نظم انتاج المعرفة على الويب:
دراسة تطبيقية للمعالجة الدلالية لجزيئات المعرفة في محركات البحث

د.سيد ربيع سيد إبراهيم
قسم المعلومات ومصادر التعلم كلية الاداب والعلوم الانسانية جامعة طيبة

تاريخ الإرسال	تاريخ القبول	تاريخ النشر
٢٠٢٣/٥/١٧	٢٠٢٣/٧/٢٧	٢٠٢٤/٤/١

المستخلص:

إن القدر من معارف الويب يمكن أن يشتمل محتوى معرفي متجانس وغير متجانس مُعالج بإمكانات بحث واستدعاء مختلفة، ومشتملاً على دلالاتٍ ومفاهيم مختلفة المستويات في التعبير عن المعارف المخصصة لها. وعليه، فإن الويب تحوي محتويات معرفية داخل مواد المعلومات تشمل معارف متداخلة وتكميلية حول الكيانات المعرفية (الموضوعات، أو الأشخاص، أو المؤسسات، أو الدول، أو الأحداث) على وسائطٍ ومنصاتٍ من الويب، الويب الاجتماعية، أو قواعد البيانات Deep Web أو على شبكة الويب المظلمة Dark Web. بل ويمكن أن يُعاد تفكيك الجزئيات المعرفية أو المفردات الدلالية لهذه الكيانات وإعادة ترتيبها وتركيبها مرةً أخرى مكونةً كيانات معرفية ذات هياكل وبنية جديدة تعكس توجهات مختلفة عن تلك التي وردت في سياقها. وهذه الهياكل المعرفية المُولَّدة يمكن أن تكون نتائج معرفية تلبي احتياجات محددة لمستخدمي نظم انتاج المعرفة. وسعت هذه الدراسة إلى تحقيق هدف استكشاف عمل نظم انتاج المعرفة وتوصيف عملها المستقبلي وو اقعها الحالي، لتحديد إمكانات وآليات التحول النوعي في انتاج المعرفة لمستخدميها، معتمدةً في ذلك على استخدام المنهج الوصفي التحليلي وأداة المعايشة والابحار في مو اقع محركات البحث. وقد خرجت الدراسة بمجموعة من النتائج أهمها تمثل في؛ أثبتت الدراسة التطبيقية على عينة الدراسة من محركات البحث الدلالية أن قدراتها على التحول لنظم انتاج المعرفة لا تكون كاملة الا إذا توفرت بها مختلف المقومات والآليات ودعائم العمل، مثل الواجهة الحوارية وقاعدة المعرفة وبناء هياكل المعرفة ومكونات المُعرِّب والوسيط لتكوين جزيئات المعرفة؛ حيث يمثل فقدان أي من هذه الجو انب بعدا تاما عن مفهوم ومؤدى نظم انتاج المعرفة، وهو ما بدا جليا في اجتماع مختلف الجو انب في محرك البحث BING وشبه اكتمالها في GOOGLE.

الكلمات المفتاحية:

نظم انتاج المعرفة، استرجاع المعرفة، محركات بحث الويب، نظم استرجاع المعلومات، قواعد المعرفة، الهياكل المعرفية، معالجة اللغة الطبيعية.

Abstract:

The amount of web knowledge can include homogeneous and heterogeneous cognitive content processed with different search and recall capabilities and include semantics and concepts of different levels in expressing the knowledge assigned to it. Accordingly, the Web contains cognitive contents within information materials that include overlapping and complementary knowledge about knowledge entities (subjects, people, institutions, countries, or events) on media and platforms from the Web, the Social Web, Deep Web databases, or on the Web. Dark Web. Indeed, the cognitive particles or the semantic vocabulary of these entities can be re-dismantled, rearranged, and recombined once again, forming cognitive entities with new structures and structures that reflect different directions from those mentioned in their context. These generated knowledge structures can be knowledge outcomes that meet the specific needs of users of knowledge production systems. This study sought to achieve the goal of exploring the work of knowledge production systems and describing their future work and current reality, to determine the capabilities and mechanisms of qualitative transformation in the production of knowledge for its users, relying on the use of the descriptive analytical approach and the tool of coexistence and navigation in search engine sites. The study came out with a set of results, the most important of which are; The applied study on the study sample of semantic search engines proved that their capabilities to transform into knowledge production systems are not complete unless they have various components, mechanisms, and pillars of work, such as the dialogic interface, the knowledge base, the building of knowledge structures, and the components of the localizer and the mediator for the formation of knowledge particles. The loss of any of these aspects represents a complete dimension of the concept and performance of knowledge production systems, this was evident in the meeting of various aspects in the BING search engine and its near completion in Google.

key words:

Knowledge production systems, knowledge retrieval, web search engines, information retrieval systems, knowledge bases, knowledge structures, natural language processing.

أولا: المقدمة

استهدفت دراسات معالجة المعلومات أساليب التنظيم الدقيق لمفردات وموضوعات مواد المعلومات، وذلك دعماً لتحقيق أفضل مستويات التو افق بين طبيعة الاستفسارات المعبرة عن احتياجات المستفيدين والنتائج المسترجعة المحققة لهذه الاحتياجات. ولأزمنةٍ مديدة، اعتمدت معالجة المعلومات على معالجة الكلمة باعتبارها مَعين المحتوى والدال على هوية الأفكار الواردة في مواد المعلومات، غير أن معالجة الكلمة توقف عند مستويات المعالجة الحرفية اللفظية لسياق المعلومات النصية، بما لم يضمن معه تحقيق أفضل درجات التحقيق في استرجاع النتائج.[١]وقد ارتقى مستوى معالجة المعلومات من الشكل اللفظي للكلمة إلى دلالة ومعنى الكلمة في التحليل الدلالي بتمثيل الكلمة لمصطلح ومفهوم يحدد جو انب الاحتياجات الموضوعية للمستفيدين، مسترجعاً النتائج بمختلف الأبعاد المطروحة في الاستفسار، ولا يزيد عن ذلك إلا بالخارج عن الموضوع مُخِلّاً بمستوى التحقيق المستهدف. وقد ساعدت قدرات معالجة اللغة الطبيعية NLP على دعم تحليل محتوى الكلمات وتحويلها إلى جو انب موضوعية لرسم صورة دلالية لسياق النص؛ حيث مكّن ذلك من الاستفسار عن الموضوعات متعددة المفردات واسترجاع نتائج ذات أبعادٍ متنوعة بأقل صياغة لاستفسارات البحث.[٢]

ومن خلال الربط بين المفاهيم ذات العلاقة باستخدام بنية اللغة الدلالية (الانطولوجيا) وقدرات الويب الدلالي، استطاعت نظم البحث دعم استرجاع المعلومات اعتماداً على الربط الدلالي بين المصطلحات والمحصلة النهائية لاتجاهات البحث أكثر من مضاهاة الكلمات ومطابقة استفسار البحث بحقول وبيانات مواد المعلومات المسترجعة. ما جعل مستخدمي الويب قادرين على استرجاع نتائج متو افقة مع توجهات المجالات الموضوعية وأساليب عرضها أكثر من الوحدات المعلوماتية المصنفة ضمن هذه المجالات والمنتمية إليها من حيث التصنيف فقط. وبالرغم من هذا كله، لم يكن لمستخدمي نظم معالجة واسترجاع المعلومات الرقمية سوى استرجاع وحدات المعلومات المتضمنة في قاعدة بيانات النظام والسابق اضافتها، ويتمثل عمل النظام في تنظيمها وتوفير أساليب استرجاعها فقط، فيما ظل يُعرف

بنظم استرجاع المعلومات.[٣]وقد استحدثت البيئة الرقمية آليات يمكنها معالجة المعلومات الرقمية بما يمكن معه توليد معلوماتٍ جديدة تمثل وحدات معلومات مستحدثة تُضاف إلى جملة ما يحويه النظام من وحدات معلوماتية. ويُطلق هنا مسمى قاعدة المعرفة على قاعدة معلومات النظام حيث تتسم قاعدة المعرفة بآفاق للمعلومات المُخلّقة تفوق قدرات السعات التخزينية المُقدرة؛ حيث تتدفق معرفة لا نهائية يمكن توليدها من مواد المعلومات ومحتواها المتضمن في قاعدة النظام.[٤]

ثانيا: مشكلة الدراسة وفرضيتها

عززت قدرات الذكاء الاصطناعي المفتوحة Open AI من معالجة المعلومات آنياً؛ حيث يمكن لنظام المعالجة تلقي الاستفسارات ومعالجتها مُحدداً جوانبها الموضوعية ومُفسراً لاتجاهاتها وسياقها، ثم القدرة على الاستدلال لطبيعة المعلومات المتوافقة وهذه الاستفسارات من جملة ما تحويه قاعدة المعلومات للنظام، من خلال انتقاء اجزائها المتناثرة بين وحدات المعلومات المحتفلة والعمل على تركيبها وتهيئتها في شكل وحدات معلوماتية مُولّدة يتم تقديمها لمستخدمي النظام.[٥]وقد تمثلت هذه القدرات في بناء نظم المحادثة بالمعرفة Knowledge-driven Conversation System؛ حيث تُبنى هذه النظم على إجراءاتٍ محددة تتمثل في؛ أولا: التقاط نقاط الموضوع من الاستفسارات الموجهة من مستخدمي النظام باستمرار المحادثة، ثانيا: الربط بين تلك النقاط ورسم العلاقات الموضوعية بينها بأكثر من اتجاه، ثالثا: تجميع المعلومات المتوافقة وهذه المسارات الموضوعية ودمجها في ترتيب يعكس الرد على مختلف المسارات في إجابةٍ واحدة. وقد تُوجت هذه الآليات الرقمية بإصدار تقنية محولات الدردشة المُدربة مسبقاً chat generative pre-trained transformer (Chat GPT). فقد بدأت هذه التقنية بمطلع عام ٢٠٢٣ م لتقديم نموذج و اقعي لروبوت يمكنه توليد المعرفة والرد بالإجابات المعرفية وليس المعلوماتية في سياق محادثة مع مستخدمي هذه النظم. وليست هذه التقنية سوى المرحلة الآنية التي انتقلت اليها معالجة المعلومات الرقمية، غير أن ممارستها مازالت في بيئة مغلقة يمكن لمجال المعلومات استثمارها في بناء نظم معالجة المعلومات و إنتاج المعرفة ليتحول علم المعلومات إلى فضاء جديد يستهدف معه التحول بمؤسساته من الاختزان والتنظيم إلى المعالجة والانتاج.[٦]

وتهتم هذه الدراسة بتناول و اقع أدوات معالجة المعلومات الرقمية وقدراتها الحالية المعززة لهذا الانتقال النوعي من أدوات معالجة واسترجاع المعلومات إلى أدوات انتاج وتخليق

المعرفة، فضلا عن التماس البنية الآلية التي يمكن أن تكون عليها معالجة المعلومات الرقمية بما يدعم انتاج المعرفة أو المحتوى الجديد المتو افق واحتياجات مستخدمي هذه الأدوات. ثم تقديم الرؤى والمقترحات التي تعززمن رسم صورة النظام الجديد لإنتاج وتدفق المعرفة من محتوى المعلومات الرقمية على الويب.

تركز الدراسة على تتبع و اقع أدوات معالجة المعلومات الرقمية على الويب فيما يتعلق بأساليب عرض المحتوى وتكشيفه وإجراءات البحث عنه، لبيان ما إذا كانت تدعم آلياتُها قدرات انتاج المعرفة أم لا. وهنا تتمحور ظاهرة الدراسة ومشكلتها البحثية في فرضية (أن أدوات معالجة واسترجاع المعلومات الرقمية على الويب لا تمتلك قدرات انتاج وتوليد المعرفة). و انطلاقا من هذه الفرضية تجمع الدراسة بين اطارين نظري وتطبيقي للتحقق من هذه الفرضية، بما يعززالوصول إلى الرؤى والتصورات المقترحة لتحقيق مفهوم أدوات إنتاج المعرفة الرقمية.

ثالثا: المنهجية:

تتبع الدراسة في مراحلها المختلفة العمل وفقا للمنهج الوصفي التحليلي بدراسة مجموعة من نظم المعالجة الدلالية للمعلومات الرقمية، وتطبيق مجموعة من الإجراءات وشواهد انتاج المعرفة لتبيان مدى جاهزية او احتياج هذه النظم عند التوجه لإنتاج المعرفة واستدعاء المحتوى المعرفي تلبيةً لاحتياجات المستفيدين. وستكون أداتا الملاحظة والابحارفي الويب مقومين أساسيين لجمع المعلومات وبناء محتوى الدراسة. وستعمل الدراسة على تبني عينة من أدوات تنظيم واسترجاع المعلومات فيما يعرف بمحركات البحث الدلالية؛ حيث كونها الأقرب لمستهدفات الدراسة من خلال اختيارعينة قصدية تتمثل في محركات البحث الدلالية على الويب. وسيسبق ذلك، الإطار النظري الممهد لرؤية واستكشاف و اقع أدوات معالجة المعلومات الرقمية يتعلق بالتوجهات الرقمية لإنتاج وتخليق المعرفة من خلال قدرات المعالجة الدلالية والسياقية للمعلومات ودعم بناء القواعد المعرفية، ثم يلي ذلك عمل آليات انتاج المعرفة ومنهجيتها داخل النظم الرقمية المخصصة لذلك. ثم تُختتم الدراسة بالطرح النهائي والرؤى التي يمكن تبنها في معالجة المعلومات للتحول إلى انتاج المعرفة من خلال توصيف نظم انتاج المعرفة بتصورالآليات والإجراءات والمقومات الداعمة لذلك.

رابعا: أهداف الدراسة:

بانتهاء هذه الدراسة يمكن أن تساهم في تحقيق جوانب مهمة لمعالجة المعلومات الرقمية؛ مثل:

١. الانتقال بتوجهات علم المعلومات من إدارة وتنظيم المعلومات إلى نطاق انتاج وتوليد المعرفة.

٢. الإضافة النوعية لهوية مؤسسات المعلومات باعتبارها ستكون مؤسسات لإنتاج وتخليق المعرفة الجديدة وليس فقط ما ترتبط به من مصادر المعلومات المقتناة داخلها.

٣. دعم أدوات تنظيم واسترجاع المعلومات على الويب وتغيير نمط العمل إلى أدوات انتاج المعرفة الرقمية باعتماد آليات نوعية في استرجاع المعرفة.

٤. رفع مستويات التحقيق للمعلومات المسترجعة من معلومات مجردة إلى سياق معرفي يحمل خبرات جديدة للباحثين عن المعلومات.

خامسا: مصطلحات الدراسة:

- الهيكل المعرفي: بناء معرفي من مجموعة من المفردات ذات الخصائص الدلالية والمفاهيمية المتجانسة.
- الجزيئات المعرفية: مجموعة المفردات والكلمات المكونة للهياكل المعرفية المرتبط معا بعلاقات موضوعية أو دلالية.
- مُولِّدات المحادثات: برمجيات تعمل بآليات بناء الحوار مع مستخدمي نظم انتاج المعرفة للرد على الاستفسارات بمجموعة من المقاطع في شكل محادثة أو حوار.
- المعالجة المعرفية: تطور معالجة اللغة الطبيعية لمواد المعلومات على الويب، واستخدام المستوى الدلالي والسياق لتحليل المفردات وتكوين علاقات الهياكل المعرفية.

