو "واحد من Is a". في حين يعبر النوع الثاني عن علاقات دلالية ينشئها منشئ الأنطولوجيا انطلاقا من تصوره ورؤيته المنطقية مع مدى معرفته بالمجال.

2. جوهر الأنطولوجيا

1.2 أنواع الأنطولوجيا

المعروف أن استعمال الأنطولوجيا يكون في إطار جمع مادة معرفية وتمثيلها بشكل منهجي منظم، بحيث يمكن الاستفادة منه حاسوبيا. وعليه اختلفت تصنيفات الأنطولوجيا من باحث لآخر وذلك حسب شموليتها وغاياتها، فذهب روسيRoussey وآخرون[9] إلى التمييز بين أربعة أصناف، تّم التعبير عنها كالآتي:

الأنطولوجيا التأسيسية العليا: وهي أنطولوجيا شاملة وعامة تستخدم في مجالات عدة؛ إذ تصبو إلى التعريف بالمفاهيم الأساسية والمجردة لكل مجال، ومن مثل هذه الأنطولوجيات نجد أنطولوجيا سومو SUMO وأنطولوجيا دولتشيDOLCE.

الأنطولوجيا العامة: وهي الأنطولوجيا غير المقيدة بمجال محدد، وإنما تتوسع لتصف معرفة عامة في مجالات عدة، وقد تحتوي على مئات الآلاف من المصطلحات والمفاهيم. ومن الأنطولوجيات التي يُزعم انتماءها في إطار هذا الصنف؛ نجد الوورد نيت WordNet باعتباره مكنزا أو قاموسا إلكترونيا ضخما يحتوي على معلومات تقدّر ب 155000 كلمة من الأسماء والأفعال والصفات. وهناك أيضا مشروع الأنطولوجيا العربية[10] لمصطفى جرار الذي استفاد من الوورد نيت بشكل غير مباشرة وذلك باتباع طريقة تمثيل البيانات نفسها للربط بين الشبكتين؛ شبكة الكلمات والشبكة الدلالية، انطلاقا من بناء المستويات العليا للأنطولوجيا وتوفير المادة اللغوية المحوسبة؛عبر رقمنة المعاجم العربية ودمج بعضها بالبعض. بيد أن هذا الزعم خاطئ إلى حد بعيد، وسيأتي تعليل ذلك في القادم من الأسطر.

الأنطولوجيا المرجعية الأساسية: وتسعى إلى تقنين مفاهيم مجال تم وصفه؛ إذ يُبنى حجرها الأساس انطلاقا من دمج واتحاد عدد من الأنطولوجيات المتعلقة بنفس المجال.

أنطولوجيا المجال: وهي أنطولوجيا مخصصة لوصف مفاهيم مجال معين من وجهة نظر معينة، وغالبا ما يكون استعمالها مقتصرا على المهتمين بهذا المجال، كما يركز هذا النوع من الأنطولوجيات على وصف العلاقات بين المفاهيم بشكل خاص. ومن أمثلتها نجد أنطولوجيا القرآن الكريم Quran Ontology المعنية بمفاهيم القرآن الكريم والأحاديث النبوية الشريفة.

أنطولوجيا التطبيق: وهي أنطولوجيا مخصصة لتطبيق معين؛ إذ لا تسعى إلى مشاركة المفاهيم بقدر ما تروم إلى "توصيف وتوثيق المفاهيم المختصة بتطبيق معين من وجهة نظر معينة"[11].

رغم اصطلاح تسمية أنطولوجيا على هذه الأنواع إلا أن هذا لا يعني أنها أنطولوجيات حقيقة، إذ إلى حدود كتابتي لهذه الكلمات لا يمكن الجزم بوجود أنطولوجيا شاملة لجميع المفاهيم المعرفية الموجودة على أرض الواقع، أو كما يطلق عليها بالأنطولوجيا العامة، وعليه فزعمهم بأن الوورد نيت ومشروع الأنطولوجيا العربية يدخل ضمن إطار البناء الأنطولوجي، هو زعم خاطئ إلى حد بعيد؛ فما هذه المشاريع إلا انعكاس لعملية رقمنة المعاجم، في حين أن البناء الأنطولوجي يستوعب المفاهيم في علاقتها بالسياق اللغوي. وبالتالي فإن أي بناء أنطولوجي يحتاج بالأساس إلى كمّ مهول من المدونات اللغوية والاستناد إلى قواعد بيانات معرفية ضخمة، والواضح أنه مراد صعب المنال لكنه ليس بالمستحيل، إذ يمكن من خلال بناء ودمج أنطولوجيات متخصصة في مجالات مختلفة أن يساعد في الوصول إلى بلورة أنطولوجيا عامة في القريب البعيد، ومن ثم بلوغ الغاية الأسمى ألا وهي التخفيف من فجوة الغموض المفاهيمي والدلالي الذي يلحق المصطلحات في بيئة التكنولوجيا والمعلومات.

1.2 مكونات الأنطولوجيا

يقوم البناء الأنطولوجي على مجموعة من العناصر والمكونات التي تشكل البنية الهندسية للأنطولوجيا،وهي مكونات مستقاة ومستوحاة أساسا من المجال الفلسفي، إذ نجد كلا من علم الفلسفة وعلوم الحاسوب يسعيان إلى تمثيل المعرفة في صورة كيانات، وأفراد، فخصائص ثم علاقات، ويستحيل الحديث عن أنطولوجيا حاسوبية بغياب أحد هذه العناصر؛ وقد فصّلها النشرتي في الآتي[12]:

- **الكيانات Entities**: تُعرف أيضا بالأصناف، وتمثّل المستوى الأول داخل الأنطولوجيا، كما تعكس مختلف الكيانات المادية والمجردة.

- **الفئات Classes:** تشمل الفئة على مجموعة من المفردات individuals تجمع بينها صفات مشتركة جعلتهم ينتمون إلى هذه الفئة التي تم احتواؤهم فيها، ولا يقتصر اشتمال الفئات على المفردات فحسب، بل تمتد لتشمل أيضا فئات فرعيةsubclassesتشكل جزءا من فئة رئيسية، ويكون التمثيل في شكل تسلسل هرمي منطقي.
- **السمات Attributes:**توصف كلا من الفئات والمفردات في الأنطولوجيا وفق الخصائص التي تميزها عن غيرها والمحددة لذاتها، كما لا تقتصر على توصيف الفئات والأفراد فقط، بل تتعدى ذلك إلى توصيف العلاقات التي تربط هذه الكيانات ككل، وذلك بتحديدها لطبيعة ونوع العلاقة.
- **العلاقات Relationship:** وتعد الجوهر الأساس الذي يؤسس لإنشاء أنطولوجيا،باعتبارها أحد أهم السمات التي تميزها؛ إذ تتمتع بالتوصيف الدلالي، الشيء الذي يخول للحاسب الآلي القدرة على تحقيق التكامل المعرفي بين مختلف الكيانات.

3. منهجية بناء أنطولوجيا عربية

إن عملية بناء أنطولوجيا حاسوبية ليست بالأمر الهين؛ إذ تحتاج وقتا وجهدا جسيما، بالإضافة إلى الخبرة في المجال المراد دراسته؛ ناهيك على أن مهمة جمع مادة الأنطولوجيا ودراستها تعطى إلى العارف والملمّ بالمجال، بهدف أن يكون البناء سليما، ومن ثم الوصول إلى نتائج دقيقة ومرضية. وقد طُرحت عدة طرق لبناء الأنطولوجيا، من قبيل SENSUS وKACTUS وOTK و Common KADS وTove وMicrokosmos وONIONS وHYSSYS. ويرى المعلوماتي كيم[13] أن جميع هذه الطرق لم تحقق نجاحا؛ إذ لم تساعد في حل الإشكالات التي تظهر إبان عملية الإنشاء الفعلي؛ وهذا راجع لكون بعضها خُصِّص لأنطولوجيات محددة، وأخرى كانت بدائية عشوائية، في حين لاتزال طرق أخرى في طور التطوير، والمتوقع أنها لن تخرج بنتائج مجدية.

