استخدام الذكاء الصنعي في البحث العلمي

د. رياض سنبل

المعهد العالي للعلوم التطبيقية والتكنولوجيا

9/11/2025

استخدام الذكاء الصنعي في البحث العلمي

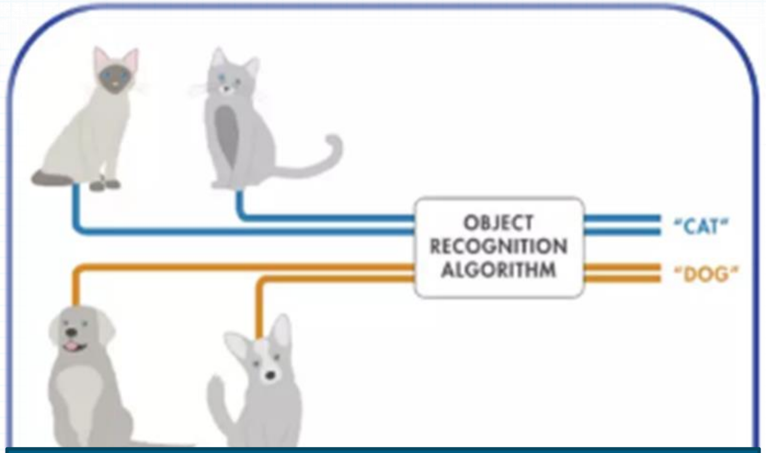
استخدام أدوات الذكاء الصنعي كأدوات مساعدة للمهام المختلفة

استخدام الذكاء الصنعي للوصول لحلول للمسائل البحثية المطروحة

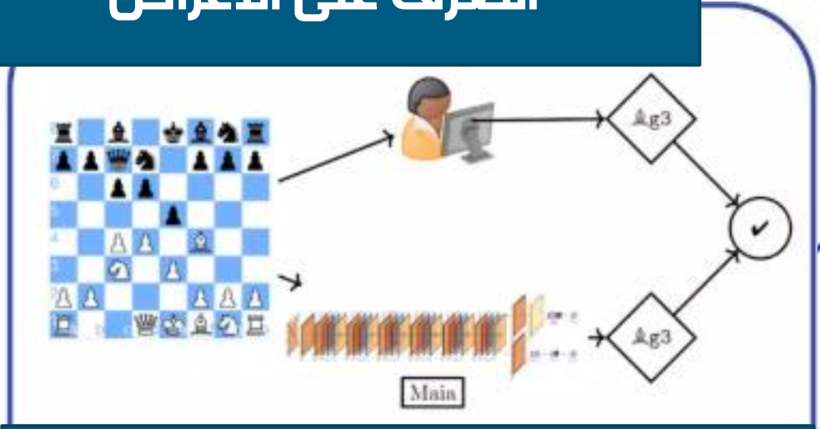
لمحة عامة عن الذكاء الصنعي (لغير المختصين)

ما هو الذكاء الصنعي

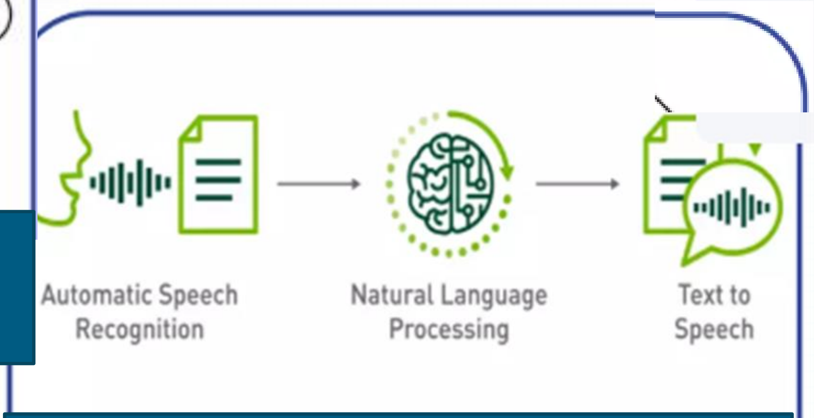
- برمجيات حاسوبية قادرة على أداء أنشطة تشبه السلوك البشري، مثل
  - التعلم من البيانات،
  - الاستدلال واتخاذ القرارات،
  - وتصحيح الأخطاء ذاتياً وتطوير أدائها مع مرور الوقت.
  - ...