سادسا: الدراسات السابقة:

ندرت الدراسات في أدب الموضوع في اللغة الانجليزية بما يعكس الحاجة لتطوير نظم استرجاع المعلومات وتحويلها إلى نظم لإنتاج المعرفة من خلال التغيير النوعي في قدرات معالجة اللغة الطبيعية داخل النظم، أيضا تشكيل الهياكل المعرفية واستخدام واجهات الحوار والمحادثة مع مستخدمي النظم. أما بالنسبة للغة العربية فلم يسفر البحث عن دراسات سابقة في أدب الموضوع، وعليه فقد تخصصت دراسات ادب الموضوع في محاور ثلاثة هي:

١. <u>نظم توليد الحوار والمحادثة</u>

ركزت العديد من الدراسات في انتاج المعرفة على برمجيات انتاج الحوار GPT generative pretrained transformers التي تستطيع انشاء حواء تفاعلي مع مستخدميها، والتقاط المفردات التي يطرحها المستخدم للرد عليها بما بتو افق معها من مقاطع أو فقرات في ذات السياق. وعدَّ الكثيرون هذه البرمجيات نظماً لإنتاج المعرفة لما لها من قدرة على تحليل المفردات والرد بفقرات ومقاطع من انتاج الذكاء الاصطناعي. ومن هذه الدراسات دراسة

Luo, C., Liu, D., Li, C., Lu, L., & Lv, J. (2021). Prediction, Selection, and Generation: Exploration of Knowledge-Driven Conversation System. *arXiv* cited at https://arxiv.org/pdf/2104.11454preprint arXiv:2104.11454.

التي عنيت بتناول نظم الحوار أو المحادثة والتركيز على قدرات تحليل المفردات و اقتراح الإجابات والردود المعرفية ذاتيا، مستخدمة آليات بناء الهياكل المعرفية للمفردات و انتقاء الجزيئات المعرفية المتجانسة في الاستفسارات والاجابات. وتختلف الدراسة الحالية عن الدراسة السابقة في تبني نظام متكامل مبني على قاعدة معرفية تستخدم آليات وإمكانات التحليل الدلالي لمعالجة جزئيات المعرفة وإعادة تركيبها.

٢. معالجة اللغة الطبيعية الدلالية

عملت بعض الدراسات على تتبع مستويات معالجة اللغة الطبيعية NLP واستحداث آليات جديدة للتحليل الدلالي لمفردات مواد المعلومات الرقمية؛ حيث يمكن بتطوير معالجة اللغة الطبيعية تمكين النظم الآلية من تحليل لغة الانسان للوصول إلى قدرات الاستخلاص الآلي والترجمة و اقتباس المعلومات والرد على الاستفسارات، بمستوى يقترب من العنصر البشري مستخدمة في ذلك تعلم الآلة machine learning. ومن الدراسات التي ركزت على هذا الجانب دراسة

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2022). Natural language processing: State of the art, current trends and challenges. Multimedia tools and applications, 1-32. cited at https://link.springer.com/article/10.1007/s11042-022-13428-4

التي عنيت بتتبع مستويات معالجة اللغة الطبيعية المتقدمة في تأريخ الحوار التفاعلي مع مستخدمي النظم الآلية والتحديات التي تواجه معالجة اللغة الطبيعية في انتاج إجابات منطقية من الآلات.

٣. **بناء هياكل المعرفة**

استهدفت بعض الدراسات الحديث عن انتاج المعرفة من خلال قدرات النظم الذكية في بناء الهياكل المعرفية؛ حيث عنيت دراسة

Rahman, M., & Cuzzocrea, A. (2020, December). Actionable Knowledge Extraction Framework for COVID-19. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 4036-4041). IEEE, https://scholar.archive.org/work/m75djnwusbfuxkjifmtg35lmxe/access/wayback/https://ieeexplore.ieee.org/ielx7/9377717/9377728/09378398.pdf

بتقنيات بناء الهياكل المعرفية لمفردات متخصصة في مجال كوفيد١٩، يمكن من خلالها تكوين علاقات موضوعية ودلالية بين المفردات المختلفة واستخدامها لإعادة تشكيل المجالات الموضوعية المتخصصة في هذا المجال. يمكن بواسطتها تحسين قدرات الرد الآلي على الاستفسارات المتعلقة بجائحة كوفيد ١٩ والمعلومات المختلفة التي تتنوع ردودها وتتباين وفقا لنوعية الاستفسار المطروح. وتختلف الدراسة الحالية في تحقيق التكاملية بين قدرات بناء الهياكل المعرفية وبين امكانات البحث والاسترجاع وفقا لاحتياجات مستخدمي الويب.

المبحث الأول: المعالجة الدلالية والسياقية للمعلومات الرقمية

اعتمدت معالجة المعلومات على اللغة الطبيعية بما لها من خصائص تتو افق وأدوات تنظيم واسترجاع المعلومات الرقمية؛ حيث تبنت نظم المعلومات الرقمية تطويروتحديث التكشيف بعيدا عن المكانز والأدوات الموضوعية المقيدة، و اقترابا من الكلمات والمصطلحات وفقا لعرضها في سياق وحدات المعلومات النصية الرقمية. وقد بدأ مفهوم معالجة اللغة الطبيعية natural language processing (NLP) مع استخدام الآلة في إجراءات التكشيف وتحليل محتوى وحدات المعلومات، بدءً من الطريقة الإحصائية بالتقاط الكلمات الأكثر تكرارا، ثم استخدام المعاجم الآلية وتحليل الكلمات إلى التحليل الدلالي ورسم شبكة العلاقات بين المصطلحات الموضوعية في محتوى مادة المعلومات.[٧]ويمثل معالجة المعلومات لغويا أول خطوة على طريق انتاج المعرفة؛ فإذا ما ألمّت نظم معالجة المعلومات بآليات التحليل الدلالي لمحتوى مواد المعلومات وإعادة تمثيل المحتوى بشكل أكثر مرونة لخدمة موضوعات أخرى ومجالا أكثر مما جاء عليه، عندئذٍ، تستطيع نظم معالجة

المعلومات التحول إلى انتاج المعرفة الرقمية والرد على الاستفسارات المختلفة بمحتوى مرن قابل للتشكيل وفقا لاحتياجات مستخدمي النظم.(٨)

وقد مثلت اجراءات تطوير معالجة اللغة الطبيعية تاريخا متعدد المراحل هدفت جميعها إلى اشتقاق المعلومات واسترجاعها بما يدعم أفضل مستويات التحقيق، بداية من مرحلة الاشتقاق المبني على الإحصاء ثم الاشتقاق المبني على ميتاداتا مواد المعلومات والاشتقاق المبني على المعاجم اللغوية ثم الشبكات العصبية، غير أن المرحلة الرئيسة التي ساهمت في بداية التفكير بإنتاج المعرفة كانت في التحليل الدلالي لمحتوى مواد المعلومات واستخدام الأنطولوجيا في بناء قواعد معرفية. وقد اكتملت دعائم التوجه لإنتاج المعرفة بمرحلة التكشيف السياقي والتحليل الحسي للغة الطبيعية، الذي استهدف المعاني والمحتوى المعرفي الذي يسكن جسد الكلمات والمصطلحات النصية في مواد المعلومات الرقمية. ولذا فإن التركيز على معالجة المعلومات هنا يتمثل في جانبي المعالجة الدلالية والسياقية للمعلومات الرقمية مستهدفة دعم قدرات إنتاج المعرفة.(٩)

١/١ مستويات المعالجة الدلالية والسياقية للمعلومات الرقمية

١/١/١ المعالجة والتنظيم الدلالي للمعلومات

يمثل الوصول إلى معنى ودلالة الكلمات والجمل أهم أهداف المعالجة الدلالية لمواد المعلومات؛ ويمكن لأنظمة المعلومات تحديد دلالات الكلمات فقط من خلال تعريف هذه الكلمات ضمن أطر وصف المعانيRDF في انطولوجيا الويب، وربط هذه المفاهيم ببعضها البعض فيما يتعلق بالعلاقات اللغوية أو الموضوعية. ومثال ذلك أن كلماتٍ مثل السفن والشراع والأسماك تدل على البحار والأنهار والماء بشكل عام، بما يعزز من معالجة الاستفسارات حول البحار والمياه والنتائج المسترجعة حولها.(١٠) وعلى ذلك فإن عمليات المعالجة الدلالية للمعلومات تتوزع على مرحلتين هما مرحلة المعالجة لمحتوى مواد المعلومات المختزنة، ومرحلة الاستفسار للرد على احتياجات مستخدمي النظم، وتعد المرحلة الأولى الأجدر باهتمام نظم المعالجة لما لها من تأثير على كشف علاقات المصطلحات ببعضها ومن ثم إيجاد الروابط المنطقية لهذا المحتوى، بل وإمكانية إعادة تأويله مرة أخرى على أكثر من كيف. ولعل أهمية الكشف الدلالي للمحتوى وإعادة تشكيله دلاليا تتمثل تمكين نظم المعالجة من التحرر من قيد المحتوى ومواد المعلومات للخروج إلى أبعاد أكثر رحابة في عرض المحتوى واستخدامه للرد على مستويات ونوعيات مختلفة من استفسارات البحث عن المعلومات، الذي يمكن أن نعتبره أول مستويات انتاج المعرفة داخل النظم.(١١)

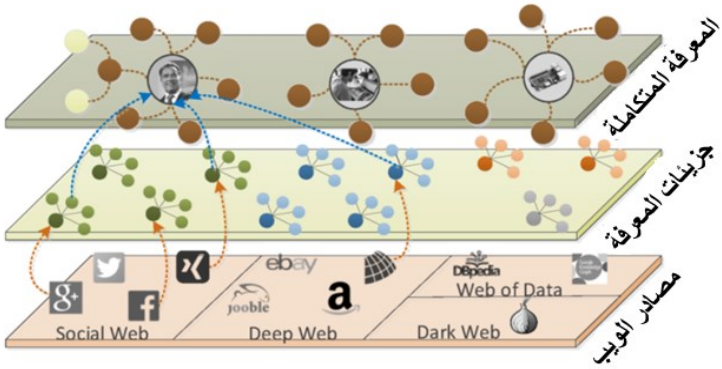
٢/١/١ المعالجة السياقية للمعلومات

تتبنى المعالجة الدلالية تحليل الكلمات المكونة للجملة ومدى تفاوت دلالات المعاني التي تشملها، أما المعالجة السياقية فإنها تعمل على فهم المعاني العامة المشتقة من السياق وعموم الجمل التي يتكون منها النص. وعلى ذلك فإن فهم السياق يمثل خط التماس مع دوائر المعرفة الكامنة في النصوص، وقدرة التحليل على توضيح هذه المعاني وتطويع استخدامها في حالات متعددة. ولعل المعالجة السياقية يمكنها توضيح الجانب المعرفي ومستهدفات المحتوى الفكرية؛ حيث يمكن للسياق الذي وردت فيه كلمات ذات دلالة على الاستياء والقلق بأن تضع تفسيراً جلياً للسؤال عن الوقت الوارد في السياق، فالمعنى الدال المباشر للسؤال عن الوقت هو تقدير الفترات الزمنية، بينما يكون التفسير السياقي أنه انصراف عن المتابعة أو الاعراض عن الاستمرار في أمرٍ ما.(١٢) وعلى ذلك، فإن المعالجات الدلالية والموضوعية تأتي مباشرة من طبيعة الكلمات والجمل المكونة للنصوص ولا يمكن الوصول إلى هذه الدلالات الموضوعية في غياب كلمات ومصطلحات معبرةً عنها بشكل واضح، بينما تتمكن المعالجة السياقية للمعلومات من الكشف عن الدلالات والأفكار التي تم استنباطها من البناء العام لوحدة المعلومات أي بشكلٍ أكثر وضوحا يمكن للمعالجة السياقية الكشف عن الأفكاروالدلالات التي ليس لها جمل مباشرة في النص بينما جاءت من بين معاني الجمل والسياق العام الذي جاء عليه النص.(١٣)

لقد دعمت تقنيات الشبكات العصبية قراءة النصوص بآليات تستطيع الوصول إلى تقسيم النصوص لأجزاء محددة والربط بينها بعلاقات موضوعية ودلالية، إلى جانب قدرة هذه الآليات على رصد الكيانات المختلفة داخل النص مثل الموضوعات والأشخاص وتحديد الإشارات الدالة عليها والربط بينها بحيث تستطيع تأويل جميع ما ذُكر عن هذه الكيانات في مواضع مختلفة من النص واستخلاص التوجهات التي أُثيرت حول هذه الكيانات.(١٤) ومن الناحية البرجماتية الو اقعية، فإن وظيفة النص والمحتوى المعلوماتي هي إيصال رسائل فكرية وموضوعية محددة قد تقترب أو تبتعد بدرجة ما عن المعاني والدلالات المباشرة التي تحملها الكلمات، ويصدق ذلك على الفجوة الهائلة بين النصوص الأدبية والعلمية التي تحويها مواد المعلومات؛ فالنصوص الأدبية غالبا ما تحمل معاني أبعد من دلالات الكلمات المستخدمة وتمثل النصوص فيها منظومة جمالية لإيصال أفكار حسية محددة. أما النصوص العلمية فإنها تحمل كلماتٍ جامدة تستخدم للدلالة المباشرة على الحقائق العلمية المستهدفة من نصوصها.(١٥)

٣/١/١ المعالجة المعرفية للمعلومات

تتجه المعالجة المعرفية إلى الدلالات المعرفية التي تجمع شتات المعلومات حول الكيان الواحد، وذلك من خلال التعرف على هذه الجزيئات ثم جمعها في قطاعات متجانسة، ومن ثم تكوين الصورة المتكاملة للكيان المعرفي. ويتناول هذا المبحث أركان انتاج واسترجاع المعرفة والتكامل من مواد معلومات الويب المختلفة، فضلاً عن تقديم رؤية شاملة لإنتاج المعرفة الدلالية والتكامل الذي يمكن أن يولد هياكلَ معرفية من جملة مصادر الويب وفقا لاحتياجات مستخدمي النظام. وتقتضي هذه الرؤية وضع آليات محددة لتمثيل كيانات المعرفة داخل محتوى مواد الويب التي تنتمي إلى نفس المجال لتكامل أجزاء المعرفة حول هذا المجال، ثم ضوابط تكاملها في بناء معرفي يتغلب على المتعارضات أو الاختلافات التي تنتج عن الربط الخاطئ بين بعض أجزاء هذه المجالات المعرفية.