أبانت بعض الدراسات والأبحاث في السنوات الأخيرة عن منهجيات معينة تعين في البناء الأنطولوجي، وتكاد أن تتفق جميعها على خطوات واحدة هدفها الأول والأخير الوصول إلى أنطولوجيا محكمة البناء، وقد خص الأكلبي[14] إحدى دراساته بالحديث عنها، فحددها في عشر خطوات. كما ذهب كلوب[15] بدوره في رسالته إلى التفصيل فيها في سبع خطوات، في حين آثرنا اختزالها في ثلاث خطوات، وهي كالآتي:

1- مرحلة الجمع والتحديد

- تحديد المجال الأنطولوجي ومصدر الأنطولوجيا والغرض منها: إذ لا بد من تعيين مجال التخصص الذي ستغطيه الأنطولوجيا وكذا مصادر المعلومات، مع تحديد الجدوى والأهداف المتوخاة منها؛
- حصر وتجميع المصطلحات الخاصة بالمجال المراد توصيفه؛ والتي تعرف بالكيانات، والفئات، والأفراد، مع تحديد معانيها وسماتها؛

- توضيح العلاقات الرابطة بين المصطلحات، دون إغفال تحديد سمات هذه العلاقات.

2- مرحلة الشروع في بناء الأنطولوجيا

- التأكد من عدم وجود أنطولوجيا في المجال المراد دراسته؛
- تعيين أداة تحرير برمجية مناسبة في عملية بناء الأنطولوجيا، وثم تصميم خريطة من مصطلحات تمثل شبكة دلالية؛
- الحصول على صيغ توكيد النصوص عبر لغة xml[16]، وكذا إطار توصيف الموارد Rdf[17] للمصطلحات الموظفة، وبالتالي إنتاج أنطولوجيا.

3- مرحلة التكامل والاندماج

- تزويد الأنطولوجيا بمعارف متنوعة لجعلها غنية معرفيا وواسعة استعمالا، ويكمن هذا التزويد في إضافة العلاقات الدلالية إلى جانب العلاقات الهرمية، وإضافة علاقات البيانات لكل مفهوم، ثم إضافة أمثلة لهذه المفاهيم والمصطلحات لإثراء المحتوى المفاهيمي، كما يمكن إضافة بعض المعارف الاستنتاجية إن وجدت؛
- تقويم الأنطولوجيا بفحصها ومراجعتها للتأكد من عدم وجود أخطاء فيها، وكذا تقويم مطابقتها للغرض الذي أنشئت من أجله؛
- الحرص على مداومة الصيانة والتطبيق؛ وهي عملية تنظيمية عبارة عن تحديث وتحسين مستمر للأنطولوجيا بتصحيح الأخطاء، مع الأخذ بعين الاعتبار آراء وتعليقات المستخدمين كنقط أساسية وقيمة من أجل تحديث التغييرات اللازمة، ثم تأتي مرحلة التطبيق وهي المرحلة التي تُعرض فيها الأنطولوجيا النظيفة على البيئة الرقمية.

II. المدونة، ودور أدوات تحليلها في جمع مادة الأنطولوجيا

1. المدونة Corpus

تعد المدونة مجال بحث ما فتئ يأخذ مكانه في الساحة اللغوية الحاسوبية، فهي عبارة عن دخيرة أو مستودع ضخم من النصوص اللغوية الطبيعية –المكتوبة أو المنطوقة –التي يمكن التعامل معها بشكل آلي، وذلك من خلال حفظها على الآلة ومن ثم التحكم في بياناتها والتعديل عليها عبر مجموعة من قواعد البيانات المخزّنة للغة. ويحدد السعيد[18] طبيعة هذه النصوص بأنها ليست تقييديّة أو عشوائية، لكنّها كتل غير منتظمة تخضع لمجموعة من الأسس والمعايير، يحددها الهدف المنشود من المدونة اللغوية.

يطلق مصطلح لسانيات المدونات Corpus linguistics على دراسة اللغة في ضوء ما هو مخزن حاسوبيا من نصوص لغوية، وعليه فهي منهج بحثي لساني جديد، يندرج تحت مضلة اللسانيات

التطبيقية. ويعنى بدراسة جانبين أساسين لهما من الأهمية بمكان؛ الأول جانب نظري يبحث في الفرضيات والإشكالات اللغوية، والآخر تطبيقي يعنى بالنمذجة الحاسوبية للظواهر اللغوية[19] . الملاحظ أن استعمال المدونات شاع على نطاق واسع، ولعل سبب ذلك يعود بالأساس في كونها "تكشف كيف يستخدم الأشخاص الحقيقيون اللغة الحقيقية"[20] ، بمعنى أنها تعكس السياقات اللغوية كما هي في الواقع اللغوي، بهدف الإحاطة بجميع الدلالات التي يكتسبها اللفظ الواحد في سياقات مختلفة، ومن ثم جعل ذكاء الآلة أكثر مرونة وفطنة بالتنوع الدلالي للألفاظ التي تزيد عادة عما هو مدون ومعروف في المعاجم. ولعل أبرز الميادين البحثية التي توظف فيها المدونة فتستفيد منها؛ نذكر معالجة اللغات الطبيعية NLP، وصناعة المعجمات والموارد اللغوية ودراسة اللهجات والترجمة بين اللغات، ناهيك عن الاستفادة التي تحققها في الأغراض التعليمية التعلمية من قبيل تعليم اللغات، وتطبيقات الذكاء الاصطناعي والتعلم الآلي وغيرهم[21] .

يتطلب البحث اللغوي المعتمد على المدونة السير على منهج معين للوصول إلى نتائج مرضية، ولعل أولى خطوات هذا المنهج تكمن في وضع القضية الإشكال في ذهن الباحث، ومن ثم طرحه لمجموعة من الإشكالات التي يتوخى من المدونة أن تعينه في الإجابة عنها، لينتقل إلى الخوض في غمار تحليلها وتوضيحها انطلاقا من كتلة من النصوص اللغوية المرقمنة في المدونة. لتأتي بعد ذلك خطوة التدوين مفسّرا إياها عالم اللسانيات المدونات توجنينيTognini بأنها عملية تدوين ما يمكن استنباطه من قضايا لغوية صرفية، وتركيبية، ودلالية وسياقية داخل المدونة، والانطلاق منها لصياغة فرضيات البحث وإشكالاته[22]، هذا يعني أن تحديد الفرضية اللغوية من شأنها أن توجه الباحث إلى نوع المدونة اللغوية وكذلك نوع النصوص الهدف وعددها، بل وحتى الأدوات البرمجية الملائمة لهذا المسعى.

يتجلى هدفنا من استعمال المدونة اللغوية في هذه الورقة، محاولة جرد مصطلحات تخصص اللسانيات الحاسوبية لهدف أسمى ألا وهو توظيف هذه المصطلحات كلبنة أساسية في عملية هندسة الأنطولوجيا، إذ كان لزاما علينا اختيار مدونة تحوي مصطلحات هذا التخصص نظرا لما تشترطه الأنطولوجيا من اتخاذ المدونة مرجعا أوليا في توثيق المصطلحات، وعليه وقع اختيارنا على مدونة "مقدمة في حوسبة اللغة العربية" التي يتوخى منها أن تسعفنا في تحصيل هذا الغرض.

