

ترويض فوضى الذكاء الاصطناعي التوليدي: دمج الرسوم البيانية المعرفية في نظم المكتبات الرقمية

د. جنيفر دي سوزا



الملخص:

ينطلق هذا المقال من أن النشر العلمي ظل منذ القرن السابع عشر متمحورًا حول الوثيقة، فبقيت المعرفة العلمية غير مهيكلية ويصعب على الآلات تفسيرها، ويدعو إلى النشر القائم على الرسوم البيانية المعرفية لجعل هذه المعرفة متوافقة مع مبادئ (FAIR). وتتخذ المؤلفات النماذج اللغوية الكبيرة مثالًا، إذ تنشت المعلومات عنها بين المقالات والمدونات ومستودعات الشيفرة. وتعرض منصة الرسم البياني المعرفي للبحث المفتوح (ORKG) التابعة لمركز (TIB)، وأسسها التقنية القائمة على (RDF) والمعرفات الدائمة، ثم تصف قالبًا من 27 خاصية لوصف هذه النماذج، وفهرسًا من المقارنات والتصورات المرئية والمراجعات، وإمكانية الاستعلام عبر (SPARQL). وتقارن هذا النهج ببطاقات النماذج في (Hugging Face)، وتناقش تحديات الجودة والاستدامة. وتخلص إلى أن هذا التمثيل قابل للنقل إلى مجالات أخرى، ويحد من عزلة المعلومات ويعزز الشفافية وقابلية إعادة الإنتاج.

الكلمات المفتاحية: الرسوم البيانية المعرفية، النشر الدلالي، النماذج اللغوية الكبيرة، مبادئ (FAIR)، المكتبات الرقمية

مقدمة:

ثمة جبل متنامٍ من البحوث. لكن الأدلة تتزايد على أننا نفرق اليوم في الوحد مع اتساع التخصص. فالباحث يترنح أمام نتائج آلاف العاملين الآخرين واستنتاجاتهم، وهي استنتاجات لا يجد الوقت لاستيعابها، فضلًا عن تذكرها، حين تظهر. ... إن السجل، إن كان له أن يكون مفيدًا للعلم، يجب أن يُوسَّع باستمرار، ويجب أن يُخزَّن، وفوق كل شيء يجب أن يُرجَّع إليه. [1]

— فانيفار بوش (Vannevar Bush)، «(As We May Think)» (1945)

يفرض الحجم الهائل للتقدم العلمي تحديات كبيرة على الباحثين، الذين يتعين عليهم غرلة كميات ضخمة من المعلومات لاستخلاص رؤى ذات معنى. ويتفاقم هذا التحدي بسبب الطابع غير المهيكل في معظمه للاتصال العلمي الحالي، وهو ما يعوق الشفافية والتكامل وكفاءة التحكيم العلمي والمساعدة الآلية. وللتغلب على هذه القيود، يجب أن تنتقل السجلات البحثية من الصيغ التقليدية غير المهيكلة إلى تمثيلات قابلة للمعالجة الآلية ومهيكلية دلاليًا، مثل الرسوم البيانية المعرفية (KGs)، بما يتيح أدوات بحث أذكى وأكثر كفاءة. ويتوافق هذا التحول أيضًا مع مبادئ (FAIR)، أي قابلية العثور والوصول والتشغيل البيئي وإعادة الاستخدام [2]، بما يضمن قابلية الاستخدام للبشر والآلات على السواء. ويسعى الرسم البياني المعرفي للبحث المفتوح (ORKG)، وهو مشروع رائد لمركز لايبنتس للمعلومات في العلوم والتقنية والمكتبة الجامعية (TIB)، إلى معالجة هذه التحديات بإتاحة التمثيل المهيكل للإسهامات العلمية ومقارنتها بحسب خصائصها الرئيسية. وكما تتيح منصات التجارة الإلكترونية مقارنة المنتجات جنبًا إلى جنب، ييسر (ORKG) الاستكشاف الفعال للمعرفة العلمية، فاتحًا فرصًا جديدة للباحثين والمكتبات الرقمية [3][4][5].