شكل رقم (١) مستويات تجميع و إنتاج جزيئات المعرفة[١٦]

ويوضح الشكل (١) طبيعة المعلومات المتناثرة هو كيان مادي (شخصية) قد توجد في قواعد البيانات المختلفة ومنصات التواصل الاجتماعي المتعددة، تحمل معلوماتٍ عنه ومنه؛ حيث يمكن ربطها جميعا لجمع المعارف المختلفة والمولدة من هذه القطاعات عن هذا الكيان. لقد كانت ومازالت أدوات البحث في جُلِّ عملياتها تستهدف استرجاع المعلومات المباشرة، مثل البحث عن الملفات الشخصية على الويب للأعلام دون ربط ذلك بمجموعات العلاقات المعرفية التي تصل هؤلاء الأشخاص والكيانات الاعتبارية الاخرى بشكل أساس. ويعزى ذلك لأنها تقتصر على جزء واحد فقط من الويب (شبكة المستندات) ولا تدمج المعرفة من مصادرها المختلفة.[١٧] ويمثل العثور على مواد وعلاقات معرفية إعادةً لإنتاج المعرفة والتكامل في مجالاتها المتنوعة، وهو ما يحتاج مناهج تكامل واسترجاع المعرفة الجديدة إلى

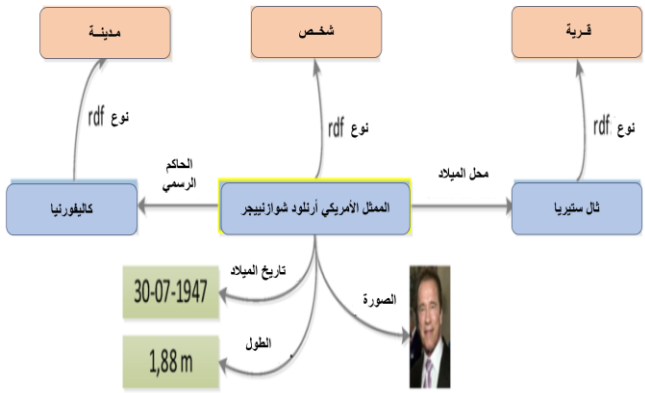
التعامل مع المعلومات غير المتجانسة الممثلة في طرائق مختلفة (منظمة أو شبه منظمة أو غير منظمة) وتنتشر عبر مصادر الويب.[١٨]

١/٣/١/١ مراحل المعالجة المعرفية للمعلومات الرقمية

تنطوي عمليات المعالجة المعرفية على مراحل متواترة تشتق من خلالها نظم معالجة المعرفة مختلف الكيانات المعرفية مثل أسماء الأشخاص والموضوعات والأعلام والأشياء، وفضلا عن الدقة التي يجب أن تتمتع بها هذه الكيانات من التحديد والوصف، فإنه يجب رسم العلاقات الموضوعية بين هذه الكيانات لرسم الشبكة المعرفية داخل النظام بشكلٍ واضح. وتتمثل مراحل المعالجة المعرفية في التالي[19]

المرحلة الأولى: اشتقاق الكيانات

تنتشر الكيانات الموضوعية داخل نصوص محتوى مواد المعلومات سواء أكان ذلك من خلال التصريح بالحديث عن هذ الكيانات أم من خلال الإشارات المستخدمة للدلالة عليها أو حتى تلك الصور الغامضة التي يمكن أن تكون عليها هذه الكيانات داخل النص ويكون الحديث عنها بشكلٍ ضمني. وعليه، فإن المهمة الأولى للمعالجة المعرفية تتمثل في رصد هذه الكيانات أو الإشارات الدالة عليها وتنقيتها وتصنيفها موضوعيا أو بحسب التخصص الواردة به، حتى يمكن من خلالها الكشف عن مستهدفات المحتوى المعلوماتي وتوجهاته المعرفية.[٢٠] ويمكن تطويع آلياتٍ مختلفة للقيام بمهمة الاشتقاق تتمثل في أولا: استخدام المعاجم اللغوية والقواميس المتخصصة في مجال محدد لرصد هذه الكيانات ومن ثم تتبع ذكرها في مجمل النص المعلوماتي وحصرها بمصطلحاتها المباشرة أو اختصاراتها أو استهلالاتها، ثانيا: آليات تعلم الآلة **machine learning** حيث يمكن لنظام معالجة المعلومات استخدام قاعدة معرفية تستكشف محتويات مادة المعلومات بمطابقة ما لديها من كيانات معرفية واشتقاقها بصورة مباشرة أو من خلال الوصف العام الذي ورد عن هذه الكيانات، وهو ما يسمح بالكشف عن الكيانات بتحليل النص موضوعيا ومن السياق في ضوء عرض المعلومات.[٢١]



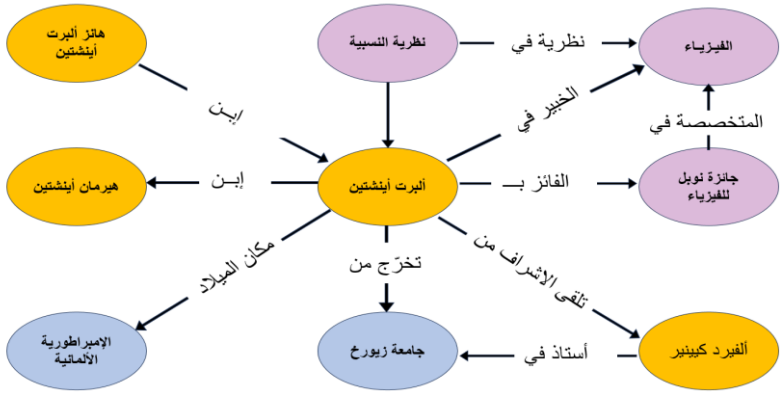
شكل رقم (٢) نموذج للكيانات المشتقة بالمعالجة المعرفية[٢٢]

ويوضح الشكل (٢) مثالا لذلك من خلال تجميع المعلومات حول شخصية الممثل الأمريكي شوازنيجر عندما كان حاكما لولاية كاليفورنيا الامريكية، فقد تم تمثيل معلومات الحكم مع المعلومات الشخصية كطول القامة والصورة الشخصية ومع المعلومات حول مكان الميلاد التي تم تجميعها من محتويات معلوماتية متباينة بمثل يعكس هيكلا معرفيا حول هذا الشخص، يمكن لأدوات انتاج المعرفة تقديمه لمستخدميها بديلا عن المعلومات المشتتة في مواد الويب. ويوضح ذلك بشكلٍ و اقعي العلاقات بين الكيانات المختلفة مثل الأماكن والاحداث والأشخاص في هيكل معرفي واحد.[٢٣]

المرحلة الثانية: اشتقاق العلاقات

تبدأ مرحلة رسم العلاقات بين كيانات المحتوى المعلومات بانتهاء تحديد مجمل هذه الكيانات ومستويات تواجدها داخل مواد المعلومات، ويمكن لنظم معالجة المعلومات رسم نوعين من العلاقات بين الكيانات؛ يتمثل أولهما في شبكة العلاقات الدلالية بين الكيانات حيث غالبا ما يحمل تصريف المصطلح الدال على الكيان ارتباطاً موضوعيا بالمصطلحات ذات العلاقة الصرفية وفقا لمبادئ اللغة العربية خاصة وباقي اللغات بنسبٍ متباينة. أما ثانيها فهي شبكة العلاقات المعرفية حيث يمكن للكيانات المعلوماتية داخل مواد المعلومات إعادة تشكيل وتوجيه المحتوى المعلومات خاصة إذا كان يضم مواد كُتبت لتوجهات و أفرع تخصصية مختلفة. ويشكل النوع الثاني خاصة أساس بناء قواعد المعرفة وآلية عمل نظم انتاج المعرفة؛ حيث يكون المستهدف هو تمثيل الكيانات المعلوماتية داخل محتوى النصوص لخدمة مجالات ومستويات معرفية لم يتم تضمينها داخل محتوى مادة المعلومات أو استهدفها النص بالتعرض لها والحديث عنها.[٢٤] ويمكن تحديد فارق الشكل بين النوعين

بمثال استخراج الكيانات التي ترتبط بعلاقاتٍ دلالية مثل السفن والبحاروالأسماك بشبكة علاقات معرفية يجمع بين هذه الكيانات المرتبطة دلاليا وكياناتٍ معرفية مثل الاقتصاد والأمن الغذائي. وعليه فإن شبكتي العلاقات الدلالية والمعرفية يمثلان أساس بناء قواعد المعرفة في نظم انتاج المعرفة، وبواسطتهما يمكن إعادة تمثل المحتوى واستخدامه لعدد غير متناهي من الإجابات والاحتياجات المعرفية للباحثين عن هذه الكيانات المعرفية.[٢٥]



شكل رقم (٣) اشتقاق العلاقات الدلالية والموضوعية في المعالجة المعرفية[٢٦]

يوضح الشكل (٣) أن بناء الهياكل المعرفية يقوم عل أساس الكيانات التي تحمل دلالات مستقلة بذاتها ويتم من خلالها الربط والاستدلال على المعارف، ثم الكيانات التي توضح طرف العلاقة الموضوعية بين مجالين مختلفين أو في نطاق المجال المعرفي الواحد بين تخصصاته. فقد أوضح الهيكل المعرفي كيفية الربط بين كيانات تنتمي لمجال الشخصيات والاعلام بكيانات معرفية أخرى تتنوع بين هيئات أكاديمية ومجالات موضوعية.[٢٧]

المرحلة الثالثة: تشكيل الهياكل المعرفية

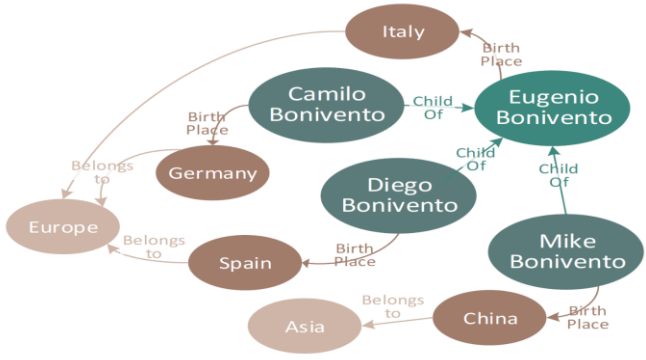
تمثل الهياكل أو الرسوم البيانية للمعرفة الناتج النهائي لعملية توليد المعرفة من جزيئات محتوى مواد المعلومات الرقمية، وذلك باعتبارها صورة جديدة للكيان المعرفي بما يحيطه ويربطه من علاقات معلوماتية ودلالية بالمجالات الأخرى.[٢٨]وقد سكَّ محرك البحث Google مصطلح الرسوم البيانية للمعرفة عام ٢٠١٢م، للتعبير عن قدرات الربط الدلالي واستخدام لغة الانطولوجيا RDF في معالجة واسترجاع المعلومات، حيث يتجه الاسترجاع إلى استدعاء كل المعارف المرتبطة بالموضوع أو الكيان موضع البحث وإعادة ترتيبها ودمجها في نتيجة واحدة بدلا من مجموعة مواد معلوماتية عن ذلك الكيان. ومن حيث التسمية تتبنى هذه الدراسة تسميتها بالهياكل المعرفية لما تعنيه من إعادة بناء المعارف وترتيبها في بنية

جديدة عن كل كيان موضوع كان أم علم أم شخص، أما من حيث المفهوم فقد تواترت مفاهيم كثيرة تعبر عن الرسوم البيانية للمعرفة استقرت دلالاتها جميعا عام ٢٠١٧م حول المفاهيم التالية:[٢٩]

- أنها عبارة عن وصف للكيانات في العالم الحقيقي وعلاقاتها المتبادلة مُنظمةً في رسوم بيانية.
- أنها عبارة عن تحديد للفئات والكيانات المعرفية في مخططات محددة.

ولعل أهم مكونات بناء الهياكل المعرفية تتمثل في ثلاثة جوانب هي؛ العلاقة، والكائن، والمجال؛ حيث تشترك هذه العوامل الثلاث في تجميع جزيئات المعرفة وتصويرها بأبعاد موضوعية متنوعة تتكامل جميعا في هيكل واحد.

ويمكن انتاج هياكل المعرفة في سياق تقنيات الويب الدلالي W3C RDF وRDFS وOWL وSPARQL؛ حيث يُبنى هيكل المعرفة على ثلاثة عناصر أساسية تتمثل في مخطط العلاقات والمعلومات غير المتجانسة والدلالات الموضوعية واللغوية لكيانات محتوى مواد المعلومات.[٣٠]وبذلك، فإن عملية تكامل البيانات الدلالية تخلق مساحة معلومات مترابطة تسهل إنتاج المعرفة المستمدة من مواد المعلومات، مما يوفر رؤية ٣٦٠ درجة للمحتوى المعلوماتي لمجموع المواد في قاعدة المعرفة. وهناك طريقتان رئيستان لدمج محتوى المعلومات؛ هما أولا: الدمج المادي، وفيه يتم نقل المعلومات إلى مستودع مركزي، وثانيا: الدمج الافتراضي؛ حيث تظل المعلومات متضمنة في موادها، ويتم تنفيذ التكامل في وقت استدعاء المعارف الناتجة عن الدمج المطابقة لاستفسار البحث. وتعد الطريقة الافتراضية أكثر ملاءمة لسيناريوهات التكامل وإنتاج هياكل المعرفة؛ حيث تتغير محتويات مواد الويب بشكل ديناميكي تفاعلي.[٣١]



شكل (٤) تشكيل هياكل المعرفة من وحدات المعلومات[٣٢]

يوضح الشكل (٤) تشكيل هياكل المعرفة من خلال آليتي التصنيف والربط؛ حيث يتم أولا تصنيف مجموعات الدلالات والسياقات وفقا لطبيعة توجهاتها الموضوعية والمعرفية، ومن ثم تمييز هذه التصنيفات بأنها مجموعات متجانسة من المعلومات. ثم يأتي بعد ذلك الربط بين مجموعات المعلومات المتجانسة و انشاء علاقات معرفية بينها وفقا لما بينها من ترابط معرفي، مكونةً فيما بينها علاقات متعددة المستويات للخروج بالهيكل المعرفي النهائي. فمثالا، تم تصنيف China, Spain, Germany and Italy الى مجموعة واحدة متجانسة تعبر في دلالاتها المعرفية والسياقية تشير إلى مجموعة معروفة من الدول.[٣٣] أيضا فإن Camilo, Diego and Mike تم تصنيفها إلى أسماء ثلاثة أشخاص بينهما علاقة الاخوة و انتماؤهم لأبٍ واحد، ثم تم الربط بين هذه الأسماء وأسماء الدول التي ولِدُوا فيها، وهذا يمثل مستوى ونوع من العلاقة اختلف عن العلاقة التي تربط الدول Spain, Germany and Italy بالقارة التي تنتمي إليها في الموقع الجغرافي وفقا لتصنيف الأماكن والدول على قارات العالم. وعليه يمكن توسيع بناء العلاقات والترابط بين مجموعات من المعلومات على مدى أكبر من التنوع في المجالات الموضوعية والمعرفية، فيما يمكن أن يمتد إلى شبكة علائقية عامة تغطي كافة مواد معلومات الويب داخل قواعد بيانات نظم انتاج المعرفة.[٣٤]

٢/١ مولدات المحادثات الاسترجاعية Conversational search transformers

تمثل مولدات المحادثات الاسترجاعية طفرةً في واجهات بحث واسترجاع المعلومات؛ حيث تحولت هذه الواجهات من استرجاع لمواد المعلومات إلى التحاور التفاعلي بين النظام والمستخدم، بل تجاوزت ذلك إلى إجراء حوار يقتضي من النظام الرجوع إلى قاعدة المعرفة واختيار مجموعة الإجابات التي تمثل استدعاء لمقاطع معرفية محددة يمكن من خلالها تلبية الاحتياجات الموضوعية بمستخلصات معرفية موجهة دون الحاجة لاستعراض مجموعات من النتائج. وتتسم هذه المولدات بمجموعة من الخصائص تتمثل في:[٣٥]

١. الاسترجاع المفتوح: حيث يتم من خلاله تمثيل عملية البحث والاسترجاع في شكل حوار متبادل بين النظام والمستخدم، يصل إلى مستويات معمقة أو متعدد المراحل في الاستفسارات والردود.

٢. الحواروالمحادثة: حيث لا تتم عملية الاسترجاع للمعارف هنا في شكل نتائج اووحدات معلوماتية إنما يتم تمثيلها في شكل حوارذات إجابات وتفاعل منطقي، بحيث تتمثل النتائج في ثنايا الحديث الدائر.

٣. التماس المعلومات: حيث تتطلب محادثة البحث عن المعرفة المستدعاة عادةً عدة دورات لتبادل المعلومات للسماح للمستخدم بتوضيح حاجاته المعلوماتية، وتقديم الملاحظات وطرح أسئلة التدقيق والتحديد. وفي هذه العملية يتم الوصول للمقاطع المعرفية المثلى من خلال سلسلة من التفاعلات بين المولد والمستخدم. تكون وهو ما يجعل هذه الإجابات بشكلٍ عام أطول منها في نظام استرجاع المعلومات.