2. اعتماد أداةنوجNooj في تحليل المدونة

أضحى إعداد برمجيات تطبيقية من الوسائل المعينة في مجال معالجة اللغات الطبيعية (NLP)، وذلك لما توفره من خدمات مهمة للبحث اللغوي. وكانت للغة العربية نصيبا من هذه البرمجيات؛ إذ حظيت باهتمام العديد من الخبراء والباحثين في مجال الهندسة اللغوية، ولعل من أبرز الأدوات اللسانية المعلوماتية التي شاع استعمالها وحققت نتائج دقيقة نسبيا في مجال معالجة اللغات

الطبيعية وخاصة فيما يخص المدونات اللغوية، نجد Aconcorde، و Antconc وNoojs؛ أما عن الأداة الأولى فهي من الأدوات التي تدعم اللغة العربية بشكل قوي، وتركز اهتماماتها على الفهرسة الآلية للنصوص، غير أنها لا تستوعب مادة كبيرة بالمقارنة مع Antconc وNoojs اللتان تستوعبان مدونات من الحجم الكبير، إلى جانب دعمهما القوي أيضا للعربية، ويحسب لهما أيضا أنهما يقبلان بالمتصاحبات اللغوية عكس الأولى. وعليه سنركز اهتمامنا في هذه الورقة على أداة Nooj في معالجة المدونة اللغوية وتحليلها.

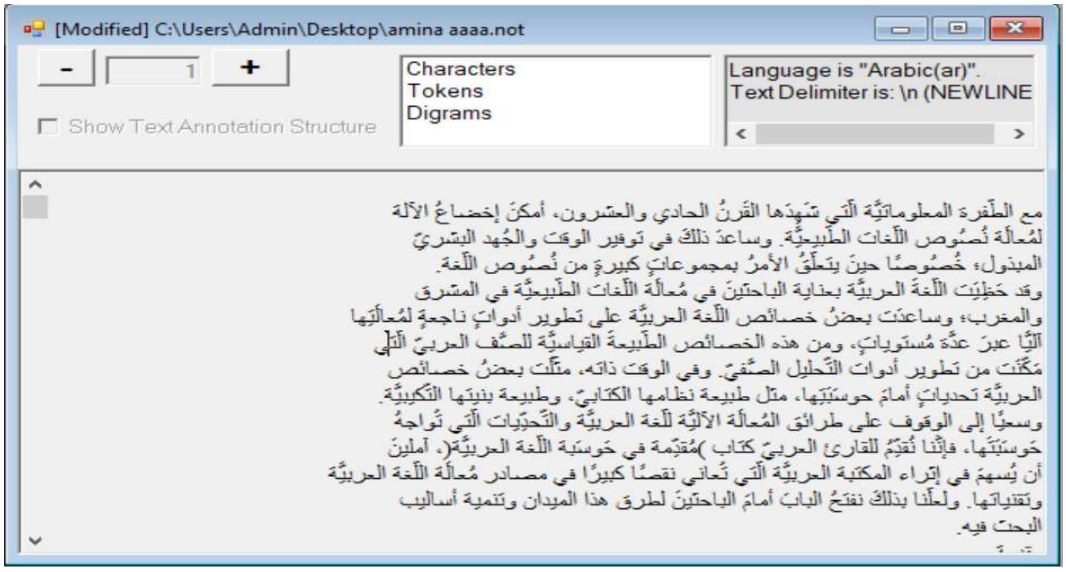
تعد أداة Nooj[23] منصة رقمية مجانية تتلاءم مع العديد من الأنظمة الحاسوبية، كما تستخدم من قِبل العديد من مختبرات البحث والبرمجة، وهذا راجع في كونها أكثر دقة من حيث النتائج في مجال التحليل اللغوي، علاوة على ذلك نجدها تدخل ضمن الأدوات التي تدعم اللغة العربية بشكل أفضل بالمقارنة مع سابقاتها. تمكّن مخبر المعالجة الآلية والتوثيق اللساني (LADL) بباريس ومختبر (ELLIADD) بجامعة FrancheComté من بلورة منصة نوج الإلكترونية اللغوية لتعمل كمحرك ومعالج لساني[24] ، إذ إنها مزودة بهيكل إحصائي يخول للمستعمل معرفة المعلومات الخاصة بالترددات الرقمية للمفردات داخل المدونة الموظفة، والاستفادة منها كذلك فيما يخص تحليل ظواهر ومقامات تداولية مختلفة.

تمثل أداة نوج من هذا المنطلق أداة واعدة للاستخدام في مجال تحليل المدونات، كما أنها تتميز بالبساطة؛ إذ تسمح لأي باحث غير متخصص في الهندسة المعلوماتية بالعمل عليها وتحليل ما يبتغي من المدونات اللغوية، نظرا لما تحتويه من تقنيات هندسية منطقية في المعالجة الآلية للغة؛ كلغة برمجة مفتوحة المصدر. تخول أداة نوج بالدرجة الأولى إمكانية معالجة المدونات المعجمية في مستوياتها الكرافيمية، والصوتية، والصرفية، والتركيبية والدلالية، مما يختصر قطع أشواط كبيرة في عملية التحليل، لتبقى الرقمنة المكثفة للغة المتطلب الوحيد لهذه الأداة.

3. توظيف أداة Nooj لتحليل مدونة "مقدمة في حوسبة اللغة العربية" آليا

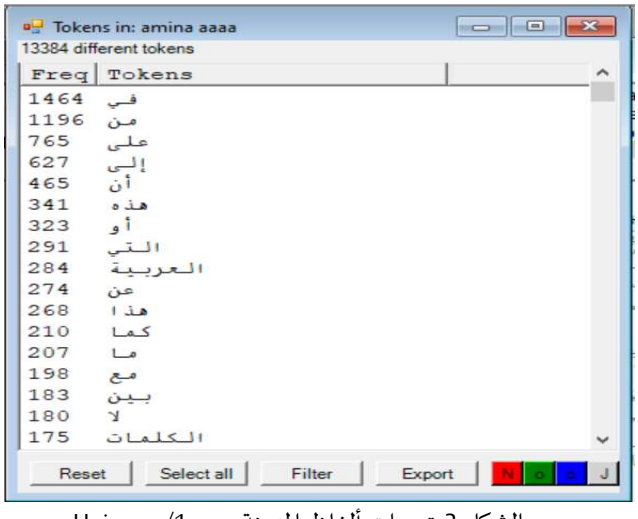
يحتاج بناء أنطولوجيا متخصصة الاستناد إلى مدونة خاصة بمصطلحات المجال، ونظرا لأننا أخذنا على عاتقنا مصطلحات اللسانيات الحاسوبية ميدانا للدراسة والتحليل، كان لزاما علينا اختيار مدونة تحوم حول هذا الموضوع، فما وجدنا غير كتاب "مقدمة في حوسبة اللغة العربية"(2019) ليعيننا على تحقيق هذه الغاية، وهو كتاب متوسط الحجم؛ إذ يضم بين دفتي 244 صفحة؛ طُرحت فيها مجموعة من فصول تناولت قضية حوسبة اللغة العربية من جوانب عدة، مما يُتوقّع أن تكون زاخرة بالمصطلحات التي نرمي مقاربتها في هذه الدراسة.

يعد توظيف أدوات حاسوبية في تحليل المدونة من العمليات التي توفر وقتا وجهدا على الباحث في عمله، وعليه اعتمدت دراستنا أداة Nooj في تحليل مدونة "مقدمة في حوسبة اللغة العربية"، محاولين بذلك حصر مصطلحات المدونة مع التركيز على ترددتها لانتقاء ما يفيدنا في مجال بحثنا الأنطولوجي.