التعرف على الأغراض



مهام تخطيط متقدمة ومعقدة



فهم اللغة، التعرف على الكلام، الترجمة، إلخ

ما هو الذكاء الصنعي

الأمن السيبراني

أنظمة التوصية

الروبوتات الذكية

تحليل البيانات واتخاذ القرار

الذكاء الصنعي في التطبيقات المؤسساتية

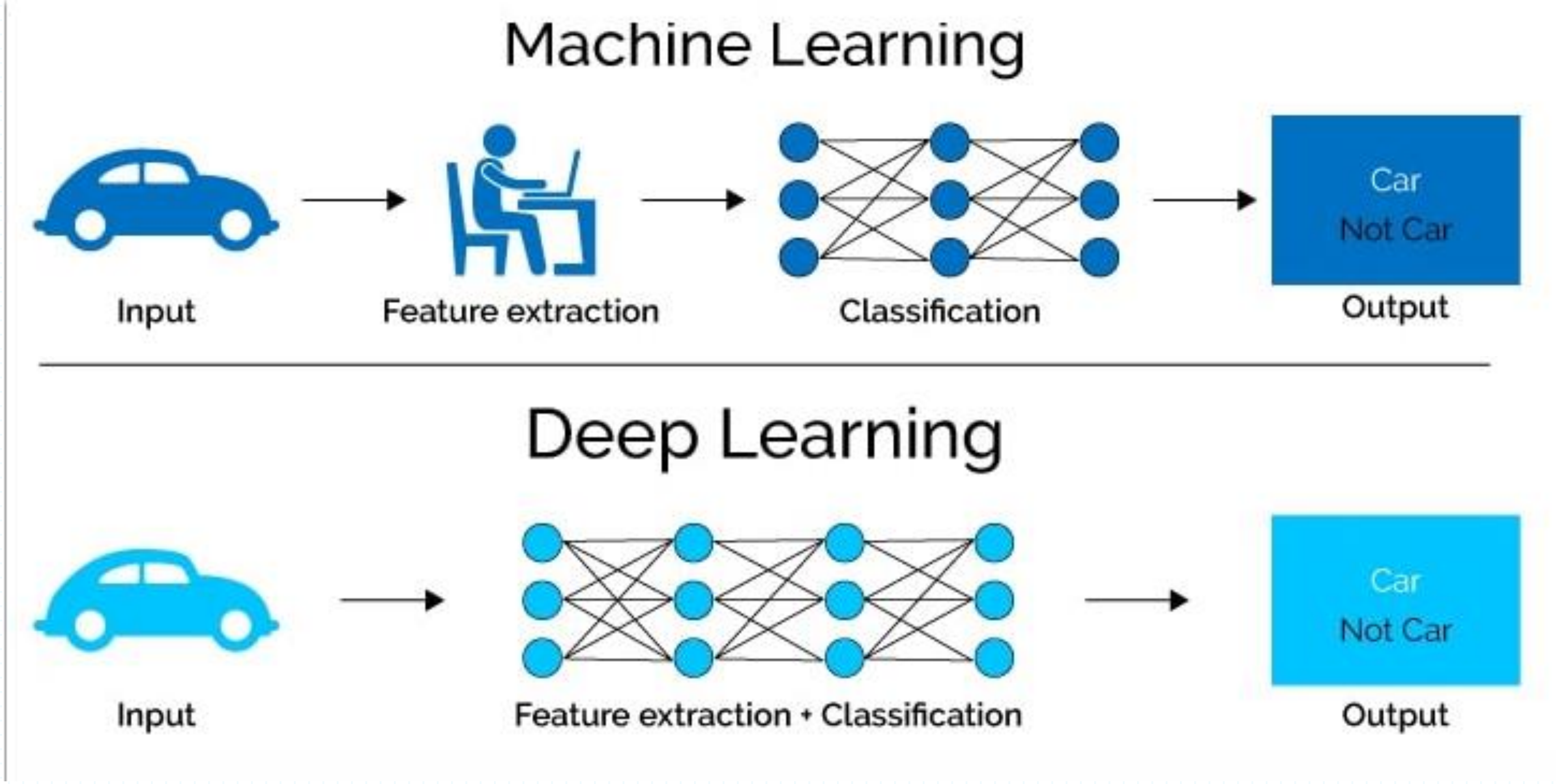
معالجة الكلام
Speech Processing

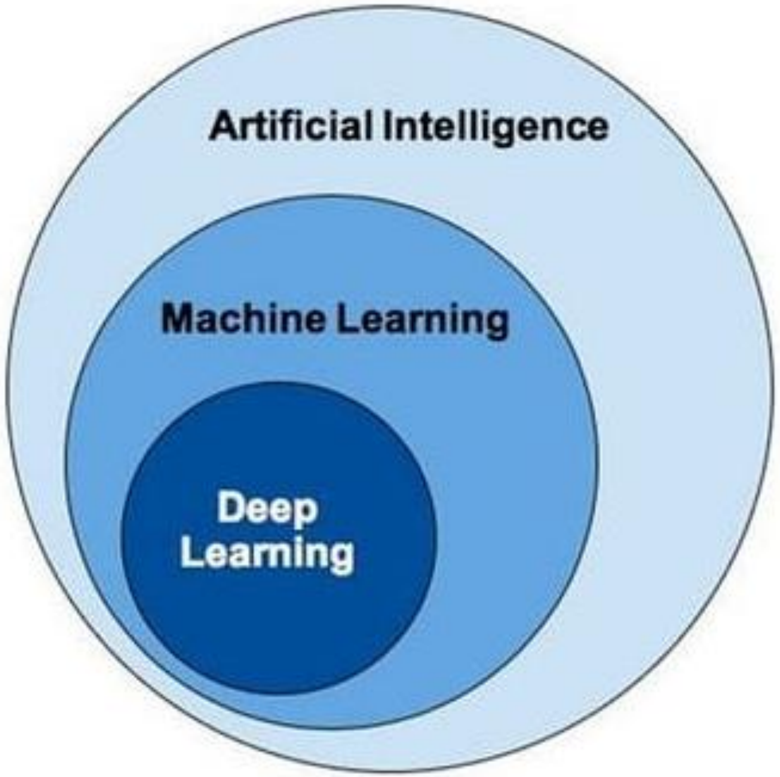
الرؤية الحاسوبية
Computer Vision

معالجة اللغات الطبيعية
Natural Language Processing

تعلم الآلة Machine Learning

تعلم الآلة vs التعلم العميق





كيف يمكن أن يساعد الذكاء الصنعي في البحث العلمي

الذكاء الصنعي ومهام البحث العلمي

- يساعد الذكاء الصنعي على إنجاز المهام بشكل أسرع وأكثر فاعلية.
- يمكن للباحث التركيز بشكل أكبر على الأعمال الأكثر أهمية (التحليل المعمق، التفكير النقدي، سؤال الأسئلة الصحيحة، ابتكار الحلول).
- كباحث.. يجب أن تقود هذه الأدوات بحرفية.. لا أن تسلمها القيادة!

الذكاء الصنعي ومهام البحث العلمي

1. تحديد مشكلة البحث وصياغتها
2. مراجعة الأدبيات
3. تصميم المنهجية أو الحل المقترح
4. جمع البيانات وتنفيذ الاختبارات
5. تفسير النتائج ومناقشتها
6. الكتابة والنشر العلمي