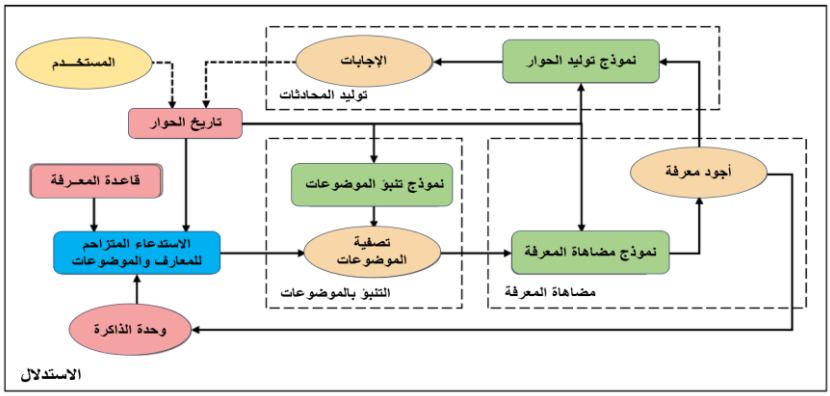
٤. احتياجات معلوماتية مؤصلة: حيث يجعل المولد المستخدم في تحديد دائم ومتواصل لما يريده من معلومات، ويتحقق ذلك بأن مستخدمي النظام يعمدون إلى صياغة استفساراتهم باستخدام الكلمات الدالة المتو افقة والاجابات المتوقعة في أذهانهم، وهو ما يجعل النظام يكون العلاقات الدلالية والموضوعية بين المقاطع المعرفية وإعادة صياغتها في هيكل معرفي يناسب توجه ودقة الإجابة المرجوة.

٥. تأريخ السياقات: يستخدم نظام المحادثة الترتيب الزمني وتسلسل الحوار مع مستخدميه بحيث يمكن له تجميع مختلف السياقات التي وردت من أول الحوار إلى آخره واستخدامه في إعادة تمثيل الإجابات المستخرجة من قاعدة المعرفة؛ بحيث تتعدد الكيانات المعرفية ويزيد من قدرة الربط على هذه الكيانات في هياكل معرفية أكثر تشابكا وتعقيدا بما يدعم الهدف المعرفي لمستخدم النظام.

١/٢/١ هندسة مولدات المحادثات

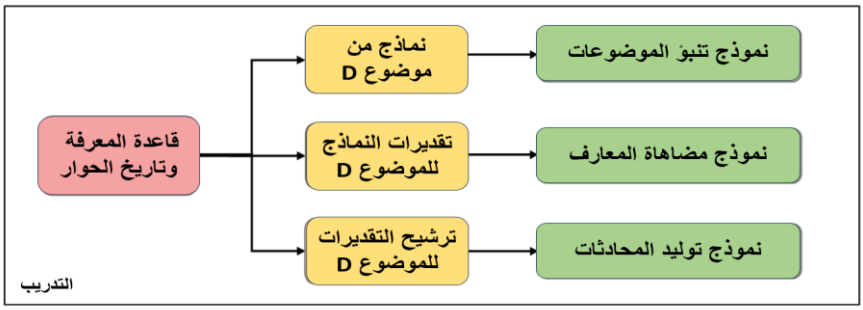
إن استخدام المعالجة المعرفية وآليات متقدمة من معالجة اللغة الطبيعية NLP قد جعل تركيز الباحثين في مجالات انتاج المعرفة تتحول إلى إجراءات وآليات توليد المقاطع والفقرات التي ستمثل النتائج المعرفية المقدمة لمستخدم النظام، غير أن ثمة جانب آخر مهم لم يحظ بنفس التركيز يتمثل في واجهة التحاور والاسترجاع التي تلعب دورها مولدات المحادثات الاسترجاعية؛ فهي تستقبل الكلمات والدلالات المعبرة عن السياق المعرفي المطلوب، فضلا عن تدقيق المعارف المطلوبة بطرح استفسارات التدقيق في شكل محادثة أو حوار ثنائي. وعليه، فإن هندسة الإجراءات وتنظيم آليات عمل المولدات يمثل نصف درجة نجاح نظام انتاج المعرفة في استدعاء المقاطع أو الفقرات المطلوبة. ولقد طورت أنظمة التعلم الآلي آليات التضمين والتركيب استنادًا إلى الأساليب الإحصائية والدلالية، بما يمكن معه وضع الكلمات

المشتقة من نصوص عدة النص الخام في سياق دال ذات هدف معرفي، وهو ما يمثل امتدادا لتطوير أنظمة فهم اللغة الطبيعية (NLU).[٣٦] ونتيجة لذلك، فإن للمولدات دورٌ مهم في تشكيل وتحديد طبيعة المقاطع المستدعاة من حيث استخراج تمثيلات الكلمات السياقية، وتوسيع وتضييق نطاق هذه الأفكار إلى الحدود المعرفية الملائمة. وقد تطورت بنية المولدات الحديثة في معالجة اللغة الطبيعية NLU بتضمين الكلمات في نماذج لغوية تحقق نتائج معرفية غير مسبوقة.[٣٧]



شكل رقم (٥) المرحلة الأولى لهندسة وبنية مولدات المحادثات المعرفية (الاستدلال)[٣٨]

يوضح شكل (٥) مرحلة الاستدلال في بنية مولدات المحادثات المعرفية التي تتمثل في استدعاء للفقرات والمقاطع لبناء الحوار المعرفي بين النظام ومستخدمه، اعتمادا على تفاعلات دلالية وعلاقات موضوعية بين الكيانات داخل محتوى قاعدة المعرفة. ويتحقق عمل النظام بتوفير دعائم نجاح عملية الاستدلال؛ حيث يعد الاستدلال أساس بناء الحوار بين النظام والمستخدم في ربط الجمل والفقرات من جانب المستخدم بتلك الكلمات والالفاظ التاريخية التي تلقاها النظام في وقتٍ سابق من محادثتهما. ويقوم النظام بإجراء استدعاء تقريبي للمواضيع بناءً على العبارات والجمل المؤرخة لديه، ثم استخدام هذه الموضوعات ووحدة الذاكرة لاستدعاء المقاطع والفقرات المقابلة لها في قاعدة المعرفة.[٣٩] ويفرز نموذج التنبؤ بالموضوع العديد من الموضوعات بناءً على العبارات والاقوال المؤرخة، ومن ثم يقارن النتائج بموضوعات الاستدعاء المُتنبأ بها والمشتقة من حديث المستخدم، ونتيجة لهذا العمل المتواصل يمكن استدعاء المقاطع والفقرات المعرفية الأنسب والأكثر ملاءمة للموضوع.[٤٠]



شكل رقم (٦) المرحلة الثانية لهندسة وبنية مولدات المحادثات المعرفية (التدريب)[٤١]

يوضح شكل (٦) مرحلة التدريب في بنية مولدات المحادثات المعرفية؛ حيث يتم استخدام المعلومات المعاد بناؤها لتدريب الوحدات بشكل منفصل. ففي مرحلة الاستدلال، يتم تجميع الكلمات المنطوقة لاستجابة المستخدم والنموذج في مجموعة معلومات مؤرخة، التي تتم معالجتها بشكل تعاوني من خلال خوارزميات استدعاء الموضوع والمعرفة الخام ونموذج التنبؤ بالموضوع ونموذج مطابقة المعرفة للحصول على أفضل الفقرات والمقاطع المعرفية قبل إرسالها إلى الحوار، ومنه إلى النموذج التوليدي للمحادثة بإخراج جملة الرد لاستكمال الحوار. ويتم اختيار المقاطع المعرفية المتو افقة مع أفضل موضوع في الاستدعاء التقريبي والألفاظ التاريخية في نموذج مطابقة المعرفة، وتصنيف المعرفة المستدعاة إلى مستويات في الدقة للحصول على أفضل المعرفة، ويتم ذلك كله من خلال تأريخ الحوار وهيكلة المعارف والمفاضلة بينها لتحقيق أفضل استجابة ممكن من النظام.[٤٢]

٢/٢/١ متطلبات بناء مولدات المحادثات المعرفية

يتطلب عمل نظام الحوار أو المحادثات الموجهة أربع وحدات:[٤٣]

الوحدة الأولى: فهم اللغة الطبيعية (NLU)؛ حيث تقوم هذه الوحدة بتحويل رسالة المستخدم الأولية إلى كتل دلالية، إلى جانب تصنيفات الموضوعات وتوجُه المستخدم.

الوحدة الثانية: تتبع حالة الحوار (DST)؛ حيث تقوم هذه الوحدة بمعايرة حالات الحوار بشكل متكرر بناءً على المدخلات الحالية وتاريخ الحوار. وتتضمن حالة الحوار إجراءات المستخدم ذات الصلة وسياق الردود المحددة لموضوعات المحادثة.[٤٤]

الوحدة الثالثة: تعلم سياسة الحوار؛ حيث تعمل هذه الوحدة بناءً على حالات الحوار التي تمت معايرتها من وحدة DST، وهنا تقرر هذه الوحدة الإجراء التالي لتوجيه الحوار.

الوحدة الرابعة: توليد اللغة الطبيعية (NLG)؛ حيث تقوم هذه الوحدة بتحويل إجراءات الحوار المحددة إلى لغة طبيعية على مستوى واجهة المحادثة، التي تمثل عادةً الشكل النهائي للاستجابة. ويتضمن عمل الوحدة أيضا تتبع حالة الحواروتعلم سياسة الحوار مدير الحوار (DM)، المتحكم المركزي في نظام الحوار الموجه المجالات المعرفية المحددة. وتتفاعل هذه الوحدة مع قاعدة معارف خارجية (KB) لاستدعاء الفقرات والجمل الأكثر اتساقا مع حوار المستخدم.(٤٥)

إن تصميم أنظمة مولدات المحادثات الاسترجاعية يستهدف إنشاء جملٍ صحيحة تلقائيًا باللغة الطبيعية، لذلك فإن أحد أهم مكونات هذه الأنظمة هو نموذج منشئ اللغة. وبمجرد بناء نموذج منشئ اللغة أو عمله، تظهر الحاجة لتحسين إنشاء سلاسل عبارات وجمل صالحة من السلاسل الفرعية التي ستتكامل فيما بعد مكونةً جملة كاملة تحمل المضمون والمعرفة المراد توصيلها للمستخدم. وبالنظر إلى عدد المتغيرات الموضوعية والسياقية والدلالية التي توضع في نظام حسابي ومنطقي، وإعطائها قيماً لغوية مقابلة لها، فإنه يمكن استخدام خوارزمية التعلم المعزز لتوليد جملة باللغة العربية أو الانجليزية، أو أي لغة طبيعية أخرى يتم استخدامها في إيصال معارف محددة وذات هدف محدد لمستخدم النظام. وعليه، فإن استخدام المعلومات المخزنة في مجموعة من متغيرات البرنامج و إنشاء جمل بلغة طبيعية تمثل هذه المعلومات يعد مهمة صعبة الأداء تتم على مرحلتين؛ حيث تتمثل المرحلة الأولى في تحديد المحتوى المعرفي الذي يريد النظام توصيله إلى المستخدم، والمرحلة الثانية هي بناء مقاطع لغوية أو فقرات صحيحة بإضافة كلمة بكلمة.(٤٦)

المبحث الثاني: نظم انتاج المعرفة (محركات المعرفة)

عملت محركات بحث الويب منذ بدايتها على تحقيق الكفاءة في استرجاع مواد المعلومات ذات الصلة بمصطلحات البحث، وتمثل ذلك في استخدامها آليات التكشيف والتحليل واستخدام الميتاداتا التي تكشف موضوع مواد الويب لإظهار أفضل درجات المطابقة والاستدعاء. غير أن استرجاع مود المعلومات يقتصر على التعامل مع مادة المعلومات بشكل عام كوحدة واحدة دون الدخول إلى محتواها المعرفي أو التدقيق في كياناتها المعلوماتية والمعارف الواردة عنها. وهنا تتمثل البداية النوعية لنظم أو محركات انتاج المعرفة التي تتنقل بالتحليل إلى مستوى المحتوى المعرفة لا مادة المعلومات، وتتحسس الكيانات المعرفية الداخلة في تفاصيل السياق وليس المصدر كوحدة عامة، بل والوصول إلى مستوى تحقيق التشابك بين هذه الكيانات المعرفية وربط معا بعلاقات موضوعية تعيد تشكيلها ورسم

توجهاتها مرة أخرى.[٤٧]كما يمكن لمحركات انتاج المعرفة تشكيل هياكل معرفية جديدة كليةً بُنيت وقت الحاجة اليها وموجهةً لخدمة الهدف المعرفي موضع اهتمام مستخدمي المحركات، بما يجعل أعمال التحليل والاستدعاء أعمالاً لحظية تتوقف على كفاءة المحرك في قدرة التحليل والربط والمعرفي وقدرة الادراك والتواصل الحواري مع مستخدميه. ولكي يتحقق ذلك في محركات انتاج المعرفة فيجب أن تتزود بقدرات التقاط الكلمات الرئيسة مع إمكانات البحث للمصادر غير المتجانسة أثناء معالجة الاستفسارات، ثم إنشاء الرسوم البيانية المعرفية عند الطلب بتطبيق نهج الربط الدلالي للجزيئات المعرفية RDF؛ حيث الرسم البياني المعرفي الناتج بدلالات الكيانات التي تم جمعها من المصادر المتكاملة، وكذلك العلاقات بين هذه الكيانات.[٤٨]

١/٢ مقومات بناء النظام

يحتاج نظام انتاج المعرفة إلى جملةٍ من مقومات العمل تتمثل في التالي:

١/١/٢ الواجهة الحوارية

تمثل الواجهة الحوارية جانب التفاعل بين النظام مستخدمه؛ حيث تتلقى هذه الواجهة مجموعة الاستفسارات ثم تمريرها إلى عملية التحليل الدلالي والتعرف على جزيئات المعرفة المطلوبة. كما أنها ترد بمجموعة المقاطع والفقرات المعرفية المنتجة من قاعدة النظام، وتمثل النتائج المعرفية النهائية الملبية لاحتياجات مستخدم النظام.[٤٩]

٢/١/٢ المحلل الدلالي

يعمل المحلل الدلالي على تحديد المردود الدلالي والمحتوى المعرفي لكل جزيء من المفردات الواردة في استفسار المستخدم، ومنها يحدد مستويات الترابط والعلاقات بين مفردات الاستفساروووضعها في تصنيف أولي يمكن توجيهه إلى قاعدة النظام لمطابقتها مع تلك المنتجة من قاعدة المعرفة.[٥٠]

٣/١/٢ قاعدة المعرفة

وهي المستودع الرئيس لجملة محتويات قاعدة المعرفة من مواد المعلومات، هي تعبر عن التفاعل الديناميكي المتجدد لجزيئات محتويات مواد المعلومات داخلها. وتمكن قاعدة المعرفة النظام من التقاط المفردات والجزيئات المعلوماتية لتشكيل الهياكل المعرفية المتخصصة في مجالات استفسار المستخدم.