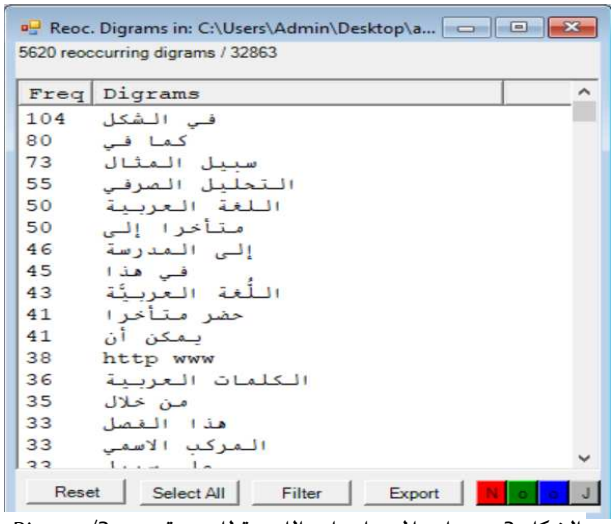


الشكل 1:التحليل اللغوي لمادة المدونة.

يمثل الشكل فوق مادة المدونة التي تم نقلها من الكتاب ووضعها على واجهة الأداة ليتم قراءتها وتفكيك عناصرها، ويرجع الهدف الأساسي من هذه العملية إلى معرفة إحصاء ترددات العناصر اللغوية التي نتوخى منها أن توجهنا إلى مادة الأنطولوجيا. ونمثل لهذه الترددات في الشكل الآتي:

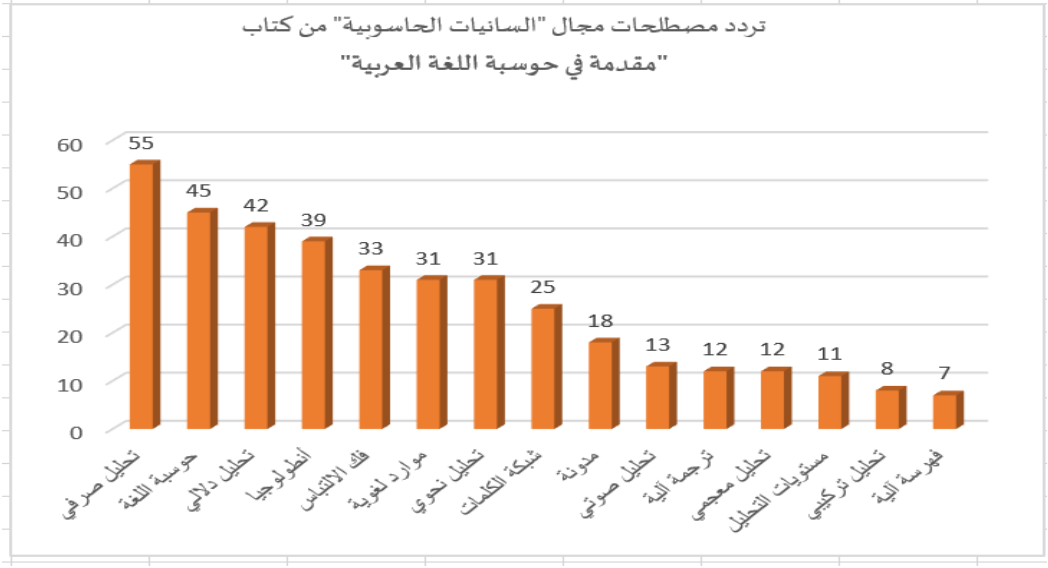


الشكل 2: ترددات ألفاظ المدونةUni gram/1 gram.



الشكل 3: ترددات المتصاحبات اللغوية للمدونةBi gram/2 gram.

تخول أداة Nooj معرفة ترددات الألفاظ اللغوية والمتصاحبات اللغوية، بل وتتعدى ذلك إلى أن تضعها بشكل تراتبي من شأنه أن يسهل عملية انتقاء مصطلحات الدراسة. وعليه وبعد إمعان النظر، ارتأينا أن نخصص خمسة عشر مصطلحا أكثر ترددا في المدونة المدروسة الخاصة بمجال اللسانيات الحاسوبية، ومن ثم إخضاعه للدراسة الأنطولوجية في القادم من النقاط. أما عن هذه المصطلحات يمكن عرضها في رسم بياني يوضح تسلسلها وترددتها كالآتي:



الشكل 4: المصطلحات الأكثر دورانا باعتبارترددتها.

III. تحرير أنطولوجيا عربية حاسوبية حول بعض مصطلحات اللسانيات الحاسوبية (دراسة تطبيقية)

1. لمحة عن أدوات بناء الأنطولوجيا

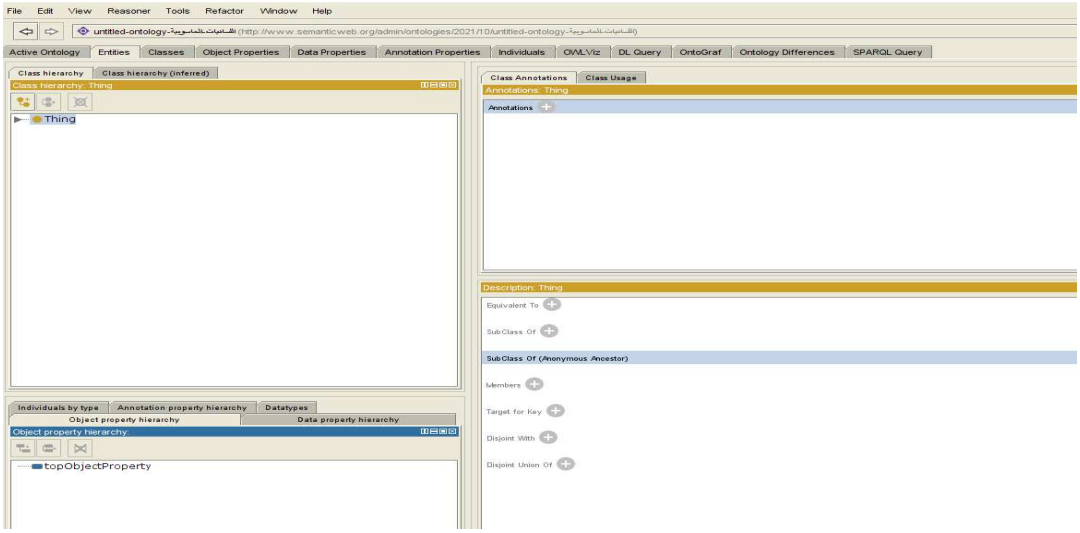
تصبو الدراسات الحديثة في مجال تنظيم المعرفة داخل البيئة الرقمية إلى تحقيق أكبر عدد ممكن من الشبكات المعرفية بلغة توصيفية تنحو نحو خدمة العالم الافتراضي ليواكب العصرنة التي نعيشها حاليا، والتي تحتاج إلى جهود جسيمة من الفكر اللغوي الإنساني والكيان الآلي المعلوماتي، حيث إن تفاعلهما من شأنه أن يقلص الغموض المفاهيمي ويفك الالتباس الدلالي الذي تعاني منه البيئة التقنية، والذي يجعلها عاجزة عن تحقيق الغاية جرَّاء ضياعها في محيط من الدلالات. ولعل اللغة العربية من أبرز اللغات الطبيعية معاناة من هذه الإشكالات، وهذا راجع إلى مجموعة من البواعث، على رأسها الاشتقاق وظاهرة التعدد الدلالي. وتخفيفا من حدة هذه الإشكالات تم ابتكار مجموعة من الأدوات والبرامج الحاسوبية المعينة على تقليص هذه الفجوة، حيث صمّمت خصيصا من أجل مساعدة الباحثين على تحرير خرائط معرفية أنطولوجية قادرة على خدمة متطلبات الشبكة العنكبوتية.