(ORKG) منصة وبنية تحتية على الويب مصممة لهيكلية المعرفة العلمية وربط عناصرها باستخدام التقنيات الدلالية. فبدلًا من الاكتفاء بالنشر القائم على النص، تمثل المنصة الإسهامات البحثية في صورة رسوم بيانية معرفية مترابطة، بما يتيح مقارنة النتائج واكتشافها ومراجعتها وتولييفها بكفاءة أكبر، وفق مبادئ (FAIR). ويستطيع الباحثون والمكتبيون تقديم أوصاف مهيكلية للأوراق البحثية، لجعل المعرفة أيسر وصولًا للبشر والآلات. والمنصة خدمة مفتوحة لمجتمع المستخدمين، وتعمل ببرمجيات مفتوحة المصدر. وتتكامل مع خدمات المعارف الدائمة، مثل (DOI) و(ORCID)، لتحسين إدارة البيانات الوصفية العلمية. كما تتيح الإسهامات متعددة التخصصات وتدعم التعاون والعلم المفتوح، وتوفر حزمة (Python) لتمكين المطورين من دمج بياناتها وإمكاناتها في تطبيقات خارجية، وتعزيز اكتشاف المعرفة وأتمتة البحث.

تتبع التطورات السريعة في الذكاء الاصطناعي التوليدي:

بدأ تطوير النماذج اللغوية الكبيرة عام 2018 مع (GPT-1) من (OpenAI) و(BERT) من (Google)، اللذين كانا رائدين في المعماريات القائمة على المحولات (transformers) لفهم اللغة. ومنذ ذلك الحين، اتسم التطور السريع للنماذج اللغوية الكبيرة بموجة من الإصدارات المبتكرة، كل منها يدفع حدود الأداء والكفاءة وسهولة الوصول [6]. وشهد عام 2023 إصدار نماذج مثل (Mistral 7B) الذي أدخل آلية الانتباه بالاستعلامات المجمعة (grouped-query attention) لتعزيز الكفاءة الحاسوبية، في حين استفاد (Mixtral 8x7B) من معماريات مزيج الخبراء المتناثر (sparse mixture-of-experts) لتحسين قابلية التوسع دون التضحية بالأداء. ومن التطورات البارزة الأخرى (LLaMA 2) من (Meta AI)، الذي حسن إمكانات الضبط الدقيق (fine-tuning) وفهم السياق، و(GPT-4) من (OpenAI)، الذي وسّع الإمكانات متعددة الوسائط توسيعًا كبيرًا، بما في ذلك إدخال الصور والنصوص. وإضافة إلى ذلك، ركزت سلسلة (Claude) من (Anthropic) على السلامة وقابلية التفسير في تقنيات المواءمة (alignment)، وأظهرت نماذج (Gemini) من (Google) تكاملًا متقدمًا مع الأدوات الخارجية، ووضعت معيارًا جديدًا للذكاء الاصطناعي الموجه نحو المهام. وثمة تركيز متزايد على التطوير مفتوح المصدر، والكفاءة الحاسوبية، والتعامل مع سياقات ممتدة، والإمكانات متعددة الوسائط. وتوضح هذه الاختراقات الوتيرة المتواصلة للابتكار، المدفوعة بمزيج من التحسينات المعمارية ومجموعات البيانات الموسعة والجهود التعاونية في مجتمع الذكاء الاصطناعي. ويبرز هذا التدفق المستمر للابتكارات الإمكانات التحويلية للنماذج اللغوية الكبيرة، كما يُبرز التحدي المتزايد المتمثل في البقاء على اطلاع في مشهد يتطور بهذه السرعة.

ويتطلب التدفق المستمر للنماذج والتقنيات ومجموعات البيانات الجديدة يقظة دائمة وقدرة على التكيف. فالمعلومات الرئيسة عن النماذج اللغوية الكبيرة كثيرًا ما تكون مبعثرة بين المقالات العلمية وتدوينات المؤسسات ومستودعات الشيفرة المصدرية، مما يصعب تتبع التطورات بكفاءة. ويمكن مقارنة النماذج اللغوية الكبيرة باستخدام تفاصيل رئيسة مثل اسم النموذج، والمؤسسة، وتاريخ الإصدار، وبيانات التدريب المسبق، واستراتيجيات الضبط الدقيق، والبصمة الكربونية. ونمذجة هذه الخصائص المتسقة تدعم المقارنة وقابلية البحث على نحو أفضل، وهي حالة استخدام ممتازة للرسم البياني المعرفي للبحث المفتوح.