٤/١/٢ الذاكرة المعرفية

تعتبر الذاكرة المعرفية من أهم مقومات عمل النظام؛ حيث تختزن هذه الذاكرة جملة الجزيئات والهياكل المعرفية التي تستمر فترة تفاعل المستخدم مع النظام ومدة الحواء المعرفي بينهما. ومن خلاله يمكن للنظام الاستعانة بتاريخ المفردات المذكورة خلال المحادثة، بل وطلب هياكل معرفية تم تشكيلها حال احتياج النظام لهذه الهياكل وإعادة تشكيلها أو اندماجها مع مكوناتٍ معرفية أخرى لتشكيل هياكل أحدث أو أكثر تخصصاً في مجالٍ محدد.(٥١)

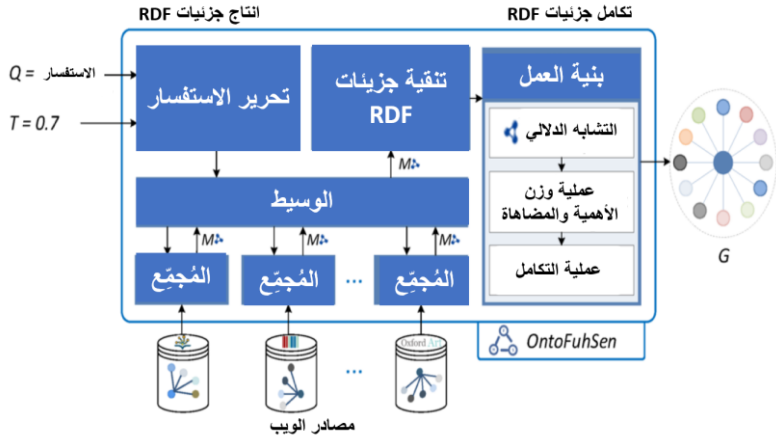
٥/١/٢ الخط المباشر مع الويب

يحتاج نظام انتاج المعرفة إلى استمرارية التحديث والتفاعل الزمني المباشر مع محتوى مواد الويب؛ حيث يمكن من خلال الاتصال المباشر بالمحتوى الرقمي إمكانية التحديث الدائم لجزيئات ومفردات قاعدة المعرفة من حيث الكم، بل و أيضا تحديثها من حيث التغييرات الدلالية التي تطرأ عليها بتوسيع أو تضييق المفاهيم الدلالية لهذه الجزيئات.

٢/٢ بنية وآليات عمل النظام

تمثل الكيانات المعرفية داخل محتوى مواد المعلومات أشكالا رسومية تصنع صورة المجالات المعرفية المختلفة، وإن العلاقات بينها هي تلك الألوان التي تحدد الشكل النهائي لهذه الصورة الكاملة. ومن ثم، فإن تحرير هذه الكيانات وإعادة تشكيل علاقاتها يمثل قدرات انتاج المعرفة من المحتوى النصي لمواد المعلومات فيما يمكن أن يُطلق عليه التنقيب عن المعرفة **text mining**.(٥٢) ويتجلى عامل التكامل بين نصوص قاعدة المعرفة في الدور الذي يلعبه بتضخيم أعداد النتائج المعرفية المُستدعاة من العدد المحدود من مواد المعلومات المصدرية لهذه القاعدة؛ حيث تتضاعف أعداد الهياكل المعرفية الناتجة عن تحرير ورسم علاقات الكيانات المعلوماتية داخل قاعدة المعرفة، فخاصية الإبدال لعلم الرياضيات تؤكد على الاعداد الهائلة من الكيانات المعرفية المستحدثة نتيجة لتغيير علاقات الربط بين الكيانات المعلومات في مواد المعلومات داخل قاعدة المعرفة. وإن مصطلح المعرفة ذاته يحمل مفهوم أعلى في تلبية الاحتياجات المعلوماتية لمستخدمي النظام؛ حيث تتعدى النتائج المسترجعة استدعاء المعلومات إلى تقديم الآليات والمهارات لمستخدمي النظام. وتهتم آليات تشكيل الهياكل المعرفية هنا داخل النظام برصد الكيانات المعرفية والدوائر المعلوماتية المتعلقة بها؛ مثل رصد النتيجة وأسباب حدوثها، والظاهرة وأعراضها، أدوات عمل الأشياء المختلفة، ثم الربط بين ذل بحيث يتم صناعة هيكل معرفي جديد داخل النظام عن كل ما يمكن أن ينتج عن ظاهرة محددة اذا توفرت الأسباب التي تم مناقشتها في مادة معلومات مستقلة،

وكيف يمكن مواجهتها او الاستفادة منها بأدوات العمل التي نصت عليها مواد معلوماتية جانب مستقل، وعلى ذلك تتضح ملامح الهياكل المعرفية عند الاستدعاء لمستخدمي النظام في الاستفادة من جملة هذا التفاعل بين المعارف المختلفة في قاعدة النظام.(٥٣)



شكل رقم (٧) بنية وآليات عمل نظام انتاج المعرفة(٥٤)

يعبر شكل (٧) بنية نظام انتاج المعرفة؛ حيث يبنى عمل النظام على آليتين هما:

١/٢/٢ الآلية الأولى: تحديد الكيانات الدلالية semantic entities identification

وبها يتم استقبال مجموعات بيانات RDF، ومنها يتم جمع الجزيئات ذات العلاقات المتجانسة وربطها بمنطق التشابه الدلالي والمعرفي، وذلك اعتمادا على ظهور العلاقات الدلالية وقابلية الاندماج بين هذه الكيانات. وعليه يكون المنتج النهائي هيكل معرفي في شكل رسم بياني RDF متكامل ترتبط فيه مجموعات المعلومات ارتباطاً ثنائي الأجزاء RDF.(٥٥)

٢/٢/٢ الآلية الثانية: تكامل الجزيئات المعرفية Molecules Integration

تعمد آلية RDF Molecules Integration إلى سياسات الاندماج لدمج جزيئات RDF المكافئة لغويًا و إنشاء رسم بياني RDF متكامل مبنياً على اللغة الدلالية لتعيين الخصائص والموارد في جزيئات RDF المكافئة، بينما تحدد سياسات الاندماج قواعد معينة لكيفية دمج الخصائص أو القيم المعينة فعليًا من أجل القضاء على التكرار مع الحفاظ على الاتساق، ويكون الاندماج الدلالي قادرًا على إنتاج كيان مدمج يهدف إلى كيانات معرفية كاملة ومتسقة.(٥٦) ويراعي مكون الدمج تسلسل الدمج بين الجزيئات الأكثر تخصصا في المجال إلى المجموعات الأعم؛ حيث يتم الاحتفاظ بالتسلسل الهرمي للمجالات المعرفية التي تمثلها الهياكل المولدة النهائية. ويتطلب الكم الهائل من البيانات التي تتم مشاركتها على الويب

تقنيات دقيقة وفعالة لاسترجاع و انتاج المعرفة التي يمكن أن تستخدمها أدوات الويب، ويعد إنشاء معرفةً تركيبية من الويب مهمة صعبة؛ خاصة حول نفس الكيان عبر مصادر بيانات الويب المختلفة المشتتة في فضاء الويب. وحتى يتم ذلك، يجب حل العديد من مشكلات إمكانية التشغيل البيني للبيانات، مثل بناء مخططات وهياكل المعرفة أو العلاقات غير الصحيحة بين التخصصات في المجال الواحد أثناء عملية التكامل.(٥٧)

وقد تطورت تقنية الويب الدلالي وأدوات معالجة اللغة الطبيعية بما يكفي لحل مشكلة تكامل المعرفة للبيانات غير المتجانسة. إن استرجاع ودمج المعرفة من مصادر بيانات الويب عملية معقدة بشكلٍ أساس بسبب طرق الاستخراج والتحويل والتحميل في كل مكان في هذه العملية وبناء المخططات ورسم العلاقات، أيضا فإن ثمة حالات يُحظر فيها التكامل المادي الكامل للبيانات (على سبيل المثال بسبب إخفاء البيانات خلف واجهات برمجة التطبيقات) أو غير مسموح به (على سبيل المثال بسبب مخاوف تتعلق بخصوصية البيانات). لذلك، هناك حاجة إلى نهج متكامل موحد و أكثر فعالية يدعم المنظمات في توليد قوى معرفية من بيانات ومعلومات متجانسة وغير متجانسة.(٥٨)

٣/٢ مكونات نظام انتاج المعرفة

يتطلب أداء عمل نظام انتاج المعرفة تو افر عدة مكونات هي كالتالي:

١/٣/٢ مكون الاستعلام Query Rewriting

يحول هذا المكون بشكلٍ أساس استعلام الأولي من مستخدمي النظام بما فيه من كلمات إلى استعلامات يفهمها المُعرِّب wrapper، وذلك باستخدام وصف مصدر البيانات في مفردات Onto، وعليه يتم تحويل الاستعلام الأولي إلى استعلام مكون من جزيئات معرفية يتم إرسالها إلى مكون محرك البحث.(٥٩)

٢/٣/٢ مكون المُعرِّب الوسيط mediator wrapper

تعمل مكونات المعرب الوسيطة بتنسيق عملية استخراج جزيئات المعلومات باستخدام أطر وصف البيانات RDF وتخزينها في شكل بياني يحفظ في ذاكرة النظام. وعليه يبدأ المعرب في تجميع الجزيئات المتو افقة وهذا الشكل البياني من قاعدة مصادر ومواد المعلومات، حتى يتم انتاج المعارف المتو افقة وهذه الدلالات المستقبلة من مستخدم النظام. ويمكن وصف عمل مكون المعرب الوسيط بأنه حلقة الوصل بين دلالات الاستعلام ودلالات تجميع وتشكيل المعرفة المنتجة من مواد المعلومات في النظام.(٦٠)

٣/٣/٢ مكون مُحَسّن الجزيئات الدلالية RDF Molecules Enrichment RDF

يعمل هذا المكون على تحسين جودة جزيئات RDF المشتقة من استعلام المستخدم ومن قاعدة النظام، وذلك من خلال إثرائها بالحقائق الجديدة المكتسبة الواردة من مواد المعلومات المستحدثة، ويتم إرفاق المعلومات الدلالية المحسنة إلى قاعدة المعرفة لزيادة فاعليتها وتحسينها المستمر. ويستمد هذا المكون قدرته على تحسين جزئيات المعرفة داخل قاعدة المعرفة من خلال التنقيب عن الجزيئات المعرفية بناءً على ذات الجزيئات الموجودة في الهياكل المعرفية، وذلك اعتمادا على خوارزميات تحليل الرسم البياني المعرفي لكل لمختلف المجالات الموضوعية.[٦١] وثمة عاملان أساسيان بناء عليهما يمكن قياس مدى درجة الجودة المضافة لجزيئات المعرفة المجمعة هما؛ مواد المعلومات المصدرية في قاعدة المعرفة، ومدى تكاملها في تحقيق المعارف المسترجعة، أيضا، حداثة مواد المعلومات وذلك نظرا للديناميكية والتحديث المستمر لمواد المعلومات التي تتصل مباشرة بقاعدة المعرفة في النظام.[٦٢]

٤/٣/٢ مكون المُجمّع Integrator

يستقبل مكون المجمع مجموعات الجزيئات التي حددها المعرب الوسيط وذلك حتى يمكن للمجمع إنشاء رسومًا بيانية لهيكل المعرفة المتوقع الرد به والمتو افق مع هذه الجزيئات و اتجاهاتها الموضوعية، وهو ما يجعل تجميع البيانات والمفردات المعرفية مهمة بسيطة نسبيًا. ويضم مكون المجمع بداخله مجموعة من المكونات الفرعية التي تدعم استدعاء المعارف المتو افقة وهيكل المعرفة المبني داخله. وهذه المكونات الفرعية تتمثل في التالي:[٦٣]

١/٤/٣/٢ المكون الفرعي للتشابه الدلالي

يستخدم هذا المكون آلياتٍ تتمثل في التعريف بدلالة كل جزيء من جزيئات المعرفة الواردة في استعلام المستخدم وتلك التي تم اشتقاقها من قاعدة المعرفة؛ حيث يتم جمع الجزيئات المتجانسة موضوعيا ودلاليا لخدمة الهيكل المعرفي العامل للمجال.

٢/٤/٣/٢ المكون الفرعي للتو افق

يعمل هذا المكون على تحديد مستويات المطابقة والتو افق بين الجزيئات المعرفية في استعلام المستخدم وبين الجزيئات المعرفية الواردة من قاعدة المعرفة بحيث يصل بها إلى أعلى مستويات التطابق والتو افق في المجال المعرفي المستهدف. وكلما ازداد هذا التو افق يقترب النص المعرفي والفقرات المستدعاة إلى أفضل رد على المستخدم في واجهة النظام.[٦٤]

٣/٤/٣/٢ المكون الفرعي للتكامل.

تتحد الجزيئات التي تم تكوينها والمتو افقة دلاليا مع استعلام المستخدم بحيث يتم ترتيبها في هيكل معرفي متكامل يراعي تداعي عرض الموضوع بشكلٍ يعكس هوية هذا الموضوع.

المبحث الثالث: محركات بحث واستدعاء المعرفة (الو اقع التطبيقي)

تتجه الدراسة التطبيقية إلى قياس و اقع أدوات معالجة واسترجاع المعلومات الرقمية على الويب، ومدى ملاءمتها لآليات وقدرات انتاج المعرفة واستعدادها لذلك. فضلا عن تحديد متطلبات ومقومات تحول محركات البحث الحالية إلى نظم انتاج المعرفة بمفهومها المقترح في هذه الدراسة. وقد اختارت الدراسة عينةً غرضية من محركات البحث الدلالية التي تمثل المستوى الأقرب من معالجة المعلومات بمقتضيات انتاج المعرفة، فضلا عن التركيز على محركين أساسين هما Bing and Google اللذان يمثلان نموذج أولي لنظم انتاج المعرفة، وفيهما الكثير من مقومات وآليات عمل انتاج المعرفة. وسيتم اجراء الدراسة التطبيقية بمقارنة خصائص العمل لإنتاج المعرفة على مختلف محركات عينة الدراسة، كما يلي:

أولا: أسلوب التفاعل

تختلف نظام انتاج المعرفة عن محركات البحث الأخرى في أسلوب التفاعل مع مستخدميها؛ حيث تتجه نظم انتاج المعرفة إلى استخدام أسلوب الحوار المعتمد على تقنية CHAT GPT، في حين تستخدم محركات البحث الدلالية الأخرى الأسلوب المعتاد في تلقي الاستفسارات واستعراض النتائج المسترجعة.(٦٥)

جدول (١) أسلوب تفاعل عينة الدراسة مع مستخدمي النظام

أسلوب التفاعل مع مستخدمي النظام	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
الحواروالمحادثات	---	√	---	---	---	---
تلقي الاستفسارات	√	√	√	√	√	√
كِلا الاسلوبين	---	√	---	---	---	---

يتضح من جدول (١) مجموعة من الملاحظات تتمثل في الآتي:

- اعتمدت واجهة محركات بحث عينة الدراسة جميعا أسلوب الاستفسارات المتضمنة للكلمات المفتاحية المعبرة عن احتياجات المستخدم من المعلومات، دون استخدام أسلوب الحوار وذلك ما يعكس افتقاد عينة الدراسة لوجود نظام إدارة الحوار والمحادثات الاسترجاعية أهم مكونات انتاج المعرفة.

- يعد نظام Bing النظام الوحيد لإنتاج المعرفة بين عينة الدراسة؛ ذلك لما له من قدرة على تلقي استفسارات البحث في شكل حواروالرد عليه بمقاطع معرفية تمثل الطرف الثاني في التفاعل المعرفي وتكوين المحادثات الاسترجاعية.