توزعت أدوات تحرير الأنطولوجيا إلى ما يُفعَّل من الموقع التابع للبرنامج فلا يسمح بتحميله بشكل مباشر، وأدوات أخرى تخول للمستعمل تحميلها على الحاسوب من الموقع الخاص بها، ونمثل لها بالأدوات الآتية:

- **أداة هوزوHOZO**: هي أداة يابانية المصدر،حيث تم تطويرها من خلال شراكة بين قسم نظم المعلومات في جامعة أوساكا اليابانية وشركة إنكيتEngeateالمحدودة، وتعد من الأدوات المفتوحة المصدر؛ إذ يمكن تحميلها[25] وتنصيبها على الحاسوب مع الاستعانة بلغة البرمجة الجافا،فتعمل على رسم الأنطولوجيا،وبناءها بشكل واضح ومنظم.
- **أداة SWOOP**:هي بدورها أداة مفتوحة المصدر، ومتاحة مجانا للتحميل[26]، تم تصميمها من طرف جامعة ميريلاند University of Maryland بالولايات المتحدة الأمريكية لإنشاء وبناء الأنطولوجيا.
- **أداة أوبو OBO**: صممت هذه الأداة في معهد بيركليBerklee للمشاريع المفتوحة المصدر في الولايات المتحدة الأمريكية عام 2007، وهي أداة كسابقاتها تعمل على بناء وتحرير الأنطولوجيا، ومن ميزاتها أنها ذو واجهة تعامل سهلة في الاشتغال.
- **أداة بروتيجيProtege**: أشهر الأدوات وأكثرها استعمالا في مجال تحرير الأنطولوجيا، صُمِّمت من طرف كلية الطب بجامعة ستانفورد بأمريكا، مجهزة بلغة الجافا ومفتوح المصدر[27] شأنها شأن البقية،

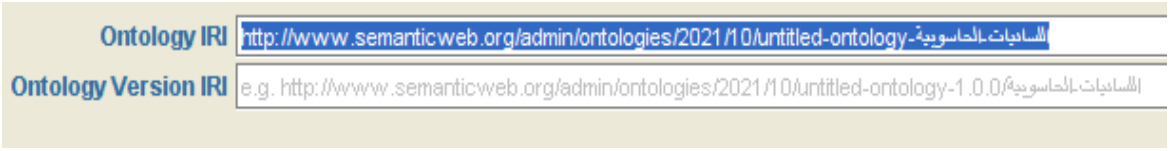
وعليه سنخص دراستنا هاته بتحرير أنطولوجيا عبر برنامج بروتيجي، لما يوفره من ميزات أبرزها أنه يسمح للمستخدم إنشاء شجرة المفاهيم آليامع تمكينه من التحديث في البيانات والإضافة عليها بكل سهولة. ناهيك على أنه مزود بكل الوسائل المتاحة لإعداد أنطولوجيا؛ من بناء الفئات والخصائص بنوعي، إلى استعلام نصوص الأنطولوجيا في الفضاء الشبكي[28].

2. خطوات بناء أنطولوجيا لبعض مصطلحات اللسانيات الحاسوبية آليا



الشكل 5: واجهة أداة Protege

إن لكل أنطولوجيا اسما تعرف به والذي غالبا ما يكون اسم المجال المشتغل عليه، وبما أن مجالنا هو اللسانيات الحاسوبية اخترناه اسما لأنطولوجيتنا، أما عملية التسمية داخل برنامج بروتيجي فتتم من خلال خانة "Ontology IRI" التي تحمل رابطا على شكل رقم تسلسي للأنطولوجيا.



الشكل 6: تسمية الأنطولوجيا

الشكل 7 :إضافة المفاهيم في مستعرض الأصناف Entities

بعد تحديدنا لاسم الأنطولوجيا، نشرع في بناء الأصناف والعلاقات والسمات الخاصة بمجال دراستنا وهي كالآتي:

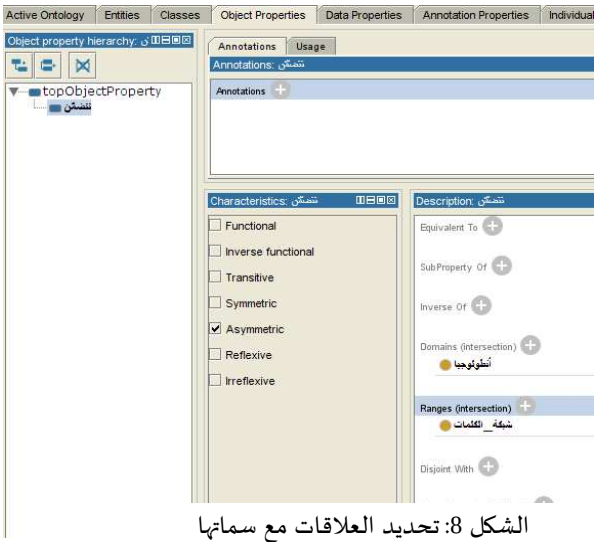
أولا: الأصناف

تتجلى أولى خطوات البناء في تنظيم المفاهيم بشكل هرمي مراعٍ الانتقال من العموم إلى الخصوص، وتتم هذه العملية عبر تبويب "Entities" الذي يخول إدراج الأصناف بشكل تلقائي، بدءا من المفاهيم الرئيسة انتقالا إلى تحديد المفاهيم الفرعية لكل صنف.

قمنا في الشكل جانبا بإضافة كيان أسميناه "اللسانيات الحاسوبية"، ثم أدرجنا تحته صنفا رئيسيا تحت اسم "حوسبة اللغة"، الذي فرعناه بدوره إلى مجموعة من الأصناف الفرعية؛ وهي مجموعة من المفاهيم الموازية تجمعها علاقة الجزء بالكل مع الصنف الرئيسي.

ثانيا: العلاقات

ننتقل بعد عرض المفاهيم إلى استعراض العلاقات الرابطة بينها، فبالرغم من أن البرنامج يقدم الربط العلائقي من تلقاء نفسه، إلا أنه لا يوفر سوى نوعين من العلاقات ألا وهي؛ علاقة التكافؤ Equivalent to وعلاقة الجزء من الكل SubClass of. وعليه يخول تبويب "Object Properties" للمُنشئ إضافة علاقات اختيارية.



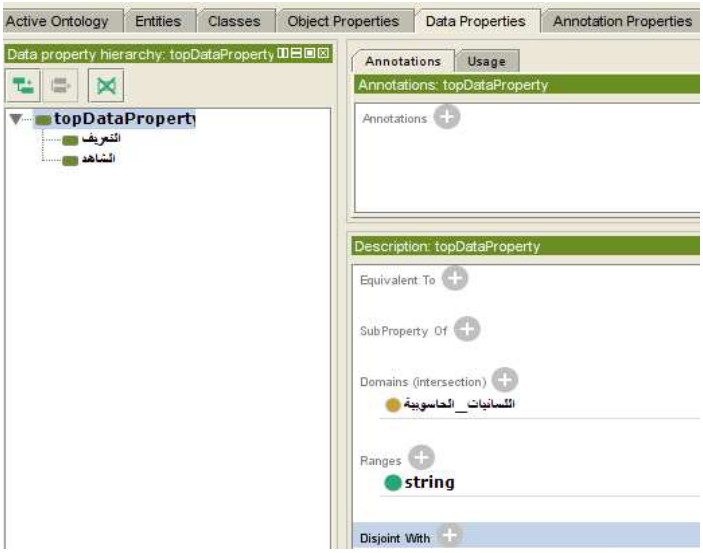
الشكل 8: تحديد العلاقات مع سماتها

علاقة التضمّن مثلا نحدد لها سمة اللاتماثلAsymmetric من مستعرض سمات العلاقات الذي يعرضه البرنامج، وهي سمة تجمع العناصر غير المتماثلة؛ بمعنى أن العلاقة بين المفاهيم هي ثابتة وفي اتجاه واحد عكس ما نجده مثلا في علاقة الترادف والتضاد.