الذكاء الصنعي ومهام البحث العلمي

تحديد مشكلة البحث وصياغتها

مراجعة الأدبيات

تصميم المنهجية أو الحل المقترح

جمع البيانات وتنفيذ الاختبارات

تفسير النتائج ومناقشتها

الكتابة والنشر العلمي

اكتشاف فجوات المعرفة وصياغة أسئلة بحثية دقيقة

استخراج المراجع ذات الصلة بسرعة وتجميعها وتلخيصها وتحليلها بكفاءة

(1) صياغة المشكلة ودراسة الأدبيات

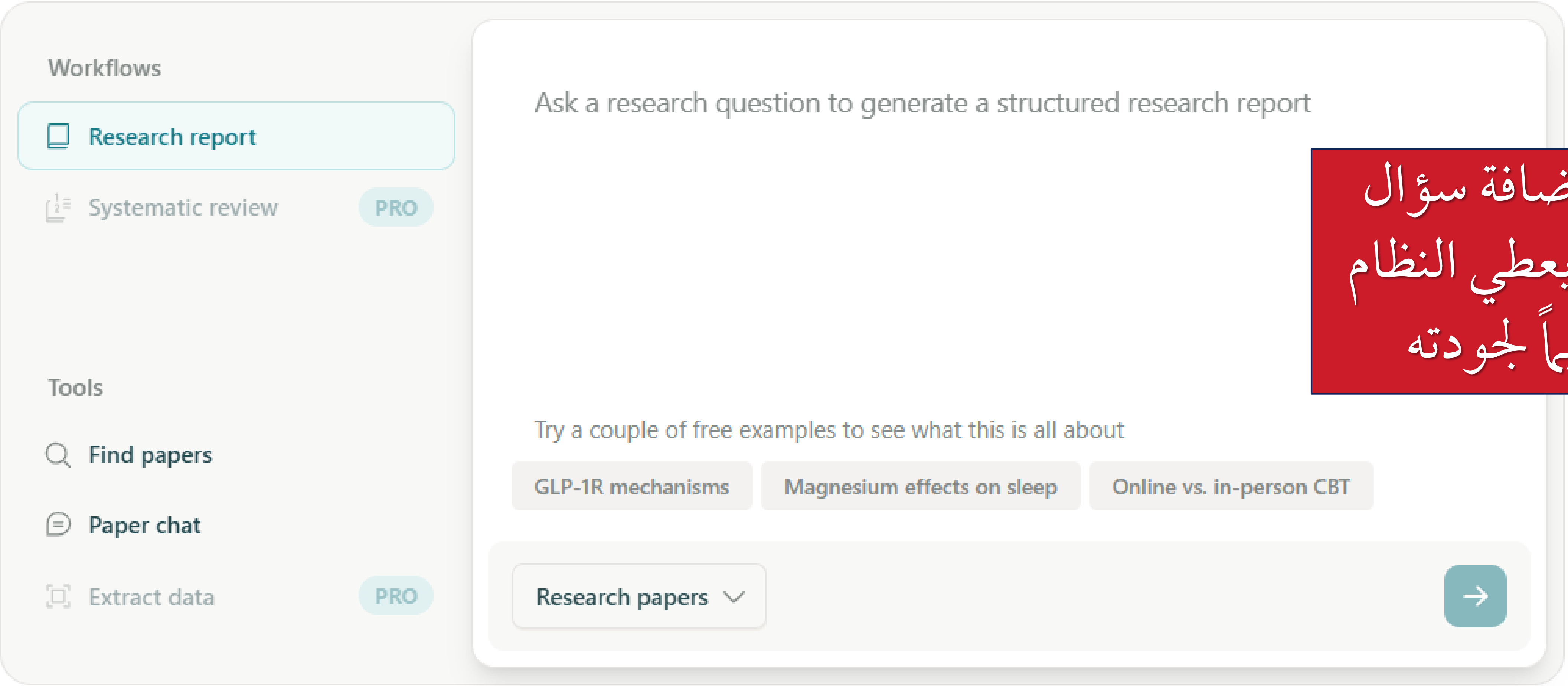
ResearchRabbit™

https://www.researchrabbit.ai/

Elicit

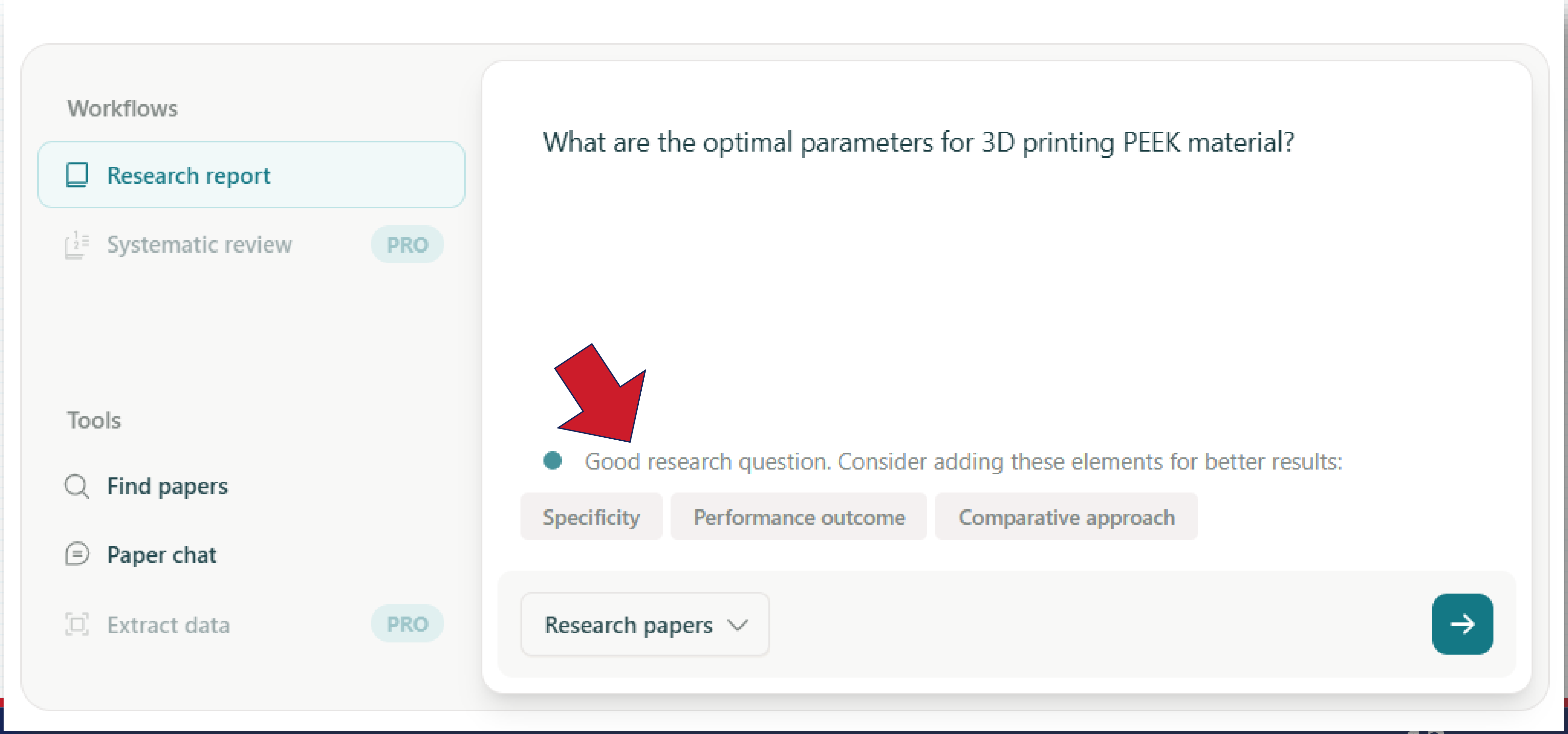
https://elicit.com/

Elicit.com

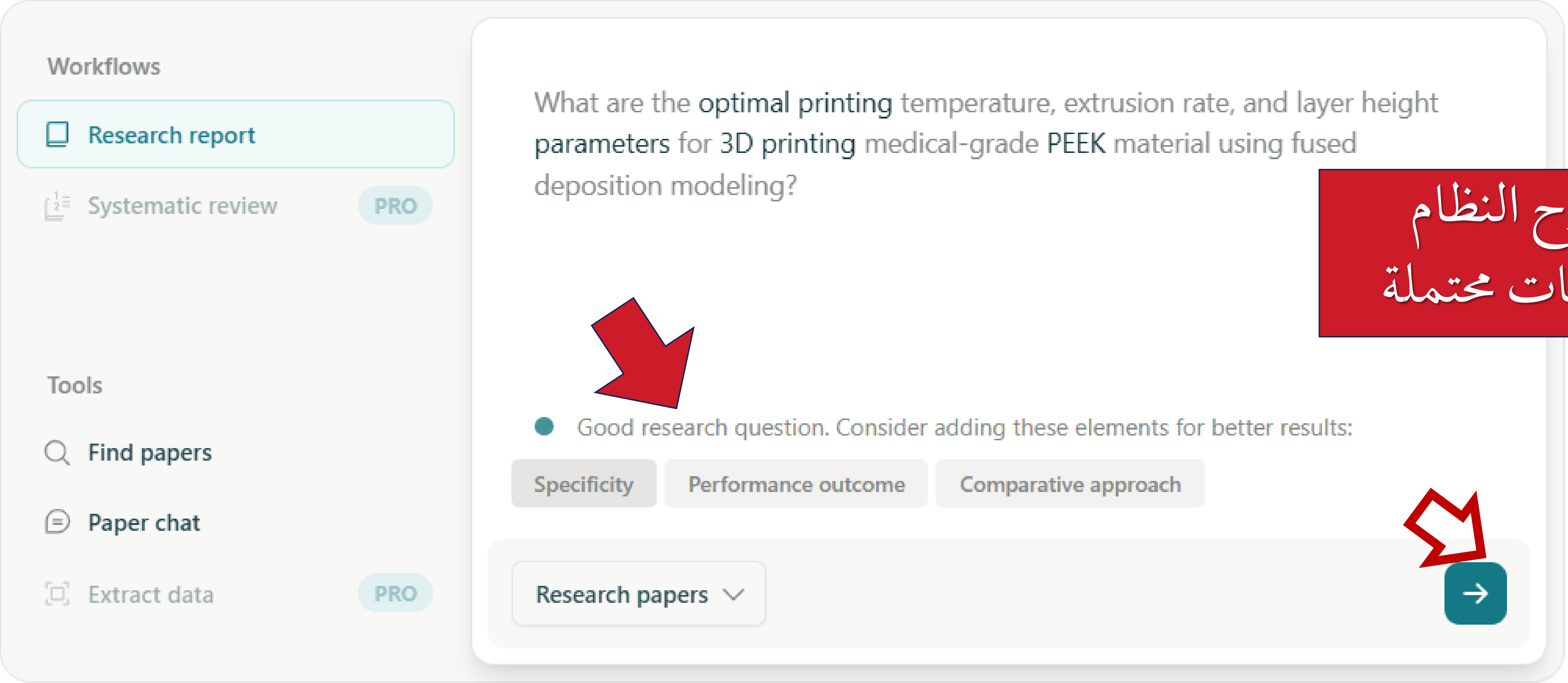


عند إضافة سؤال بحث يعطي النظام تقييماً لجودته

Elicit.com

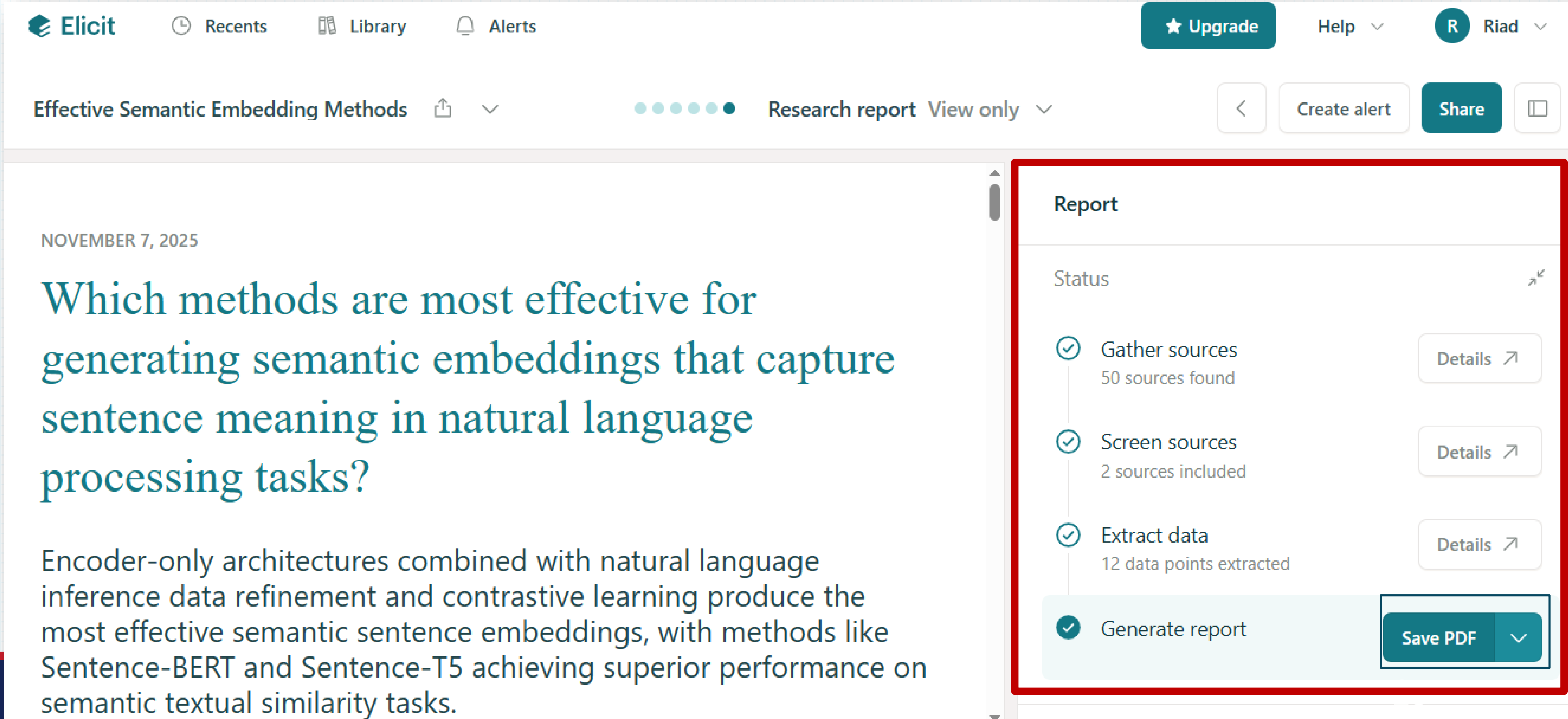

Workflows
Research report
Systematic review
PRO
Tools
Find papers
Paper chat
Extract data
PRO
What are the optimal parameters for 3D printing PEEK material?
Good research question. Consider adding these elements for better results:
Specificity
Performance outcome
Comparative approach
Research papers

Elicit.com



يقترح النظام
تحسينات محتملة

Elicit.com


Elicit
Recents
Library
Alerts
Upgrade
Help
R
Riad
Effective Semantic Embedding Methods