ثانيا: النتائج المعلوماتية والمعرفية

تعكس النتائج المستدعاة على واجهة النظام طبيعة المعالجة المعلوماتية التي يتبناها النظام مع مواد المعلومات ومحتوياتها؛ حيث تعبر الفقرات والمقاطع المستدعاة للإجابة على استفسارات مستخدم النظام عن تجزئة وتكامل المعارف، أما النتائج المسترجعة بشكل مستقل فهي تعبر عن معالجة المواد دون محتواها ويكون عمل النظام معتمدا على تكيف المادة واستدعائها وفقا للمفردات المعبرة عنها دون التعمق في تحليل محتواها المعرفي.[٦٦]

جدول (٢) النتائج المعلوماتية والمعرفية لعينة الدراسة

النتائج المعلوماتية والمعرفية	نظام انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
استدعاء المواد	√	√	---	---	---	---
استدعاء المعرفة	√	√	√	√	√	√
استدعاء النوعين	---	√	---	---	---	---

يتضح من جدول (٢) الملاحظات التالية:

- انتهجت عينة الدراسة أسلوبا عاما في استدعاء المحتوى المعلوماتي للرد على مستخدميها، يتمثل في استدعاء الملفات والمو اقع والمصادر التي تتصل وموضوع البحث. غير أن نظامي Bing and google فقط استخدما أسلوبا آخر يتمثل في الردود المعرفية المباشرة التي تكفي المستخدم الانتقال الى مواد المعلومات ذاتها لاستعراض تفاصيل المعلومات.
- استخدم نظام Bing أسلوبا متقدما يتمثل في الردود المفتوحة لاستقبال التعليق عليها، ثم معاودة الرد مرة أخرى فيما يعبر عنه بالحوار أو المحادث المفتوحة. ويعبر ذلك عن مفهوم البناء المعرفي للردود والاجابات المعلوماتية التي تتعدى مستوى مواد المعلومات المنفردة لتصل إلى المحتوى المعرفي المجمع وفقا لهيكل وبنية دلالية موضوعية محددة.

ثالثا: الواجهات الحوارية ونظم توليد المحادثات

أوجدت شركة ميكروسوفت منتجا جديدا من خلال رعايتها لمجموعة الذكاء الاصطناعي مفتوح المصدر open AI الذي تمخض عن دراساته وبرمجياته تقنية مولدات المحادثات الاسترجاعية CHAT GPT، وقد استثمرت ميكروسوفت هذا المنتج في إعادة أحياء محرك البحث BING بميزة تنافسية استطاعت النيل من صدارة محرك البحث GOOGLE كأفضل محركات بحث الويب استخداما.[٦٧]ولا تمصل الواجهات الحوارية مجرد تقنية حديثة صالحة لاسترجاع المعلومات فحسب، بل إنها تمثل نصف قدرة العمل التي يحتاجها محرك البحث للتحول إلى نظام لإنتاج المعرفة، بما لها من قدرات تفكيك مفردات الحوار مع مستخدم النظام وإعادة بنائها مرةً أخرى بناء على علاقاتها الموضوعية والدلالية.

جدول (٣) الواجهات الحوارية ونظم توليد المحادثات لنظم عينة الدراسة

الواجهات الحوارية ونظم توليد المحادثات	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
واجهة حوار	---	√	---	---	---	---
فراغ البحث	√	√	√	√	√	√
واجهة تفاعلية	√	√	√	√	√	√

يتضح من جدول (٣) المشاهدات التالية:-

- تفرَّد نظام Bing من بين محركات العينة بواجهة حوارية صريحة لإجراء المحادثات الاسترجاعية بينه وبين مستخدميه، يمكن من خلالها مناقشة جوانب وتفاصيل المجال المعرفي والاحتياجات المعرفية للمستخدم. وقد عرض النظام لجوانب تساعد المستخدم على الاستمرار في الحوار وبنائه بمنطق تدقيق المعلومات المسترجعة من خلال اقتراح أسئلة مخصصة يمكن للمستخدم الاستعانة بها في اكمال الحوار، أيضا البدايات محددة الوجهة في الرد بحيث إذا كان لدى المستخدم المزيد من التخصيص يمكن مناقشته في المقاطع الحوارية اللاحقة.
- توقفت باقي نظم عينة الدراسة وعلى رأسها نظام google عند المستوى الدلالي الذي يفتح أمام المستخدم واجهات توضيحية لشرح المفاهيم الأساسية في الموضوع مستخدما جزء من محتوى معرفي محدد، ثم استكمال سياق النتائج بمجموعات النتائج المسترجعة.

ويدلل ذلك على افتقار نظم عينة الدراسة لقدرات فتح الحوار وتبادل جو انب التفاعل بينها وبين المستخدم.

- تعكس النتائج المسترجعة و اقع عينة الدراسة التي لم تستطع الخروج من نطاق التعامل مع مواد الويب باعتبارها مصادر معلومات مستقلة، وعدم قدرتها على الدخول إلى عمق المحتوى وتحليله بتجزئته لمقاطع معرفية.

رابعا: أسلوب المعالجة والتحليل الدلالي

تتبنى نظم انتاج المعرفة أعلى مستويات المعالجة للغة الطبيعية؛ حيث تحتاج إلى إتمام المعالجة المعرفية المفاهيمية لجزيئات المعلومات داخل قاعدة المعرفة، وبدون هذا المستوى لا يمكن للنظام التمكن من تحليل الدلالات والمفاهيم لمفردات الاستفسارات ومحتويات مواد المعلومات.[٦٨] وهو ما يجعل مستوى المعالجة مؤشراً على قدرة النظام في الوصول إلى درجات تخصيب المفردات الدلالية و إنتاج المعرفة.

جدول (٤) أسلوب المعالجة والتحليل الدلالي في عينة الدراسة

أسلوب المعالجة والتحليل الدلالي	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
المعالجة الطبيعية	√	√	√	√	√	√
المعالجة الدلالية	√	√	√	√	√	√
المعالجة المعرفية	√	√	---	---	---	---

يعكس جدول (٤) بعض الجو انب مثل:

- ارتقى نظام Bing إلى مستوى المعالجة المعرفية التي تتعامل مع جزيئات محتويات مواد المعلومات، من حيث تخصيصها الموضوعي والدلالي ثم إعادة جمعها وتمثيلها في هياكل معرفية جديدة، يمثل كل منها بناء موضوعي مختلف أو ذات وجهة محددة. ويعتبر هذا المستوى من المعالجة المعرفية الداعم لنظم انتاج المعرفة ودعم واجهات المحادثات الاسترجاعية لها.
- تمتلك باقي عينة الدراسة مستويين مختلفين من المعالجة الموضوعية للغة الطبيعية يمثل المستوى الأدنى فيها المعالج الطبيعية للمفردات تحليلها الموضوعي، ثم المستوى الأعلى في معالجة وتشكيل دلالات ومفاهيم المفردات في مواد المعلومات وقدرة إدراك مفهوم السياق التي وردت فيه. غير أن كلا المستويين لا يدعم بناء نظام انتاج المعرفة الذي يحتاج إلى تجزئة وتكوين المفردات المعرفية في هياكل موضوعية دلالية جديدة.

خامساً: مدى توافر مكونات معالجة المعرفة

يبنى نظام انتاج المعرفة بتكامل مجموعة من المكونات يمكن لغياب أيا منها التأثير سلباً على قدرة النظام في تحليل وإنتاج الهياكل المعرفية. وتتباين هذه المكونات من حيث أهميتها داخل النظام، غير أنها جميعا تقع في مستوى الضرورة لتحقيق أهداف انتاج المعرفة.(٦٩)

جدول (٥) توافر مكونات معالجة المعرفة في عينة الدراسة

مكونات معالجة المعرفة	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
مكون الاستعلام	√	√	√	√	√	√
مكون المعرب الوسيط	---	√	---	---	---	---
مكون محسن الجزئيات الدلالية	√	√	---	---	√	√
مكون المُجمِّع	---	√	---	---	---	---
المكون الفرعي للتشابه الدلالي	---	√	---	---	---	---
المكون الفرعي للتوافق	---	√	---	---	---	---
المكون الفرعي للتكامل	---	√	---	---	---	---

يتضح من جدول (٥) بعض الدلائل التالية:

- لم يتضح من إجراءات البحث واسترجاع النتائج في عينة الدراسة اعتمادها على مكونات أو برمجيات تمكنها من انتاج المعرفة، وهي المكونات التي يتصل عملها بتحليل وتشكيل هياكل المعرفة مثل مكونات المجمع والوسيط والمكون الفرعي للتكامل بين جزيئات المعرفة. وذلك فيما عدا نظام Bing، ويؤكد ذلك على توقف عمل عينة الدراسة عند مستوى نظم استرجاع المعلومات وليس نظم انتاج المعرفة.
- إن النتائج المسترجعة في هذا الجانب تؤكد على تكامل المكونات الفرعية لنظم انتاج المعرفة؛ حيث لا يمكن تحول النظام من استرجاع المعلومات لإنتاج وإدارة المعرفة بدون توافر كل مكونات العمل الدلالية التي تختص بتفتيت جزيئات المحتوى وإعادة تشكيلها، وبناء المقاطع والفقرات المعرفية.
- يمثل مستوى تحليل المفاهيم والدلالات والكيانات المعلوماتية وتكامل الهياكل المعرفية فارق العمل بين نظم استرجاع المعلومات التي تعمل مع مادة المعلومات كوحدة واحدة، وبين

نظم انتاج المعرفة التي تصل الى معالجة محتوى المعرفة في داخل مادة المعلومات، بل وتطويعه وفقا للاحتياجات الموضوعية المختلفة.

سادسا: القاعدة المعرفية لمحركات البحث

تمثل قاعدة المعرفة المرتكز الرئيس لنظام انتاج المعرفة؛ حيث يستطيع النظام من خلالها تخزين المفردات الدلالية المؤرخة للتفاعل مع المستخدم، وأيضا استدعاء مختلف الجزيئات والمفردات الدلالية والمفاهيمية التي تتشكل بها هياكل المعرفة وتتكون بها الصورة النهائية للمقاطع والفقرات والمستخلصات المعرفية المستدعاة لتلبية احتياجات المستخدم.[٧٠] ويمكن قياس وجود وفاعلية قاعدة المعرفة داخل نظم عينة الدراسة من خلال تتبع الإجابات والنتائج المستدعاة على واجهة الحوار؛ حيث تختلف النتائج المباشرة المتكاملة عن النتائج المشتتة المتفرقة التي لا تعكس وجود قاعدة معرفية من الأساس.

جدول (٦) القاعدة المعرفية لمحركات بحث عينة الدراسة

القاعدة المعرفية لمحركات البحث	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
معارف متكاملة	---	√	---	---	---	---
نتائج موجهة	√	√	---	---	---	√
مصادر متفرقة	√	√	√	√	√	√

يتضح من جدول (٦) الآتي:

- أن مفهوم القاعدة المعرفية لم يتو افر بنيةً وتطبيقا الا في نظام Bing؛ حيث توصف قاعدة بيانات النظام بأنها معرفية حال ما استطاعت دعم النظام بمفردات ومحتويات معلوماتية متكاملة وليست وحدات او مصادر معلومات مشتتة. أي أن نظم استرجاع المعلومات التي تمثلها عينة الدراسة تعمل على تخزين واستدعاء مواد ومصادر معلوماتية، بينما تحتوي قاعدة المعرفة على كيانات وهياكل معرفية ديناميكية.
- يوصف النظام بأنه نظام انتاج معرفة حال انعكس ذلك على النتائج المستدعاة؛ حيث يمكن الوصول إلى استدعاء المعارف المتكاملة والمقاطع والفقرات المباشرة في إجابة استفسار البحث مع وجود قاعدة المعرفة التي تكون وتنشئ النصوص في الحوار والمحادثة مع المستخدم. وتعمل نظم استرجاع المعلومات على تحسين إجراءات استرجاع المعلومات من خلال تخطي حاجة النتائج المتفرقة والروابط الخارجية إلى مستوى النتائج الموجهة في اقتطاع أجزاء من سياقات مواد المعلومات للرد المباشر على استفسار البحث.

سابعا: خصائص تلقي الاستفسارات

تتسم نظم انتاج المعرفة باستخدام لغة الحوار عن غيرها من نظم استرجاع المعلومات ومحركات البحث، حيث تتميز عملية انتاج المعرفة بأنها عملية تفاعلية قائمة على التواصل المباشر مع مستخدم النظام والتعمق في تحليل وتحديد احتياجاته المعرفية بناء على استمرارية الحديث والحوار والتماس جوهر الحاجات المعرفية والرد عليها بما يو افقها من قاعدة المعرفة. وهو النهج الذي لا تتبعه مطلقا نظم استرجاع المعلومات؛ حيث تتلقى الاستفسارات في شكل كلمات مفتاحية أو أسئلة موجهة لاسترجاع معلومات.(٧١)

جدول (٧) خصائص تلقي الاستفسارات في عينة الدراسة

خصائص تلقي الاستفسارات	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
الحوار	---	√	---	---	---	---
الأسئلة الموجهة	√	√	√	√	√	√
الكلمات المفتاحية	√	√	√	√	√	√

يتضح من جدول (٧) أن أسلوب استقبال نظم عينة الدراسة يعكس دلالةً قاطعة على طبيعة عمل النظام الداخلية ومدى قدرة النظام على استرجاع المعلومات أو انتاج المعارف؛ حيث قدم نظام Bing فقط أسلوب التحاور مع مستخدميه استقبالا للاحتياجات المعرفية والرد عليها بما يو افقها من إجابات في شكل استكمال الحديث. وهو يما يعكس امتلاك النظام لقاعدة معرفة ونظام لإنتاج وإدارة المحادثات الاسترجاعية، خلاف باقي عينة الدراسة التي توقفت عند حد استقبال المفردات المفتاحية لصياغة الاستفسار والرد على ذلك باسترجاع المعلومات في شكل مفردات أو رو ابط خارجية أو نتائج متفرقة يجمعها فقط القرب أو البعد عن الموضوع.

ثامنا: خصائص عرض النتائج المعرفية

تتميز النتائج المعرفية عن النتائج المعلوماتية في خصائص التكامل والمباشرة والايجاز، خلاف ما تكون عليه النتائج المعلوماتية المسترجعة واجهات عرض النتائج لمحركات البحث من تشتت وتباين في تلبية احتياجات مستخدم النظام. ويمثل تزود النتائج المسترجعة بالصفات المعرفية دليلا على قدرة النظام على انتاج وحدات المعرفة وإمكانية التحول من استرجاع المواد والمصادر إلى انتاج المعارف.(٧٢)

جدول (٨) خصائص عرض النتائج لمحركات بحث عينة الدراسة

خصائص عرض النتائج المعرفية أو المعلوماتية	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
استدعاء المقاطع	---	√	---	---	---	---
استدعاء النتائج	√	√	√	√	√	√
استدعاء الروابط	√	√	√	√	√	√
استدعاء الملفات	√	√	√	√	√	√

يتضح من جدول (٨) أن شكل النتائج المستدعاة في النظام يعبر بوضوح عن هوية النظام كونه نظام قادرا على انتاج وتشكيل المقاطع والفقرات المعرفية، وهو ما يجعله نظاما لإنتاج المعرفة بوضوح. أما وقوف النظام عند استدعاء نتائج وروابط وملفات يجعل النظام متوقفا عند حد استرجاع مواد المعلومات. وقد تحققت هوية انتاج المعرفة فقط في نظام Bing.