الخلفية: الرسم البياني المعرفي للبحث المفتوح:

من الاعتبارات التقنية الأساسية عند نشر الرسوم البيانية المعرفية تحديد الكيانات والسمات والعلاقات البارزة لنماذج الذكاء الاصطناعي التوليدي بوصفها موارد معيارية يمكن الوصول إليها على الويب عبر معرفات الموارد الموحدة (Uniform Resource Identifiers, URIs). وينبغي أيضًا تمثيل العبارات المعلوماتية بصيغة تسلسلية عن النماذج اللغوية الكبيرة في صورة ثلاثيات (triples) مهيكلية تتوافق مع صياغة إطار وصف الموارد (Resource Description Framework, RDF) التي يوصي بها اتحاد الشبكة العالمية (W3C) [7]. والمفردات المعيارية والبيانات الوصفية ونشر البيانات على الويب باستخدام معرفات (URI) قابلة للاستدعاء (dereferenceable) أمور حاسمة لتشغيل البياني الدلالي والربط بمجموعات البيانات الخارجية، بما يتيح منظومة حقيقية للبيانات المترابطة (linked data).

وهذه الوظائف مدعومة دعمًا كاملاً في (ORKG)، وهو منصة نشر علمي من الجيل التالي. وفي هذا السياق، ليس (ORKG) برمجية فحسب، بل هو أيضًا فضاء أسماء (namespace) لتحديد المفردات العلمية عبر معرفات (URI)، وواجهات بديهية لتعريف خصائص الموارد وأصنافها وقوالبها، وتقنيات خلفية لتمثيل الرسم البياني المعرفي بصياغة (RDF). وإضافة إلى ذلك، تدعم بنيته الدلالية القابلة للمعالجة الآلية عروض مقارنة متقدمة، على غرار مقارنات المنتجات في التجارة الإلكترونية، وتتيح استعلامات (SPARQL) للحصول على عروض مخصصة أو مجمعة للمعرفة العلمية المهيكلية.

ويعتمد (ORKG) نموذجًا تعاونيًا ديناميكيًا لإنشاء المعرفة مستوحى من ويكيبيديا. يستطيع فيه أي شخص الإسهام بالبيانات باستخدام المفردات التي يختارها لتمثيل كيانات مترابطة عبر روابط (RDF)، فيتشكل رسم بياني عالمي للبيانات قابل للاكتشاف. وتتيح طبيعته المفتوحة تحرير محتوى المنتجات البحثية، مثل الأوراق والمقارنات، وتحديثه وإعادة نشره، مع تخزين جميع الإصدارات للرجوع إليها مستقبلاً. وتتيح المعرفات الدائمة، مثل (DOI)، الإحالة الدقيقة إلى إصدارات بعينها، في حين تضمن البيانات الوصفية للمصدر والنشأة (provenance)، التي تسجل تفاصيل مثل المنشئ وتاريخ الإنشاء والأساليب، إمكانية التتبع والجدارة بالثقة وتقييم الجودة.

وتعزز معرفات (DOI) المخصصة لمكونات الرسم البياني قابلية الاكتشاف عبر البنى التحتية الببليومترية العلمية العالمية مثل (DataCite) و (Crossref)، كما يمكن العثور على الموارد عبر محركات البحث. ويضمن الوصول عبر بروتوكولات (HTTP) وواجهات (REST) البرمجية وواجهة مستخدم، مع إمكانية الوصول إلى البيانات الوصفية بمعزل عن بيانات الرسم البياني. ويحقق (ORKG) التشغيل البياني باستخدام (RDF)، وهو المعيار الذي يوصي به (W3C) لتمثيل المعرفة المقروءة آلياً، ويعزز قابلية إعادة الاستخدام بتوليد البيانات الوصفية للمصدر والنشأة تلقائياً ونشر بيانات الرسم البياني بترخيص (CC BY-SA). وتدعم هذه الخصائص مجتمعةً نشر المعرفة العلمية على نحو فعال وجدير بالثقة ومتوافق مع مبادئ (FAIR).