Research report
View only
Create alert
Share
NOVEMBER 7, 2025
Which methods are most effective for generating semantic embeddings that capture sentence meaning in natural language processing tasks?
Encoder-only architectures combined with natural language inference data refinement and contrastive learning produce the most effective semantic sentence embeddings, with methods like Sentence-BERT and Sentence-T5 achieving superior performance on semantic textual similarity tasks.
Report
Status
Gather sources
50 sources found
Details
Screen sources
2 sources included
Details
Extract data
12 data points extracted
Details
Generate report
Save PDF

Elicit.com

Which methods are most effective for generating semantic embeddings that capture sentence meaning in natural language processing tasks?

Encoder-only architectures combined with natural language inference data refinement and contrastive learning produce the most effective semantic sentence embeddings, with methods like Sentence-BERT and Sentence-T5 achieving superior performance on semantic textual similarity tasks.

Abstract

Sentence embeddings that capture meaning most effectively arise from encoder-only architectures refined by natural language inference (NLI) data and augmented with contrastive learning. Reimers and Gurevych (2019) present Sentence-BERT, which uses a siamese/triplet network with pooling to yield fixed-length vectors. Their method scores 11.7 points higher than InferSent and 5.5 points ahead of the Universal Sentence Encoder on semantic textual similarity (STS) benchmarks, while cutting inference time for 10,000 sentences from 65 hours (BERT) to 5 seconds on GPU and training in under 20 minutes.

Ni et al. (2021) detail Sentence-T5, which employs both encoder-only and encoder-decoder variants with a dual encoder and multi-stage contrastive approach. The encoder-only model exceeds previous baselines on SentEval and SentGLUE tasks and, when scaled from 220 million to 11 billion parameters, achieves throughput rates of 274 examples per second on a Cloud TPU-v3. Both approaches also show robust transfer learning performance across tasks such as sentiment analysis, paraphrase detection, and natural language inference.