تاسعا: مقومات العمل لنظم انتاج المعرفة

تحتاج نظم انتاج المعرفة الة مجموعة من مقومات العمل إلى جانب مكوناتها الرئيسة لبناء النظام والتمكن من معالجة محتوى المعلومات الرقمية وتشكيل الهياكل المعرفية واجراء التفاعل والحوار المعرفي مع مستخدم النظام. ويمكن من خلال اختبار نظم ومحركات عينة الدراسة يمكن استجلاء وجود هذه المقومات أم لا بتتبع وتطبيق الاستفسارات استرجاع النتائج على واجهاتها.[٧٣]

جدول (٩) توافر مقومات انتاج المعرفة لعينة الدراسة

مقومات العمل لنظم انتاج المعرفة	نظم انتاج المعرفة					
	GOOGLE	BING	HAKIA	KNGINE	SWOOGLE	DuckDuckGo
الواجهة الحوارية	---	√	---	---	---	---
المحلل الدلالي	√	√	---	---	√	√
قاعدة المعرفة	---	√	---	---	---	---
الذاكرة المعرفية	---	√	---	---	---	---
الخط المباشر	√	√	√	√	√	√

يتضح من جدول (٩) بعض الملاحظات التالية:

- يستلزم لتحول النظام إلى قدرات انتاج المعرفة تو افر مجموع من المقومات يتكامل عملها مع عمل المكونات الفرعية للنظام لدعم تحليل و انشاء الهياكل المعرفية واسترجاع الفقرات والمقاطع التي تحمل الإجابات المباشرة.
- تو افرت مختلف مقومات انتاج المعرفة لنظم معالجة معلومات الويب في نظام Bing؛ حيث يستطيع النظام اجراء التفاعل بالحوار والمحادثة من خلال تخصيص واجهة حوارية مع مستخدميه، كما يمكن له تكامل المعرفة المنتجة طوال فترة التفاعل مع المستخدم من خلال الذاكرة المعرفية التي تؤرخ للمفردات الدلالية السابقة واللاحقة في حوار المستخدم.
- إن بعضا من مقومات العمل مثل الخط المباشر والذاكرة المعرفية يتعلق وجودها بتحسين نتائج المعرفة المستدعاة وليس بقدرة النظام على انتاج المعرفة بالأساس؛ حيث يدعم مقوم الخط المباشر قدرات النظام على التحسين الدائم لطبيعة المفردات الدلالية والمفاهيمية ومدى حداثتها ومواكبتها لتطورات المعارف الآنية. أيضا فإن الذاكرة المعرفية تؤثر في تحسين النتائج المستدعاة بربط المفردات والدلالات الحالية بالدلالات والمفاهيم السابقة في بداية حوار المستخدم، مما يدعم تكامل النتائج أولا وأخيرا وترابطها حول موضوع المحادثة.

نتائج وتوصيات الدراسة:

خرجت الدراسة من جملة ما تم عرضه في الاطارين النظري والتطبيقي بالنتائج والتوصيات التالية:

أولا: وضعت نظم انتاج المعرفة مؤسسات المعلومات أمام تحدٍّ نوعي كون هذه النظم تتعدى حاجز مادة المعلومات إلى إدارة المعرفة وتلبية احتياجات مستخدميها في شكل معارف ومحتوى مباشر يتوافق والاحتياجات المعلوماتية التي يرغبونها. وعليه فإن وقوف مؤسسات المعلومات عند مستوى إدارة مواد المعلومات ومعالجتها يعرضها للانعزال التام عن واقع معالجة وانتاج المعرفة. وعليه، توصي الدراسة بإحداث تغيير نوعي على المستويين الأكاديمي والمهني في أقسام ومؤسسات المعلومات تتماشى وطبيعة التوجهات الرقمية في معالجة وإنتاج المعرفة، سواء أكان ذلك في طبيعة الخطط الدراسية في الأقسام الاكاديمية أو في نوعية البرمجيات والتجهيزات اللازمة لإدارة مؤسسات المعلومات.

ثانيا: تواجه نظم استرجاع المعلومات على الويب وخاصة محركات البحث عاصفة رقمية تكاد تطيش معها القدرات والآليات التي تستخدمها محركات البحث في معالجة واسترجاع

مواد الويب، وذلك أن آليات انتاج المعرفة تتقدم بخطوة هائلة في قدرات تلبية احتياجات مستخدمي الويب؛ فبدلا من النتائج المستقلة المتفرقة البعيدة القريبة من الموضوع، تقدم نظم انتاج المعرفة رداً موجزاً في شكل مقاطع أوفقرات مباشرة شاملة كل الجو انب المعرفية لمجال الاستفسار، دون التنقل أو التشتت بين مواد ومو اقع الويب. وعليه، فإن الدراسة توصي بتوجيه الانتباه إلى الدراسات التطبيقية لمحركات البحث وتوصيفها في إطارنظم انتاج المعرفة وتحديد متطلبات التحول الرقمي والمعلوماتي في أداء محركات البحث.

ثالثا: لا تزال المعالجة المعرفية للمحتوى الرقمي معتمدةً على تطور آليات وتقنيات اللغة الطبيعية؛ حيث لوحظ الارتقاء بمستوى معالجة المفردات في مجتوى مواد المعلومات من اللفظي الى الدلالي إلى التحليل المعرفي وتكوين شبكات وهياكل معرفية لمجالات المعلومات. وعليه، فإن الدراسة توصي بتبني التوجهات الآنية للمعالجة المعرفية من علم المعلومات بمزيد من الدراسات التي تؤصل لقدرات الهياكل المعرفية والتصنيف المعرفي لجزيئات المعلومات وتحويل قواعد البيانات إلى مفهوم قواعد المعرفة.

رابعا: إن ثمة فارق نوعي بين محركي البحث Bing and Google كأول نظامين يحققان مفهوم انتاج المعرفة وبين باقي محركات البحث الدلالية التي تعمل على التحليل الدلالي لمفردات مواد الويب في مرحلة التكشيف أو البحث والاسترجاع. وقد أوضحت الدراسة التطبيقية أن النتائج المستدعاة في نظم انتاج المعرفة تختلف نوعيا عن تلك التي تعمل بها محركات البحث الدلالية، التي تقترب في الخصائص مع محركات البحث التقليدية من حيث استدعاء نتاج مشتتة متفرقة غير متكاملة الجو انب للمجال الموضوعي الواحد. وعليه، توصي الدراسة ببداية التحول النوعي لمختلف محركات البحث بمفهومها الحالي إلى الو اقع النوعي لعمل محركي البحث Bing and Google.

خامسا: مازالت قدرات انتاج المعرفة وممارساتها في طور العمل التجريبي والتحديث المستمر من جانب شركات عالمية مثل ميكروسوفت وجوجل، باعتبارأن انتاج المعرفة يمثل مضمارا للسباق والتنافس التجاري بين الشركتين تتحقق من خلاله قوة وضعف النفوذ في مجال استرجاع المعلومات وإنتاج المعرفة. وعليه، توصي الدراسة بتوجيه شركات التقنيات والمعالجة الرقمية العربية جزءا من قدراتها للمشاركة والمنافسة فيما يخص انتاج المعرفة من المحتوى الرقمي العربي على الويب، ومن ثم تحصيل النفع المادي والقومي للمحتوى الرقمي العربي؛ حيث لن يكون ذلك في نطاق اهتمام الشركات الأجنبية.

سادسا: أوضحت نتائج الدراسة التطبيقية أن انتاج المعرفة مازال قاصرا على معالجة المحتوى النصي فقط؛ حيث تكون الاستعلامات والمقاطع والمعارف المسترجعة في شكل السياق النصي فقط. وقد خلت عينة الدراسة التطبيقية من أدوات بحث المعلومات والمعرفة من استخدام وسائط أخرى بجانب النص في إدارة الحوار مع مستخدم النظام. وترى الدراسة أن استخدام وسيطي الصوت والصورة ليسا ببعيدين عن التفعيل في الوقت القريب، وسيكون الوسيط الصوتي الأول في تحقيق التفاعل واسترجاع المعرفة؛ حيث تمتلك الآلات الآن قدرات إدراك المفردات الصوتية وتحليلها دلالياً وهو ما يمكن أن يتم ترجمته إلى الوسيط النصي لقاعدة المعرفة النصية أو الصوتية فيما بعد.

سابعا: أثبتت الدراسة التطبيقية أن نظم توليد المحادثات الاسترجاعية وواجهات الحوار مع المستفيد الداعم الأول لبناء نظم انتاج المعرفة بمفهومها ومؤداها الصحيح؛ حيث لا يمكن للنظام إيجاد التفاعل المعرفي في تلقي المفردات واستدعاء المقاطع المعرفية وتشكيلها إلا على واجهة توليد المحادثات الاسترجاعية. وعليه، فإن نظام توليد المحادثات وقاعدة المعرفة بآليات تشكيل الهياكل المختلفة، يمثلان المكونان الاساسيان لبناء نظام انتاج المعرفة، ويأتي معهما باقي المكونات التي تكمل بنية النظام.

ثامنا: أثبتت الدراسة التطبيقية على عينة الدراسة من محركات البحث الدلالية أن قدراتها على التحول لنظم انتاج المعرفة لا تكون كاملة الا إذا توفرت بها مختلف المقومات والآليات ودعائم العمل، مثل الواجهة الحوارية وقاعدة المعرفة وبناء هياكل المعرفة ومكونات المعرب والوسيط لتكوين جزيئات المعرفة؛ حيث يمثل فقدان أي من هذه الجوانب بعدا تاما عن مفهوم ومؤدى نظم انتاج المعرفة. وهو ما بدا جليا في اجتماع مختلف الجوانب في محرك البحث BING وشبه اكتمالها في محرك البحث GOOGLE. وقد انتفت جل هذه المقومات في باقي محركات عينة الدراسة.

تاسعا: أظهرت النتائج والاجابات التي خرجت مع الدراسة التطبيقية لعينة الدراسة أن انتاج المعرفة في أول مراحله كتغير نوعي لنظم استرجاع المعلومات؛ فالتعمق في المقاطع والاجابات الحوارية المعرفية المستدعاة من النظام تؤكد جودة التحليل الدلالي والربط المفاهيمي بين مفردات الاستفسارات، غير أنها لا تلبي احتياجات مستخدمي النظام في الإجابات المعرفية التي تتضمن مهارات عمل أو إرشادات لأداء مهام أو رداً مباشرا على حل مشكلة منطقية أو مضامين إنسانية واجتماعية. فهي تمثل مستوى أعلى من انتاج الأداء تسعى إليه نظم انتاج المعرفة ولم تبلغه بعد.

المراجع :

1- Collarana, D., Galkin, M., Lange, C., Grangel-González, I., Vidal, M. E., & Auer, S. (2016). FuhSen: A Federated Hybrid Search Engine for building a knowledge graph on-demand .https://www.researchgate.net/profile/Diego-Collarana/publication/301765490_FuhSen_A_Federated_Hybrid_Search_Engine_for_building_a_knowledge_graph_on-demand/links/57a070b908ae94f454eae3a2/FuhSen-A-Federated-Hybrid-Search-Engine-for-building-a-knowledge-graph-on-demand.pdf

2- Landolsi, M. Y., Hlaoua, L., & Ben Romdhane, L. (2023). Information extraction from electronic medical documents: state of the art and future research directions. Knowledge and Information Systems, 65(2), 463-516 . https://link.springer.com/article/10.1007/s10115-022-01779-1

3- Cochez, M., Ristoski, P., Ponzetto, S. P., & Paulheim, H. (2017). Global RDF vector space embeddings. In The Semantic Web ISWC 2017: 16th International Semantic Web Conference, Vienna, Austria, October 2125, 2017, Proceedings, Part I 16 (pp. 190-207). SpringerInternationalPublishing. https://jyx.jyu.fi/bitstream/handle/123456789/56299/1/cochezetalglobalrdfembedding.pdf

4- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2022). Natural language processing: State of the art, current trends and challenges. Multimedia tools and applications, 1-32. https://link.springer.com/article/10.1007/s11042-022-13428-4

5- Collarana, D. (2020). Strategies and Techniques for Federated Semantic Knowledge Integration and Retrieval (Vol. 44). IOS Press. https://books.google.com/books?hl=ar&lr=&id=OkvYDwAAQBAJ&oi=fnd&pg=PP1&dq=STRATEGIES+ANDTECHNIQUES+FOR+FEDERATED+SEMANTIC+KNOWLEDGE+INTEGRATION+ANDRETRIEVAL&ots=dZiYJDRm3H&sig=LINCK8Z2nFd5YV_b6CPtEHSNXlE

6- Lampropoulos, G., Keramopoulos, E., & Diamantaras, K. (2020). Enhancing the functionality of augmented reality using deep learning, semantic web and knowledge graphs: A review. Visual Informatics, 4(1), 32-42. https://www.sciencedirect.com/science/article/pii/S2468502X20300012

7- Sousa, S., & Kern, R. (2022). How to keep text private? A systematic review of deep learning methods for privacypreserving natural language processing. Artificial Intelligence Review, 1-66. https://link.springer.com/article/10.1007/s10462-022-10204-6

8- Uc-Cetina, V., Navarro-Guerrero, N., Martin-Gonzalez, A., Weber, C., & Wermter, S. (2022). Survey on reinforcement learning for language processing. Artificial Intelligence Review, 133. https://link.springer.com/article/10.1007/s1046202210205-

9- Gu, J., Stefani, E., Wu, Q., Thomason, J., & Wang, X. E. (2022). Vision-and-language navigation: A survey of tasks, methods, and future directions. arXiv preprint arXiv:2203.12667. https://arxiv.org/pdf/2203.12667

10- Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R., & Ruberti, G. A. (2018). Ontology-Based Data Access and Integration. https://muei.etseib.masters.upc.edu/165/essi/ca/docs/slides-obda-2010-02-08/@@download/file/slides-OBDA-2010-02-08.pdf

11- Lin, B. (2022). Knowledge management system with nlp-assisted annotations: A brief survey and outlook. arXiv preprint arXiv:2206.07304. https://arxiv.org/pdf/220607304

12- Beek, W., Schlobach, S., & van Harmelen, F. (2016). A contextualised semantics for owl: sameAs. In The Semantic Web. Latest Advances and New Domains: 13th International Conference, ESWC 2016, Heraklion, Crete, Greece, May 29-June 2, 2016, Proceedings 13 (pp. 405-419). Springer International Publishing. https://www.cs.vu.nl/~frankh/postscript/ESWC2016.pdf

13- Joshi, K. P., & Saha, S. (2020). A Semantically Rich Framework for Knowledge Representation of Code of Federal Regulations. Digital Government: Research and Practice, 1(3), 117. https://dl.acm.org/doi/pdf/10.1145/3425192

14- Barsce, J. C., Palombarini, J. A., & Martínez, E. C. (2017, September). Towards autonomous reinforcement learning: Automatic setting of hyper-parameters using Bayesian optimization. In 2017 XLIII Latin American Computer Conference (CLEI) (pp. 19). IEEE. https://arxiv.org/pdf/1805.04748

15- Endris, K. M. (2020). Federated Query Processing over Heterogeneous Data Sources in a Semantic Data Lake (Doctoral dissertation, Universitätsund Landesbibliothek Bonn). https://core.ac.uk/download/pdf/322961743.pdf

16- Collarana, D. (2020). OP.CIT.