ترويض فوضى الذكاء الاصطناعي التوليدي: فهرس للنماذج اللغوية الكبيرة على (ORKG):

يبدأ نشر المعلومات عن النماذج اللغوية الكبيرة على (ORKG) بتحديد خصائص معيارية لوصفها وصفًا مهيكلًا قابلاً للمقارنة. واستلهامًا من (Hugging Face)، وهي مركز رئيس للنماذج اللغوية الكبيرة مفتوحة المصدر وبطاقات نماذجها، طوّرت المؤلفة مفردات ومخططًا معياريين لدعم التقاط المعرفة وتمثيلها الدلالي.

قالب (ORKG) لمواصفات النماذج اللغوية الكبيرة:

التسميات المعيارية، لضمان الاتساق، أنشأنا مفردات مضبوطة (controlled vocabulary) تشكّل أساس مخطط مواصفات النماذج اللغوية الكبيرة (LLM Specification Schema). واشتقت الخصائص من بطاقات النماذج في (Hugging Face)، مع إضافة محمولات (predicates) خاصة بالنماذج اللغوية الكبيرة إلى فضاء أسماء (ORKG) على الويب لتوسيع المفردات، ولمواءمة هذه الخصائص مع الأنطولوجيات الخارجية، استُخدمت علاقة (same-as) في (RDF) لإثبات التكافؤ بين المحمولات.

استخدام قوالب (ORKG). لتسجيل معلومات النماذج اللغوية الكبيرة تسجيلًا موحدًا، أنشأنا قالبًا يعتمد استمارات الإدخال (form-based) يتألف من خصائص محددة مسبقًا. ويستفيد نظام القوالب هذا من أنماط الخصائص المتكررة في الرسوم البيانية الفرعية في (ORKG)، بما يتيح تحديد السمات ذات الصلة على نحو متسق عبر إسهامات متعددة. ويتضمن القالب 27 خاصية مهيكلية تصف النماذج اللغوية الكبيرة وصفًا شاملاً، منها:

- التعريف: اسم النموذج، وعائلة النموذج، والمؤسسة، وتاريخ الإنشاء.
- تفاصيل التطوير: معمارية التدريب المسبق، ومهمة التدريب المسبق، ومدونة التدريب المسبق، وحجم مدونة التدريب (بمليارات الرموز (tokens))، وتاريخ انتهاء نطاق بيانات التدريب، والابتكار.
- الضبط الدقيق: مهمة الضبط الدقيق وبيانات الضبط الدقيق للتحسين.
- المواصفات التقنية: المُحسّن (optimizer)، والمُرمّز (tokenizer)، وعدد المعاملات، والحد الأقصى لعدد المعاملات (بالملايين)، وطول السياق (بالرموز).
- اللغة والاستخدام: اللغة المدعومة، والامتداد، والتطبيق.
- البنية التحتية والأثر: العتاد المستخدم، ووصف العتاد، والكربون المنبعث (tCO2eq).
- الشفافية والترخيص: إتاحة الشيفرة المصدرية، والتدوين، والترخيص، والمشكلة البحثية المعالجة.

وتضمن هذه الخصائص تمثيلًا شاملاً للنماذج اللغوية الكبيرة، بما يتيح مقارنات مهيكلية وتحليلًا مفصلاً. فعلى سبيل المثال، تُدعم الشفافية بسمات مثل إتاحة الشيفرة المصدرية، في حين يلتقط الأثر البيئي من خلال بيانات انبعاثات الكربون.

ومخطط (ORKG) لمواصفات النماذج اللغوية الكبيرة، الذي يضم هذه الخصائص السبع والعشرين، متاح على منصة (ORKG). ويمكن هذا المخطط الباحثين من تسجيل النماذج اللغوية الكبيرة ومقارنتها وتحليلها بدقة واتساق، بما يعزز الشفافية وسهولة الوصول في دراسة الذكاء الاصطناعي التوليدي.