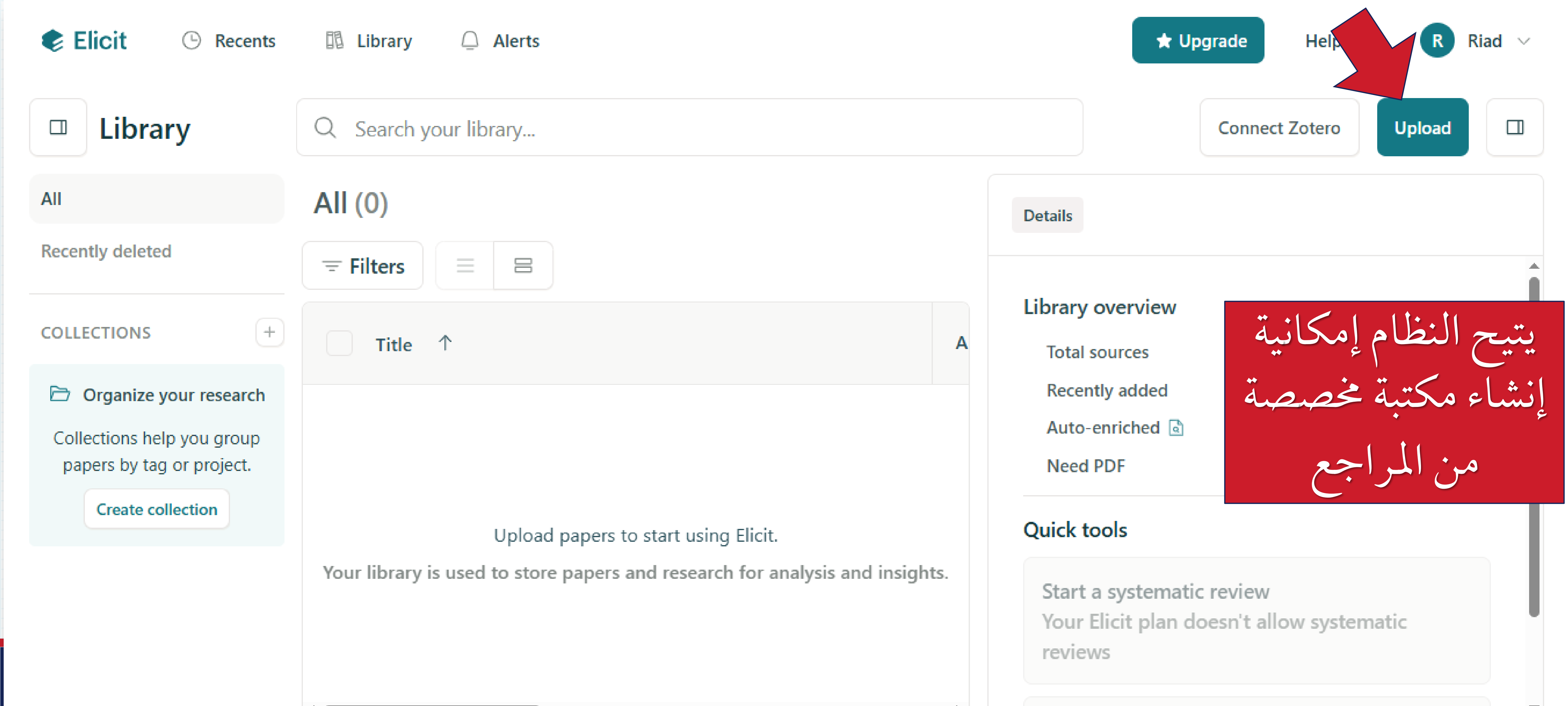
Paper search

We performed a semantic search using the query "Which methods are most effective for generating semantic embeddings that capture sentence meaning in natural language processing tasks?" across over 138 million academic papers from the Elicit search engine, which includes all of Semantic Scholar and OpenAlex.

We retrieved the 50 papers most relevant to the query.

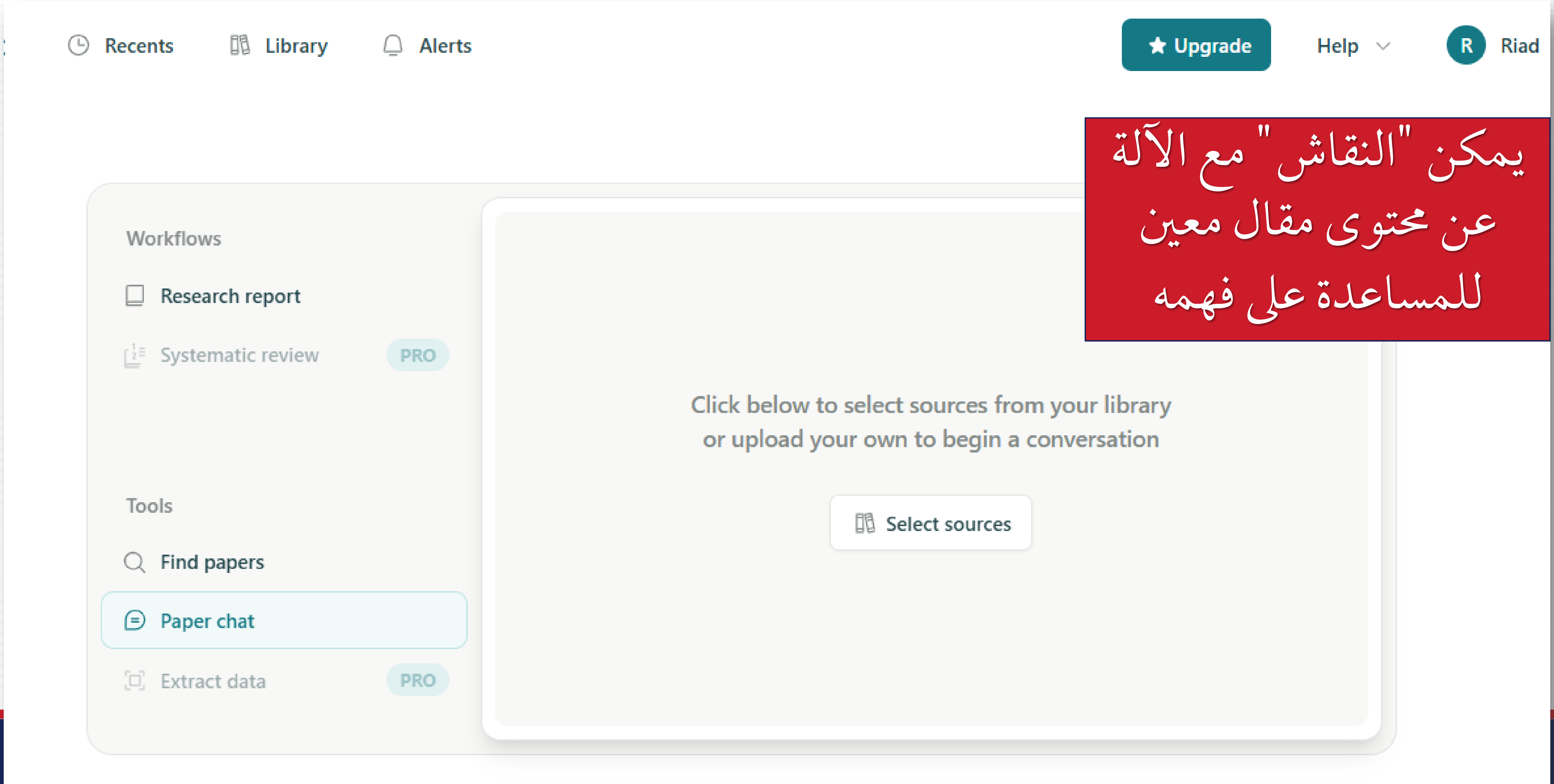
Screening

Elicit.com



يتيح النظام إمكانية إنشاء مكتبة مخصصة من المراجع

Elicit.com



يمكن "النقاش" مع الآلة
عن محتوى مقال معين
للمساعدة على فهمه

Elicit.com

Recents Library Alerts ★ Upgrade Help R

Expert-Driven Reinforcement Learning

Share

يمكن "النقاش" مع الآلة
عن محتوى مقال معين
للمساعدة على فهمه

Chat with papers what is the main contribution for this paper?

what is the main contribution for this paper?