لدينا في الشكل جانبا على سبيل المثال؛ المجال "أنطولوجيا" يتضمن المدى "شبكة الكلمات"، وعلاقة التضمن هاته ذات خاصية اللاتماثل، أي أن المفاهيم لا يسمح لها بأن تتبادل موضعها وبالتالي العلاقة مستقرة ذو اتجاه واحد.

ثالثا: السمات

تلي مرحلة إدراج العلاقات مرحلة تخصيص السمات، وذلك عبر التبويب الخاص بها Data Properties، وبما أننا اعتمدنا متصاحبات لغوية في مادة الأنطولوجيا سنخصص للأصناف سمتين فقط، وأخص بالذكر سمة "التعريف" و"الشاهد".

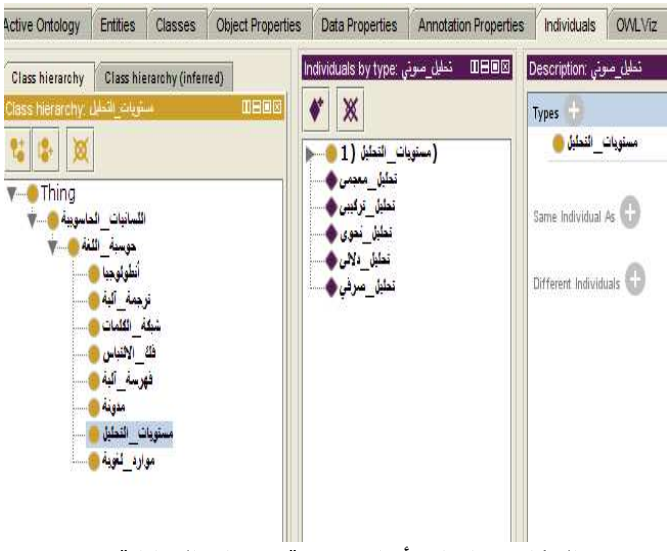


الشكل 9: إضافة السمات مع تحديد نوعها

تلي عملية تحديد نوع السمات تعيين خاصية هذه الأخيرة، وذلك عبر تحديد ما إذا كانت هذه الأصناف تعبر عن رقم أو حرف أو نص، وبما أننا حددنا السمات في "التعريف" و"الشاهد" فكلاهما يأخذان سمة النص المعبر عنها بخاصية String، كما هو موضح في الشكل جانبا.

رابعا: الأفراد

تأتي عملية إدراج الأفراد بعد عملية تعيين السمات؛ وأقصد بها إدراج المفاهيم الفرعية للأصناف الرئيسية، وتتم هذه العملية عبر مستعرض الأفراد Individuals by classes كما هو في الشكل جانبا.



الشكل 10: إضافة أفراد صنف "مستويات التحليل"

نختار من مستعرض الأصناف الصنف الأساس الذي نريد إنشاء فرد منه، ونمثل لذلك بصنف "مستويات التحليل" فندرج تحته كلا من "تحليل صوتي"، و"تحليل تركيبي"، و"تحليل نحوي"، و"تحليل صرفي"، "تحليل دلالي" و"تحليل معجمي"، باعتبارهم أفرادا من صنف "مستويات التحليل".

- العلاقة بين الأفراد

بعد تحديد الأفراد تأتي عملية ربطها بعلاقات عبر تبويب Object property assertions، ونمثل لهذه الخطوة بالأفراد "تحليل معجمي" و"تحليل دلالي" اللذان يرتبطان بعلاقة "الاستلزام".

الشكل 11: تعيين علاقة "استلزام" بين الأفراد "تحليل معجمي" و"تحليل دلالي"

- سمات الأفراد

نضيف سمات الأفراد عبر خيار Data property assertions، الذي يخول إضافة السمات مع المعلومات الخاصة بها كما هو موضح في الشكل جانبا.

الشكل 12: إضافة سمات فرد "تحليل نحوي"

نأخذ على سبيل التمثيل، الفرد المسمى ب "تحليل نحوي" ونضيف له سمة التعريف ثم الشاهد الذي نستمده من المدونة الموظفة، فيظهر على الشكل الآتي.

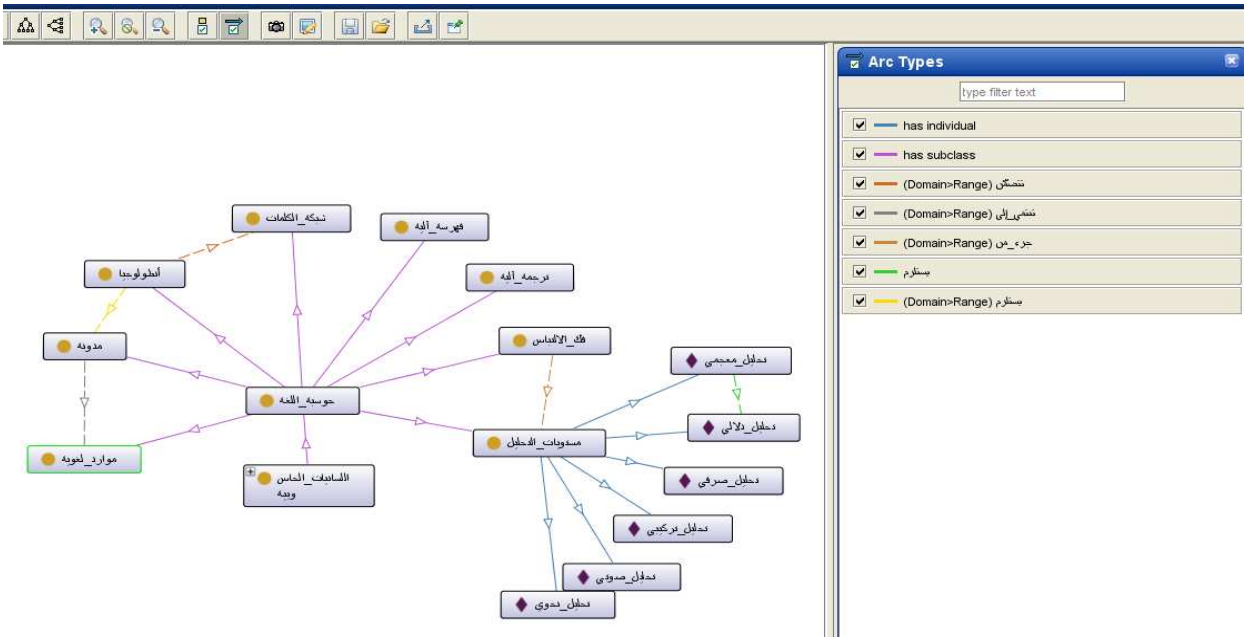
الشكل 13: تشغيل المفكر لاختبار صحة الأنطولوجيا

خامسا: التأكد من صحة الأنطولوجيا

تُظهر هذه المرحلة مدى صحة الأنطولوجيا وصحة بنائها، وذلك عبر تبويب Reasoner، ومن ثم النقر على الأيقونة الخاصة بتشغيل المفكر Start reasoner بعد التأكد من أن المفكر على الاختيار "HermiT".