فهرس (ORKG) للأوراق والمقارنات والتصورات المرئية والمراجعات المتعلقة بالنماذج اللغوية الكبيرة:

تُنشر المعلومات عن أوراق النماذج اللغوية الكبيرة في صورة أوصاف دلالية مهيكلية باستخدام وحدة الأوراق (papers module)، ثم تُجمع في مقارنات (ORKG) للبحوث (ORKG Comparisons). وهذه المقارنات النوع الأساسي من محتوى المنصة، وتقدم عرضًا مكثفًا لأحدث البحوث في مسألة بعينها. وتُنظم الإسهامات في جداول يمكن مقارنتها وتصفيتها بحسب خصائص مختلفة. وتضم إحدى المقارنات المنشورة أوصافًا مهيكلية قابلة للمعالجة الآلية لـ 92 نموذجًا لغويًا كبيرًا، مع بيانات وصفية مثل العنوان والمنشئ ووصف موجز. ويمكن نشر المقارنات بأي مستوى من التفصيل؛ فمثلًا يمكن جمع الأوصاف المهيكلية للنماذج المبكرة من (OpenAI)، أي (GPT-1) و 2 و 3، في مقارنة مستقلة.

تتيح المعلومات الدلالية الدقيقة القابلة للمعالجة الآلية في (ORKG) تفاعلات ديناميكية، مثل توليد تمثيلات بصرية من بيانات المقارنات. ومن الأمثلة مقارنة أعدتها المؤلفة بعنوان «A Catalog of DeepMind's LLMs including their seminal Chinchilla model»، وتضم (AlphaFold) و (Gopher) و (Chinchilla) و (GopherCite) و (Flamingo) و (Gato) و (Sparrow). ويمكن مقارنة عدد معاملات أكبر نسخة من كل نموذج، بالملايين، واستكشاف أسئلة مثل: ما أكبر نموذج أصدرته (DeepMind)؟ وما أقصى تفاوت بين أحجام نماذجها؟ ويوضح ذلك كيف تقدم المعرفة القابلة للمعالجة الآلية رؤى سريعة في مسائل بحثية محددة.

وأخيرًا، يقدم (ORKG) «المراجعات» (Reviews)، وهي أداة مصممة لتأليف مقالات المراجعة ونشرها، تيسر إنشاء المجتمع لمقالات ديناميكية حية باستخدام المقارنات من الرسم البياني المعرفي للبحث المفتوح لتقديم معرفة قابلة للمعالجة الآلية. فمثلًا، تجمع المراجعة المتاحة على المنصة، والواردة ضمن المراجع، مقارنات متعددة للنماذج اللغوية الكبيرة. وتقدم رؤى تفصيلية في خصائص النماذج اللغوية الكبيرة التي طورتها مؤسسات مختلفة، مع دمجها كذلك في مقارنة أوسع لجميع النماذج. وهذا النهج مثالي لمقالات المراجعة، إذ يتيح تحليلات مفصلة للدراسات، أو في هذه الحالة، مقارنات للنماذج اللغوية الكبيرة التي جرت مراجعتها.

قابلية الاستعلام:

يمكن استخدام استعلامات (SPARQL) لتوليد عروض أصغر ومحددة للبيانات. فمثلًا، يمكن ترجمة الاستعلام أدناه إلى سؤال بلغة طبيعية على النحو التالي: ما النماذج اللغوية الكبيرة التي أنشئت في عام 2023؟ ويعيد هذا الاستعلام جدولًا (يقتصر على أول 100 نتيجة) بعنوانين الأوراق والنماذج اللغوية الكبيرة المصاحبة لها وتاريخ إنشاء كل نموذج. ومن التوسعات الطبيعية لهذا الاستعلام تغيير السنة، بل القدرة على تحديد نطاق من التواريخ بصيغة (yyyy-mm-dd) يُبحث ضمنه. وهذه نقطة استعلام مفتوحة متاحة لأي مستخدم للويب.

```
PREFIX orkgp: <http://orkg.org/orkg/predicate/>
PREFIX orkgc: <http://orkg.org/orkg/class/>
PREFIX orkgr: <http://orkg.org/orkg/resource/>
SELECT ?paper_label ?model_label ?date_created
WHERE {
  orkgr:R609337 orkgp:compareContribution ?contribution .
  ?contribution orkgp:HAS_MODEL ?model .
  ?model rdfs:label ?model_label .
  ?paper orkgp:P31 ?contribution .
  ?paper rdfs:label ?paper_label .
  ?contribution orkgp:P49020 ?date_created .
  FILTER(STRSTARTS(STR(?date_created), "2023"))
```