17- Bąk, J., Blinkiewicz, M., & Ławrynowicz, A. (2017, October). User-friendly visual creation of R2RML mappings in SQuaRE. In Proceedings of the Third International Workshop on Visualization and Interaction for Ontologies and Linked Data co-located

with the 16th International Semantic Web Conference ISWC 2017, Vienna, Austria (Vol. 1947, pp. 139-150). https://ceur-ws.org/Vol-1947/paper13.pdf

18- Masum, M., Shahriar, H., Haddad, H. M., Ahamed, S., Sneha, S., Rahman, M., & Cuzzocrea, A. (2020, December). Actionable Knowledge Extraction Framework for COVID-19. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 40364041). IEEE. https://scholar.archive.org/work/m75djnwusbfuxkjifmtg35lmxe/access/wayback/https://ieeexplore.ieee.org/ielx7/9377717/9377728/09378398.pdf

19- Bąk, J., & Blinkiewicz, M. (2016). SQuaRE: A Visual Tool for Creating R2RML Mappings. https://ceur-ws.org/Vol-1690/paper62.pdf

20- Bai, L., Rossi, L., Cui, L., & Hancock, E. R. (2016, December). A transitive aligned Weisfeiler-Lehman subtree kernel. In 2016 23rd International Conference on Pattern Recognition (ICPR) (pp. 396-401). IEEE. https://eprints.whiterose.ac.uk/115960/1/ICPR16_0434_MS_1_.pdf

21- Petrova, E., & Pauwels, P. (2020). Semantic Enrichment of Association Rules Discovered in Operational Building Data. August, São Paulo, Brazil. https://www.researchgate.net/profile/EkaterinaPetrova10/publication/340038053_Semantic_Enrichment_of_Association_Rules_Discovered_in_Operational_Building_Data/links/5f3b960592851cd302014e22/Semantic-Enrichment-of-Association-Rules-Discovered-in-Operational-Building-Data.pd

22- Collarana, D. (2020). OP.CIT.

23- Ramezani, M., Feizi-Derakhshi, M. R., & Balafar, M. A. (2022). Text-based automatic personality prediction using KGrAt-Net: a knowledge graph attention network classifier. Scientific Reports, 12(1), 21453. https://www.nature.com/articles/s41598-02225955-z

24- Bagosi, T., Calvanese, D., Hardi, J., Komla-Ebri, S., Lanti, D., Rezk, M., ... & Xiao, G. (2014). The ontop framework for ontology-based data access. In The Semantic Web and Web Science: 8th Chinese Conference, CSWS 2014, Wuhan, China, August 8-12, 2014, Revised Selected Papers 8 (pp. 67-77). Springer Berlin Heidelberg. http://www.inf.unibz.it/~dlanti/bibliography/papers/2014-csws-ontop.pdf

25- Gao, H., Wang, Y., Shao, J., Shen, H., & Cheng, X. (2022). User Identity Linkage across Social Networks with the Enhancement of Knowledge Graph and Time Decay Function. Entropy, 24(11), 1603. https://www.mdpi.com/1099-4300/24/11/1603/pdf

26- Ramezani, M., Feizi-Derakhshi, M. R., & Balafar, M. A. (2022). OP.CIT.

27- Arenas, M., Grau, B. C., Kharlamov, E., Marciuška, Š., & Zheleznyakov, D. (2016). Faceted search over RDF-based knowledge graphs. Journal of Web Semantics, 37, 55-74. https://ora.ox.ac.uk/objects/uuid:961f284b-9723-4341-ad4c-65b51fe14ab4/download_file?file_format=pdf&safe_filename=main.pdf&type_of_work=Journal+article

28- Arenas, M., Grau, B. C., Kharlamov, E., Marciuska, S., & Zheleznyakov, D. (2014, October). Enabling Faceted Search over OWL 2 with SemFacet. In OWLED (pp. 121-132).https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=96b2d4f9ce73af8d23f83baf61f9869221ca8b54

29- Hagendorff, T. (2021). Forbidden knowledge in machine learning reflections on the limits of research and publication. Ai & Society, 36, 767-781. https://link.springer.com/article/10.1007/s00146-020-01045-4

30- Collarana Vargas, D. (2019). Strategies and Techniques for Federated Semantic Knowledge Retrieval and Integration (Doctoral dissertation, Universitäts und LandesbibliothekBonn). https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/7906

31- Dalton, J., Xiong, C., & Callan, J. (2021). Cast 2020: The conversational assistance track overview. Technical report, technical report. https://www.cs.cmu.edu/afs/cs.cmu.edu/Web/People/callan/Papers/trec2021-dalton.pdf

32- Lbid.

33- conversational search assistant with transformers. In Advances in Information Retrieval: 43rd European Conference on IR Research, ECIR 2021, Virtual Event, March 28–April 1, 2021, Proceedings, Part I (pp. 130-145). Cham: Springer International Publishing. https://arxiv.org/pdf/2101.08197

34- Luo, C., Liu, D., Li, C., Lu, L., & Lv, J. (2021). Prediction, Selection, and Generation: Exploration of Knowledge-Driven Conversation System. arXiv preprint arXiv:2104.11454. https://arxiv.org/pdf/2104.11454

35- Ni, J., Young, T., Pandelea, V., Xue, F., & Cambria, E. (2022). Recent advances in deep learning-based dialogue systems: A systematic survey. Artificial intelligence review, 1-101. https://arxiv.org/pdf/2105.04387

36- Qu, C., Yang, L., Chen, C., Qiu, M., Croft, W. B., & Iyyer, M. (2020, July). Open-retrieval conversational question answering. In Proceedings of the 43rd International ACM SIGIR

conference on research and development in Information Retrieval (pp. 539-548). https://dl.acm.org/doi/pdf/10.1145/3397271.3401110

37- Følstad, A., Araujo, T., Law, E. L. C., Brandtzaeg, P. B., Papadopoulos, S., Reis, L., ... & Luger, E. (2021). Future directions for chatbot research: an interdisciplinary research agenda. Computing, 103(12), 2915-2942. https://link.springer.com/article/10.1007/s00607-021-01016-7

38- Ni, J., Young, T., Pandelea, V., Xue, F., & Cambria, E. (2022). OP. CIT.

39- Zaib, M., Zhang, W. E., Sheng, Q. Z., Mahmood, A., & Zhang, Y. (2022). Conversational question answering: A survey. Knowledge and Information Systems, 64(12), 3151-3195. https://link.springer.com/article/10.1007/s10115022-01744-y

40- Wang, L., Lu, J., Zhou, G., Pan, H., Zhu, T., Huang, N., & He, P. (2022). Representation Learning Method with Semantic Propagation on Text-Augmented Knowledge Graphs. Computational Intelligence and Neuroscience, 2022. https://www.hindawi.com/journals/cin/2022/1438047/

41- Ni, J., Young, T., Pandelea, V., Xue, F., & Cambria, E. (2022). OP. CIT.

42- Atef Mosa, M. (2021). Predicting semantic categories in text based on knowledge graph combined with machine learning techniques. Applied Artificial Intelligence, 35(12), 933-951. https://www.tandfonline.com/doi/pdf/10.1080/08839514.2021.1966883

43- Kejriwal, M. (2022). Knowledge graphs: A practical review of the research landscape. Information, 13(4), 161. https://www.mdpi.com/20782489/13/4/161/pdf

44- Aasman, J., & Mirhaji, P. (2018, January). Knowledge graph solutions in healthcare for improved clinical outcomes. In CEUR Workshop Proceedings (Vol. 2180). CEUR-WS. https://einstein.pure.elsevier.com/en/publications/knowledge-graph-solutions-in-healthcare for-improved-clinical-out

45- Lygerakis, F., Kampelis, N., & Kolokotsa, D. (2022). Knowledge Graphs' Ontologies and Applications for Energy Efficiency in Buildings: A Review. Energies, 15(20), 7520. https://www.mdpi.com/1996-1073/15/20/7520/pdf

46- Gao, M., Lu, J., & Chen, F. (2022). Medical Knowledge Graph Completion Based on Word Embeddings. Information, 13(4), 205. https://www.mdpi.com/2078-2489/13/4/205/pdf

47- Jahanshahi, H., Kazmi, S., & Cevik, M. (2022). Auto response generation in online medical chat services. Journal of Healthcare Informatics Research, 6(3), 344-374. https://link.springer.com/article/10.1007/s41666-022-00118-x

48- Liang, L., Li, Y., Wen, M., & Liu, Y. (2022). KG4Py: A toolkit for generating Python knowledge graph and code semantic search. Connection Science, 34(1), 13841400. https://www.tandfonline.com/doi/pdf/10.1080/09540091.2022.2072471

49- Cortis, K., & Davis, B. (2021). Over a decade of social opinion mining: a systematic review. Artificial intelligence review, 54(7), 48734965. https://link.springer.com/article/10.1007/s10462-021-10030-2

50- Albagli-Kim, S., & Beimel, D. (2022). Knowledge Graph-Based Framework for Decision Making Process with Limited Interaction. Mathematics, 10(21), 3981. https://www.mdpi.com/22277390/10/21/3981/pdf

51- Montenegro, J. L. Z., & da Costa, C. A. (2022). The hope model architecture: a novel approach to pregnancy information retrieval based on conversational agents. Journal of Healthcare Informatics Research, 6(3), 253-294. https://link.springer.com/article/10.1007/s41666022-00115-0

52- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), 5731-5780. https://link.springer.com/article/10.1007/s1046202210144-1

53- Li, H. (2022). Piano Teaching Knowledge Graph Construction Based on Cross-Media Data Analysis and Semantic Network. Computational Intelligence and Neuroscience, 2022. https://www.hindawi.com/journals/cin/2022/5499593/

54- Collarana, D. (2020). OP.CIT.

55- Gongane, V. U., Munot, M. V., & Anuse, A. D. (2022). Detection and moderation of detrimental content on social media platforms: current status and future directions. Social Network Analysis and Mining, 12(1), 129. https://link.springer.com/article/10.1007/s13278-022-00951-3

56- Zhu, G., Hao, M., Zheng, C., & Wang, L. (2022). Design of Knowledge Graph Retrieval System for Legal and Regulatory Framework of Multilevel Latent Semantic Indexing. Computational Intelligence and Neuroscience, 2022. https://www.hindawi.com/journals/cin/2022/6781043/

57- Xu, J., Mayer, W., Zhang, H., He, K., & Feng, Z. (2022). Automatic Semantic Modeling for Structural Data Source with the Prior Knowledge from Knowledge Base. Mathematics, 10(24), 4778. https://www.mdpi.com/article/10.3390/math10244778

58- Wang, L., Lu, J., Zhou, G., Pan, H., Zhu, T., Huang, N., & He, P. (2022). Representation Learning Method with Semantic Propagation on Text Augmented Knowledge Graphs.

Computational Intelligence and Neuroscience, 2022. https://www.hindawi.com/journals/cin/2022/1438047/

59- Huck, J. (2022). Knowledge Graphs, Metadata Practices, and Badiou's Mathematical Ontology. KULA, 6(3), 117. https://www.erudit.org/en/journals/kula/1900v1-n1-kula07188/1091366ar.pdf

60- Jiang, S., Liu, Y., Zhang, Y., Luo, P., Cao, K., Xiong, J., ... & Wei, J. (2022). Reliable semantic communication system enabled by knowledge graph. Entropy, 24(6), 846. https://www.mdpi.com/10994300/24/6/846/pdf

61- Nian, Y., Hu, X., Zhang, R., Feng, J., Du, J., Li, F., ... & Tao, C. (2022). Mining on Alzheimer's diseases related knowledge graph to identity potential AD related semantic triples for drug repurposing. BMC bioinformatics, 23(6), 1-15. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859022-04934-1

62- Silva, M. C., Eugénio, P., Faria, D., & Pesquita, C. (2022). Ontologies and Knowledge Graphs in Oncology Research. Cancers, 14(8), 1906. https://www.mdpi.com/20726694/14/8/1906/pdf

63- Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). Deep Learning applications for COVID-19. Journal of big Data, 8(1), 1-54. https://link.springer.com/article/10.1186/s40537-020-00392-9%23authTaghi_M_-Khoshgoftaar

64- Opdahl, A. L., Al-Moslmi, T., Dang-Nguyen, D. T., Gallofré Ocaña, M., Tessem, B., & Veres, C. (2022). Semantic Knowledge Graphs for the News: A Review. ACM Computing Surveys, 55(7), 138. https://dl.acm.org/doi/pdf/10.1145/3543508

65- Khalili, A., Van Andel, P., Van Den Besselaar, P., & De Graaf, K. A. (2017, December). Fostering serendipitous knowledge discovery using an adaptive multigraph-based faceted browser. In Proceedings of the Knowledge Capture Conference. https://research.ld-r.org/papers/kcap_fosteringSerendipity.pdf

66- Mudgal, S., Li, H., Rekatsinas, T., Doan, A., Park, Y., Krishnan, G., ... & Raghavendra, V. (2018, May). Deep learning for entity matching: A design space exploration. In Proceedings of the 2018 International Conference on Management of Data (pp. 19-34). https://pages.cs.wisc.edu/~anhai/papers1/deepmatcher-tr.pdf

67- Collarana, D., Galkin, M., Lange, C., Scerri, S., Auer, S., & Vidal, M. E. (2018). Synthesizing Knowledge Graphs from Web Sources with the MINTE^++ Framework. In The Semantic Web–ISWC 2018: 17th International Semantic Web Conference, Monterey,

CA, USA, October 8–12, 2018, Proceedings, Part II 17 (pp. 359-375). Springer International Publishing. https://www.researchgate.net/profile/Diego-Collarana/publication/325381996_Synthesizing_Knowledge_Graphs_from_web_sources_with_the_MINTE_framework/links/5b290b4b0f7e9b1d003575ca/Synthesizing-Knowledge-Graphs-from-web-sources-with-the-MINTE-framework.pdf

68- Collarana, D., Galkin, M., Traverso-Ribón, I., Lange, C., Vidal, M. E., & Auer, S. (2017, January). Semantic data integration for knowledge graph construction at query time. In 2017 IEEE 11th International Conference on Semantic Computing (ICSC) (pp. 109-116). IEEE. https://www.researchgate.net/profile/Diego-Collarana/publication/315867282_Semantic_Data_Integration_for_Knowledge_Graph_Construction_at_Query_Time/links/5a33ac10458515afb691e06a/Semantic-Data-Integration-for-Knowledge-Graph-Construction-at-Query-Time.pdf

69- Ostrowski, D., & Kim, M. (2017, January). A semantic based framework for the purpose of big data integration. In 2017 IEEE 11th International Conference on Semantic Computing (ICSC) (pp. 305-309). IEEE. https://ieeexplore.ieee.org/abstract/document/7889553/

70- Fuenmayor, L., Collarana, D., Lohmann, S., & Auer, S. (2017, October). FaRBIE: A Faceted Reactive Browsing Interface for Multi RDF Knowledge Graph Exploration. In VOILA@ ISWC (pp. 111-122). http://ceur-ws.org/Vol-1947/paper10.pdf

71- Freire, N., Charles, V., & Isaac, A. (2018, June). Evaluation of Schema. org for aggregation of cultural heritage metadata. In The Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, Proceedings (pp. 225-239). Cham: Springer International Publishing. https://www.researchgate.net/profile/Nuno-Freire-2/publication/325529501_Evaluation_of_Schemaorg_for_Aggregation_of_Cultural_Heritage_Metadata/links/5b22540da6fdcc69745ee712/Evaluation-of-Schemaorg-for-Aggregation-of-Cultural-Heritage-Metadata.pdf

72- Traverso, I., Vidal, M. E., Kämpgen, B., & Sure-Vetter, Y. (2016, September). GADES: a graph-based semantic similarity measure. In Proceedings of the 12th International Conference on Semantic Systems (pp. 101-104). https://dl.acm.org/doi/abs/10.1145/2993318.2993343

73- Sequeda, J. F., & Miranker, D. P. (2017). Towards a Pay-as-you-go Methodology for Ontology and Mapping Engineering in Ontology Based Data Access. In ISWC (Posters, Demos & Industry Tracks). https://ceur-ws.org/Vol-1963/paper612.pdf