سادسا: رسم الأنطولوجيا

بعد الانتهاء من كل المراحل فوق، تتبقى المرحلة الأخيرة المتمثلة في تمثيل هذه المفاهيم والعلاقات شجريا على شكل شبكة دلالية، وتتم هذه العملية عبر تبويب OntoGraf، الذي يخول بدوره اختيار الشكل الأنسب في التشجير، لنحصل في الأخير على أنطولوجيا لبعض من مصطلحات اللسانيات الحاسوبية، ممثلة على الشكل الآتي:



الشكل14: صورة لأنطولوجيا بعض مصطلحات مجال اللسانيات الحاسوبية

3. تقييم فعالية الأنطولوجيا وأدوات تحريرها

تعد مقاربة الأنطولوجيا آليا مقاربة واعدة من شأنها أن تسهل عمل الكشف عن دلالة المفاهيم من خلال الربط بين أصنافها بعلاقات تخول الوصول إلى معرفة دقيقة، كما تجدر الإشارة إلى أن عملية البناء هذه تسهم في رفع مهارة التفكير والإبداع الفني لدى الناشئ، إذ تجعل منه أكثر مرونة في اختيار الألفاظ وتحديد ما بينها من علاقات في ضوء ما يوظفه من موارد لغوية. ومن بين أهم ما يمكن أن نسجله للأنطولوجيا ويحسب لها أنها تعين على تنظيم وتصنيف المعرفة لكي تتكامل مع باقي التطبيقات وقواعد البيانات والبرمجيات الأخرى، إذ نجدها تساهم في إنشاء معاجم موضوعية، وكذلك إضافة الفهرسة الموضوعية عبر تطوير الفهرسات الآلية، بحيث يصبح اعتماد الترتيب الموضوعي للفهرس عوض الترتيب التقليدي، ناهيك عن الدور الفعال الذي يمكن أن تقوم به في حل معضلة الترجمة الآلية عبر تزويدها بأنماط لغوية عدة وسياقات دلالية دقيقة ومتنوعة يجعل منها أكثر قدرة على فك اللبس الدلالي.

وجلي بنا الحسم بأن العلاقة الوثيقة التي تربط الأنطولوجيا بعلم الدلالة ونظرياته حقيقة لا مراء فيها، إذ من شأنها شرح بنية اللغة بقدرتها على تفتيت اللفظ إلى ذراته الدلالية والوصول إلى نواته ثم إعادة تجميعه مرة أخرى وإخضاعه لهندسة المعرفة[29] لتكون بذلك وسيلة فعالة من وسائل التحليل المتقدم للمحتوى، والتي من شأنهاالنهوض باللغة على المستوى التقني وذلك بتخزين المعلومات واسترجاعها وتبادلها.

بيد كل هذه الإنجازات التي تقوم بها الأنطولوجيا على المستوى التقني، إلا أنها لا تخدم اللغة العربية بقدر ما تخدم اللغات الأخرى، وهذا راجع بالأساس إلى أدوات التحرير البرمجية للأنطولوجيا، فبالرغم من كونها تساعد على إدراج الأصناف والعلاقات والسمات والربط بينهم، إلا أن ذلك يتم على مستوى بيئة لغة إنجليزية مصممة بخوارزميات إنجليزية جُهِّزت لها بالأساس، فلا نجد أي مبادرات في ترجمة واجهة أدوات تحرير الأنطولوجيا إلى اللغة العربية لتسهيل عملية الإنشاء الأنطولوجي في العالم العربي، بل ولا نجد اهتماما بالمجال الأنطولوجي إن صحّ التعبير، مع العلم أنه من بين أدق الميادين المسعفة إلى خدمة اللغة العربية والارتقاء بها على المستوى المعلوماتي، ولعل القصور التي تعاني منه منظومة التعليم في مراحل ما قبل الجامعة يعود بالأساس إلى اتساع الفجوة بين واقع اللغة والمقررات التعليمية الذي خلفه ضعف ربط المجال التقني بالتعليم، مما أثّر سلبا على التنمية المعرفية لدى المتعلمين. وبيت القصيد أن تمثيل الأنطولوجيا للواقع اللغوي يسهم في تطوير التعليم وبلورة مقررات اللغة العربية.

خاتمة

بيت القصيد في هذا المقام، أن الأنطولوجيا مجال معرفي خصب والأساس الفقري للويب الدلالي، فهي عبارة عن خريطة مفاهيمية على شكل شبكة دلالية واسعة، وقاعدة معرفية أساسية أساسها تصنيف المفاهيم والفئات تصنيفا هرميا معقدا ضمن مجال محدد، وذلك لتمثيل المعلومات الموجودة في مدونة ما، وهذه القاعدة المعرفية تتعدى حدود أشكال الكلمات لتتشكل من مجموعة عقد أو حلقات متداخلة ومترابطة، تُصنَّف فيها المفاهيم بطرق لغوية دلالية وأخرى منطقية رياضية، تسعى نحو نمذجة العلاقات والحد من الغموض المفاهيمي والدلالي في عالم المعرفة الحاسوبية.وعليه يمكن إجمال النتائج المتوصل إليها في النقاط الآتي:

- تمثل الأنطولوجيا خريطة معرفية من شبكات كلمات ترتبط فيما بينها بشبكة دلالية، فيشكلان هذان الأخيران بنية الأنطولوجيا وعمودها الفقري.
- لا يمكن الحديث عن الأنطولوجيا بعيدا عن المفاهيم والسمات التي تحدد هوية هاته الأخيرة، مع التركيز على العلاقات لما لها من الأهمية بمكان في ربط كيانات العالم المعرفي الواقعي.
- تسهم الأنطولوجيا في تعزيز كفاءة التطبيقات الحاسوبية لمعالجة اللغات الإنسانية؛ كمحركات البحث والترجمة الآلية.
- تمكن الأنطولوجيا من عملية تنظيم الوثائق وسهولة استرجاعها، كونها تعد بمثابة صياغة واقع استخدام اللغة.
- تعين الأنطولوجيا على تحديد العلاقات بين المفاهيم وبرمجتها بهدف توصيف مجال موضوعي أو معرفي محدد، بطريقة تمكّن الأنظمة الحاسوبية من الاتصال والتواصل مع بعضها البعض.

- تصميم منهجية محكمة في بناء الأنطولوجيا من شأنه أن يسهّل فلسفة العلوم بصفة عامة، فيدفع الفرد إلى الإبداع والابتكار.
- تمثل المدونة جزءا من التحرير الأنطولوجي، فهي إثبات لمادته والبرهنة على صحتها، لما توفره من سياقات لغوية ملموسة على أرض الواقع.
- الفرق بين الأنطولوجيا وشبكة الكلمات، يعود إلى اعتماد الأولى للمدونة، والثانية للمعجم، وأن الأولى قائمة على عملية الحوسبة، بيد أن الثانية نتيجة لعملية الرقمنة.
- تعد أداة تحرير الأنطولوجيا Protege من الأدوات الداعمة بقوة للغة العربية، إذ تسمح بإدراج الأصناف والسمات والعلاقات مع تحديد خواص كل منهم.
- الاهتمام بالأنطولوجيا يعود بالنفع على واقع اللغة العربية، إذ يسهم في التقليل من ذائقة الالتباس الدلالي الذي تعاني منه اللغة العربية في الساحة الرقمية.

الهوامش:

[1] العويشق، عريب بنت عبد الله (2017). هندسة الأنطولوجيا. في: هند بنت سليمان الخليفة (تحرير) علم الدلالة والأنطولوجيا من منظور حوسبة اللغة العربية (27 – 45). الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.

[2] الويسي، ياسين الحسين (1010). مقدمة فلسفية. سوريا: دار الفرق، ص 19.

[3] السعيد، المعتز بالله (2017). كيف نبني مدونة لغوية موسمة تركيبيا للغة العربية بطريقة نصف آلية. الدراسات اللغوية. مج 19، (3)، ص314.