}

LIMIT 100

ومع أن هذه الورقة تعرض حالة الاستخدام الخاصة بالنماذج اللغوية الكبيرة، فإن إمكانيات ربط المعلومات على الويب وفرصه كبيرة.

الأعمال ذات الصلة:

أقرب تمثيل للمعلومات الملتقطة عبر قالب (ORKG) لنماذج المحولات أو النماذج اللغوية الكبيرة هو بطاقات النماذج (model cards) الصادرة لهذه النماذج على (Hugging Face Hub) [8]. غير أن بطاقات النماذج والتمثيل الدقيق للمعرفة في الرسوم البيانية المعرفية مثل (ORKG) يختلفان اختلافاً كبيراً من حيث قابلية الاكتشاف وقابلية إعادة الإنتاج والمشاركة. فبطاقات النماذج، مرة أخرى، وثائق يقرأها البشر تقدم شروحاً للمشروعات، لكن قابلية اكتشافها تقتصر على البحث بالكلمات المفتاحية في المستودعات، وكثيراً ما تفتقر إلى البنية الدلالية اللازمة للاكتشاف الآلي الدقيق. ومع أن التوصيات الخاصة ببطاقات النماذج تتضمن ملء القيم ذات الصلة لخصائص النموذج، فإن هذا التمثيل المهيكل، وإن كان قد يحسّن قابلية القراءة البشرية للمرور السريع على خصائص النموذج، يفتقر إلى قابلية المعالجة الآلية لأنه يظل مضمّناً في وثيقة. وفي المقابل، توفر الرسوم البيانية المعرفية تمثيلات مقروءة آلياً ومهيكلية وغنية دلاليّاً تعزز قابلية الاكتشاف عبر الاستعلام المتقدم ووسم البيانات الوصفية. وهي تتفوق في قابلية إعادة الإنتاج بالتقاط بيانات وصفية مفصلة ومعيارية، وتتبع التحديثات، وربط الأساليب ومجموعات البيانات بشفافية. وفضلاً عن ذلك، تتيح الرسوم البيانية المعرفية المشاركة والتكامل السلسين داخل أنظمة مترابطة، بما يعزز التعاون والتشغيل البيئي، خلافاً للطبيعة الثابتة والمستقلة لأي وثيقة نصية، أيّا كانت الصيغة المقررة لكتابتها من أجل قابلية قراءة بشرية أفضل.

ورغم مزاياه، يطرح تطبيق نهج مهيكل قائم على الرسوم البيانية المعرفية لتمثيل النماذج تحديات، ولا سيما في ضمان جودة المقالات المؤلفة من المراجعات المهيكلية. فمع أن (ORKG) يعزز قابلية الاكتشاف وقابلية إعادة الإنتاج من خلال الجوانب الرئيسة للمحتوى العلمي المقروءة آلياً، يظل إشراف الخبراء ضرورياً لتحويل البيانات المهيكلية إلى سرديات عالية الجودة. والاستدامة عامل رئيس آخر؛ فخلافاً للجهود القائمة على المتطوعين، يستفيد (ORKG) من تمويل حكومي بوصفه المشروع الرائد لمركز (TIB)، بما يضمن تطويره على المدى الطويل. ويتيح هذا الدعم تحسينات مستمرة في الأتمتة وسهولة الاستخدام، مما يقلل الجهد المطلوب من الباحثين للإسهام بالمعرفة المهيكلية. ومن خلال دمج الاستخلاص بمساعدة الآلة، وتحفيز الإسهامات، والمواءمة مع مسارات عمل النشر العلمي، يعزز (ORKG) تبنيّه على نطاق أوسع. وبذلك يوفر إطاراً مستداماً وقابلًا للتوسع لتحسين قابلية اكتشاف النماذج وقابلية إعادة إنتاجها ومشاركة المعرفة.

الخاتمة:

يسلط هذا المقال الضوء على الإمكانيات التحويلية لاستخدام الرسوم البيانية المعرفية، وتحديدًا الرسم البياني المعرفي للبحث المفتوح (ORKG)، لتمثيل المعلومات عن تقنيات الذكاء الاصطناعي التوليدي، مثل النماذج اللغوية الكبيرة (LLMs)، وإدارتها. فباعتماد نهج مهيكل ودلالي وقابل للمعالجة الآلية، يضمن (ORKG) أن تلتزم المعرفة الناتجة بمبادئ (FAIR)، أي قابلية العثور والوصول والتشغيل البيئي وإعادة الاستخدام.

والتمثيل القائم على الرسوم البيانية المعرفية الذي يعرضه هذا المقال لا ينطبق على النماذج اللغوية الكبيرة فحسب، بل يسهل نقله إلى مجالات علمية أخرى. فمنصة (ORKG) تجسد تنوع استخدامات المكتبة الرقمية الحديثة، إذ تقدم بنية تحتية عامة تمكّن المستخدمين من أي مجال من نمذجة بحوثهم في صورة معرفة مهيكلية باستخدام القوالب والأوراق والمقارنات والتصورات المرئية والمراجعات. ويمكن تحرير هذه السجلات الحية وتنسيقها تعاونيًا، بما يضمن أن تظل المعرفة ديناميكية ومحدّثة ومتاحة للبشر والآلات على السواء.

ومن خلال معالجة تحديات المعرفة المجزأة غير المهيكلية، يحد هذا النهج من نشوء جزر معلوماتية معزولة، ويعزز قدرًا أكبر من الشفافية وقابلية إعادة الإنتاج وقابلية الاستخدام. وفي نهاية المطاف، تمهد الاستفادة من (ORKG) والمنصات المماثلة الطريق لمستقبل أكثر ترابطًا وقابلية للمعالجة الآلية في نشر المعرفة العلمية واكتشافها.

المراجع:

- [/https://orkg.org](https://orkg.org)
 - [/https://gitlab.com/TIBHannover/orkg](https://gitlab.com/TIBHannover/orkg)
 - [/https://orkg.readthedocs.io](https://orkg.readthedocs.io)
 - <https://orkg.org/comparison/R1355351>
 - <https://orkg.org/review/R640001>
1. Bush, V. (1945). As We May Think. The Atlantic. Retrieved from <https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>
 2. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, et al. 2016. The FAIR Guiding Principles for Scientific Data Management and Stewardship. Scientific Data. 3(1):1-9.
 3. Auer S, Oelen A, Haris M, Stocker M, D'Souza J, Farfar KE, Vogt L, Prinz M, Wiens V, Jaradeh MY. 2020. Improving Access to Scientific Literature with Knowledge Graphs. Bibliothek Forschung und Praxis. 44(3):516-529.
 4. Stocker M, Oelen A, Jaradeh MY, Haris M, Arab Oghli O, Heidari G, Hussein H, Lorenz AL, Kabenamualu S, Farfar KE, Prinz M, Karras O, D'Souza J, Vogt L, Auer S. 2023. FAIR scientific information with the Open Research Knowledge Graph. FAIR Connect. 1(1):19-21. <https://doi.org/10.3233/fc-221513>.
 5. D'Souza J, Hussein H, Evans J, Vogt L, Karras O, Ilangovan V, Lorenz AL, Auer S. 2024. Quality Assessment of Research Comparisons in the Open Research Knowledge Graph: A Case Study. J LIS.it. 15(1):126-143.
 6. Fourrier, C. (2023). 2023, Year of Open LLMs. Hugging Face. Retrieved from <https://huggingface.co/blog/2023-in-llms>

7. Cyganiak, R., Wood, D., & Lanthaler, M. (2014). RDF 1.1 Concepts and Abstract Syntax. World Wide Web Consortium (W3C). Retrieved from <https://www.w3.org/TR/rdf11-concepts/>
8. Hugging Face. (n.d.). Model Cards. Retrieved from <https://huggingface.co/docs/hub/en/model-cards>