[4] كان هذا التعريف بمثابة البؤرة التي تفشت لتُعتمد من قبل مجموعة من التعريفات الخاصة بالأنطولوجيا.

[5] Gruber, T (1993).Toward principles for the Design of Ontologies Used for Knoweledge. international journal Human-Computer Studies. (43), 901 – 928.

[6] Wilson, R (August 2014). *The Role of Ontologies in Teaching and Learning, P16.*

[7] Ibid, P02.

[8] (2006: 26).

[9] (Roussey & others, 2011: 9-38)

[10] وصل إجمالي المعاجم المحوسبة لهذا المشروع إلى 150 معجما عربيا مختلفا، سواء كانت معاجم اللغة العامة أو المعاجم العلمية المتخصصة، والمنتظر من هذا المشروع أن يكون طويل الأمد.

[11] العويشق، عريب بنت عبد الله (2017). هندسة الأنطولوجيا. في: هند بنت سليمان الخليفة (تحرير) علم الدلالة والأنطولوجيا من منظور حوسبة اللغة العربية، ص32.

[12] النشرتي، مؤمن سيد (2014). نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في إطار التنقيب عن البيانات : دراسة تحليلية. جريدة Cybrarians. (35)، ص17-18.

[13] Kim, J & Choi S (2007). Evaluation of Ontology Development Methodology with CMM-i. *Fifth International Conference on Software Engineering Research, Management and Applications.* 823 – 827.

[14] الأكلبي، علي بن ذيب (2015). دور الأنطولوجيا في دعم محركات البحث الدلالية في البيئة العربية. مكتبة الملك فهد الوطنية. مج 21، (1)، ص203.

[15] Kaloub, A (2013). *Automatic Ontology- Based Document Annotation for Arabic Information Retrieval.* A Thesis Submitted as Partial Fulfillment of the Requirements for the Degree of Master in Information Technology, 8 – 10

[16]XML هي لغة الترميز القابلة للتوسع وهي لغة تستخدم لوصف بنية ومحتوى الوثائق الإلكترونية المتاحة على الويب، كما وتتميز بالمرونة الشديدة.

[17]RDF أي إطار توصيف مصادر الويب، يعد أحد أبرز التقنيات المستخدمة مع تطبيقات الويب الدلالي باعتباره نموذج معياري لوصف البيانات وتنسيق عملية تبادلها باستخدام نظام الميتاداتا Metadata.

[18]السعيد، المعتز بالله (2019). المدونات اللغوية. في: محسن رشوان والمعتز بالله السعيد (تحرير) الموارد اللغوية الحاسوبية، الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، ص53.

[19]العنزي، بدرية بنت براك (2020). المعالجة الآلية للتطور الدلالي وفق لغويات المدونة المحوسبة: دراسة دلالية حاسوبية. اللسانيات العربية. (11)، ص187.

[20] Egbert, J & Baker, P. (2020). *Using Corpus Methods to Triangulate Linguistic Analysis*. New York: Routledge, P4.

[21]السعيد، المعتز بالله (2016). توظيف المدونات اللغوية في تطوير مقررات اللغة العربية لمراحل التعليم العام. التخطيط والسياسة اللغوية. (3)، ص59.

[22] Tognini, B, E (2000). *Lexis in contrast*. In: Granger & Altenberg (editors). Studies in corpus linguistics. Benjamins, Amesterdam, P3-41.

[23]http://www.nooj-association.org/files/app/NooJApp.zip%D8%9B

[24]غازي، عز الدين (2016). بيئة نوج Nooj اللسانية، أداة لمعالجة وتعليم اللغات: تطبيقات صرفية معجمية على العربية. اللسان العربي. (77)، 229 – 267.

[25]http://www.hozo.jp/

[26]https://www.semanticweb.org/wiki/Swoop.html

[27]https://protege.stanford.edu/

[28] المالكي، طارق (2015). أنطولوجيا حاسوبية للنحو العربي. طنطا: دار النابغة للنشر والتوزيع، ص55.

[29]العويشق، عريب بنت عبد الله (2017). هندسة الأنطولوجيا. في: هند بنت سليمان الخليفة (تحرير) علم الدلالة والأنطولوجيا من منظور حوسبة اللغة العربية (27 – 45). الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، ص24.

قائمة المصادر والمراجع

الأجنبية:

- Authoul, A (2012) .*constitution d'une ressource sémantique arabeàpartir de corpus multilingues alignés*.France:THÈSE Pour obtenir le grade de Docteur De L'université De Grenoble.
- Catherine, R & Francois, P & Myoung, A & Oscar, C (2011). An introduction to ontologies and ontology engineering. *In Ontologies in Urban Development Projects*. London: Springer
- Egbert, J & Baker, P. (2020). *Using Corpus Methods to Triangulate Linguistic Analysis*. New York: Routledge.
- Gruber, T (1993).Toward principles for the Design of Ontologies Used for Knoweledge. international journal Human-Computer Studies. (43).
- Kaloub, A (2013). *Automatic Ontology- Based Document Annotation for Arabic Information Retrieval*. A Thesis Submitted as Partial Fulfillment of the Requirements for the Degree of Master in Information Technology.

- Kim, J & Choi S (2007). Evaluation of Ontology Development Methodology with CMM-i. *Fifth International Conference on Software Engineering Research, Management and Applications*.
- Tognini, B, E (2000). *Lexis in contrast*. In: Granger & Altenberg (editors). Studies in corpus linguistics. Benjamins, Amesterdam.
- Wilson, R (August 2014). *The Role of Ontologies in Teaching and Learning*.

العربية:

- الأكلبي، علي بن ذيب (2015). دور الأنطولوجيا في دعم محركات البحث الدلالية في البيئة العربية. مكتبة الملك فهد الوطنية. مج 21، (1).
- السعيد، المعتز بالله (2015). شبكة الكلمات العربية لأغراض الصناعة المعجمية. المجلة المصرية لهندسة اللغة. مج 87، ج 2.
- السعيد، المعتز بالله (2016). توظيف المدونات اللغوية في تطوير مقررات اللغة العربية لمراحل التعليم العام. التخطيط والسياسة اللغوية. (3).
- السعيد، المعتز بالله (2017). كيف نبني مدونة لغوية موسمة تركيبيا للغة العربية بطريقة نصف آلية. الدراسات اللغوية. مج 19، (3).
- السعيد، المعتز بالله (2019). المدونات اللغوية. في: محسن رشوان والمعتز بالله السعيد (تحرير) الموارد اللغوية الحاسوبية (51 – 88). الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.
- السعيد، المعتز بالله؛ ومحسن رشوان (2019). مقدمة في حوسبة اللغة العربية. الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.
- العنزي، بدرية بنت براك (2020). المعالجة الآلية للتطور الدلالي وفق لغويات المدونة المحوسبة: دراسة دلالية حاسوبية. اللسانيات العربية. (11).
- العويشق، عريب بنت عبد الله (2017). هندسة الأنطولوجيا. في: هند بنت سليمان الخليفة (تحرير) علم الدلالة والأنطولوجيا من منظور حوسبة اللغة العربية (27 – 45). الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية.
- غازي، عز الدين (2016). بيئة نوج Nooj اللسانية، أداة لمعالجة وتعليم اللغات: تطبيقات صرفية معجمية على العربية. اللسان العربي. (77).
- المالكي، طارق (2015). أنطولوجيا حاسوبية للنحو العربي. طنطا: دار النابغة للنشر والتوزيع.
- النشرتي، مؤمن سيد (2014). نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في إطار التنقيب عن البيانات : دراسة تحليلية. جريدة Cybrarians. (35).
- الويسي، ياسين الحسين (1010). مقدمة فلسفية. سوريا: دار الفرق.