Main Contribution

The main contribution of this paper is the proposal of **Supervised Reinforcement Learning (SRL)**, a novel training framework that addresses the limitations of existing methods for training small-scale language models on multi-step reasoning problems 1 .

Key Innovation

SRL reformulates problem solvi logical "actions" and trains the model to generate an internal mmitting to each action 2 . The

Add new step

(1) صياغة المشكلة ودراسة الأدبيات

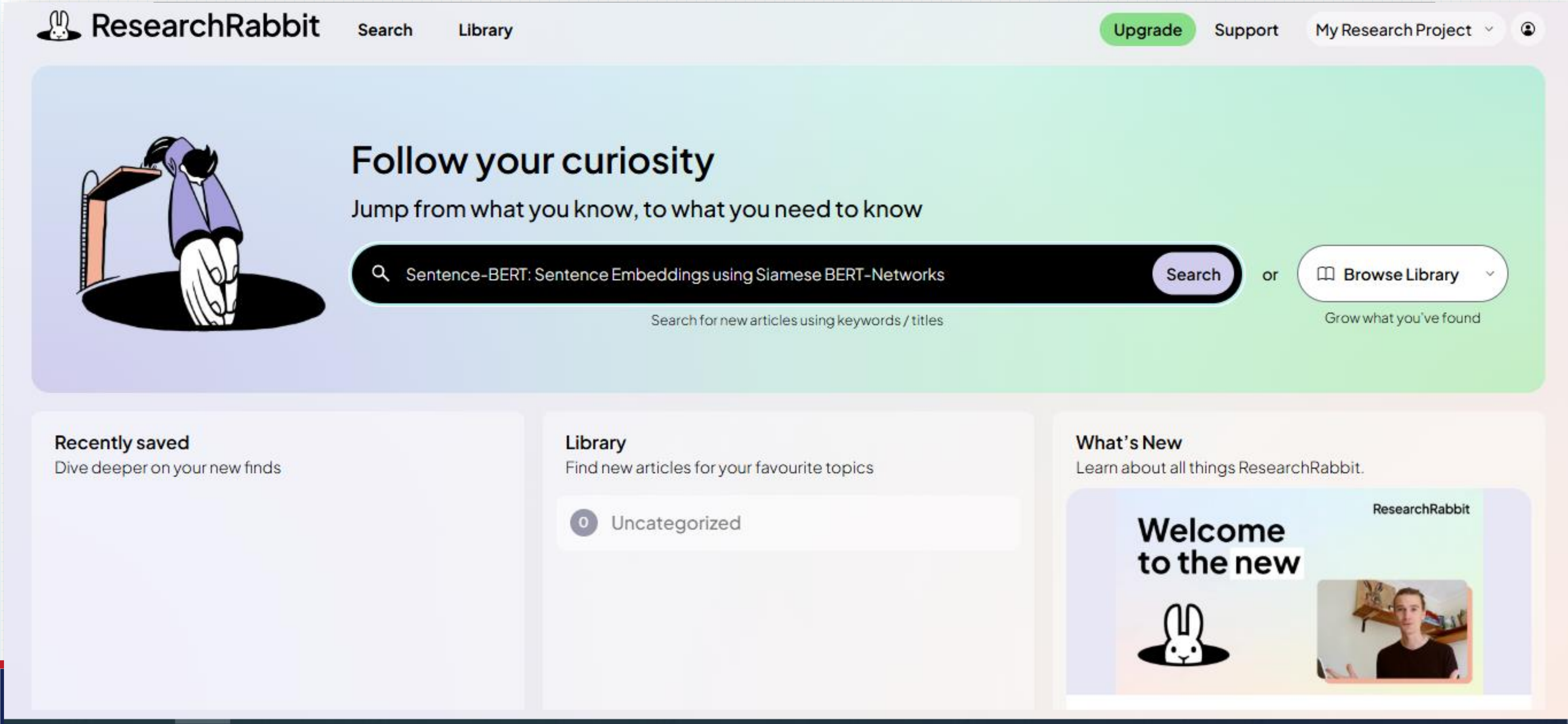
ResearchRabbit™

https://www.researchrabbit.ai/

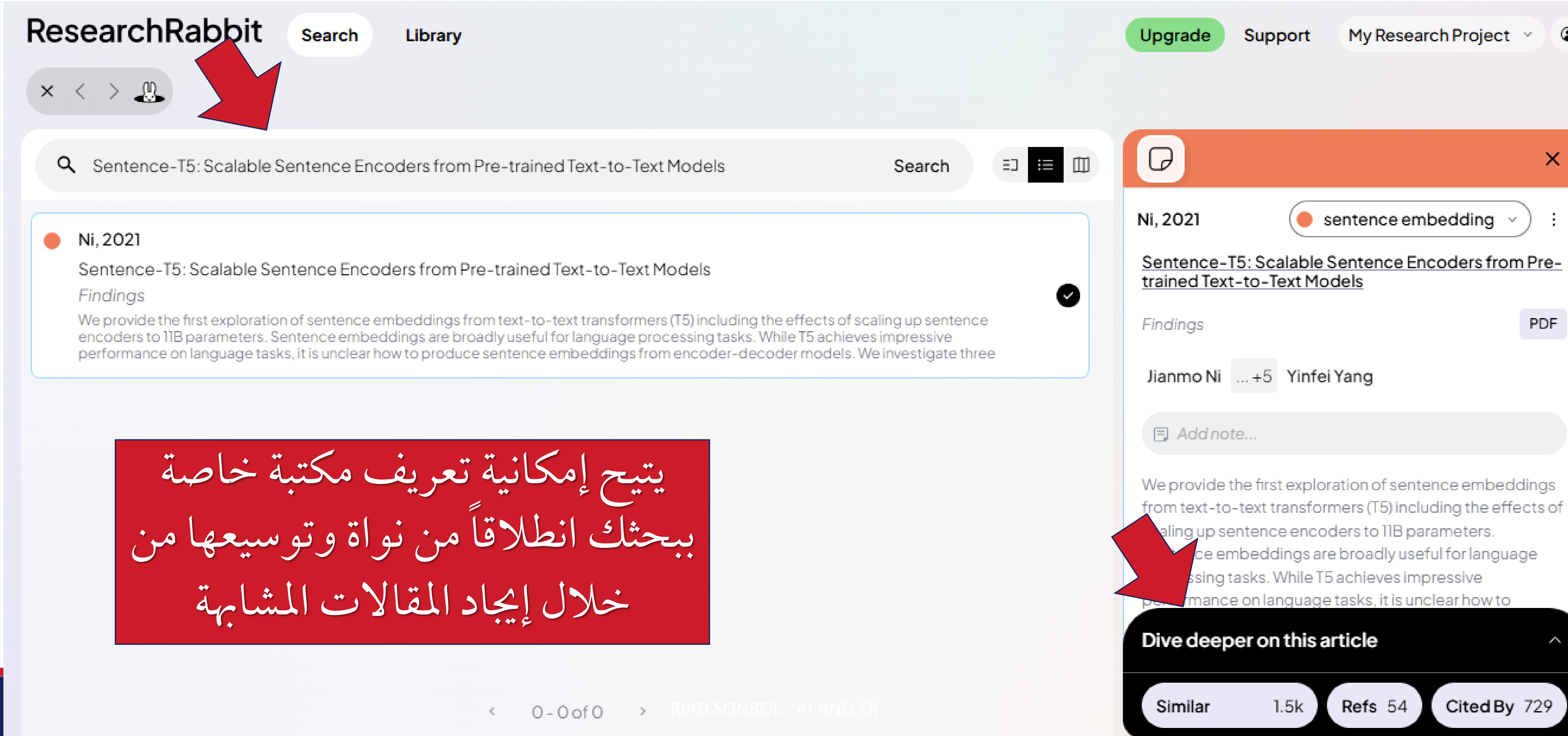
Elicit

https://elicit.com/

ResearchRabbit.ai



ResearchRabbit.ai

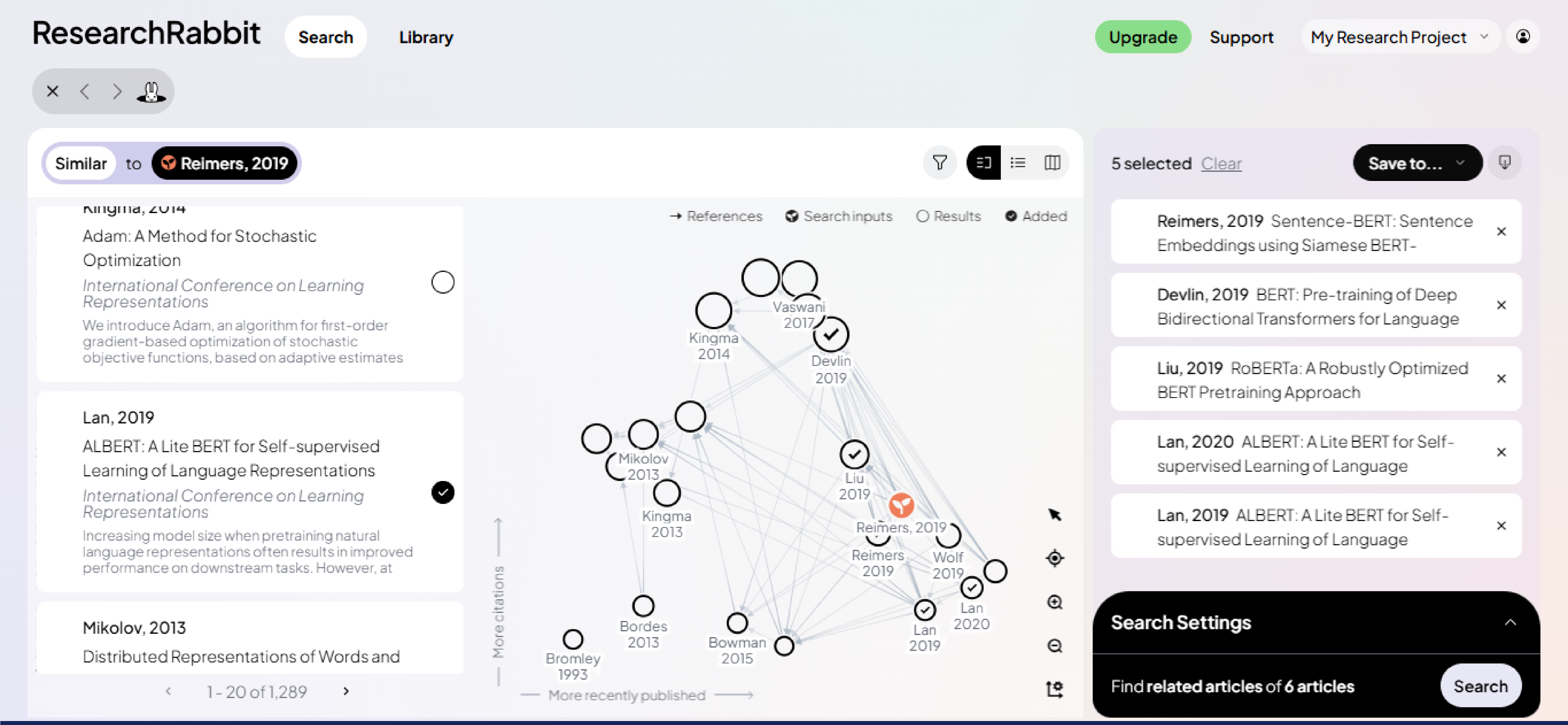


يتيح إمكانية تعريف مكتبة خاصة
ببحثك انطلاقاً من نواة وتوسيعها من
خلال إيجاد المقالات المشابهة

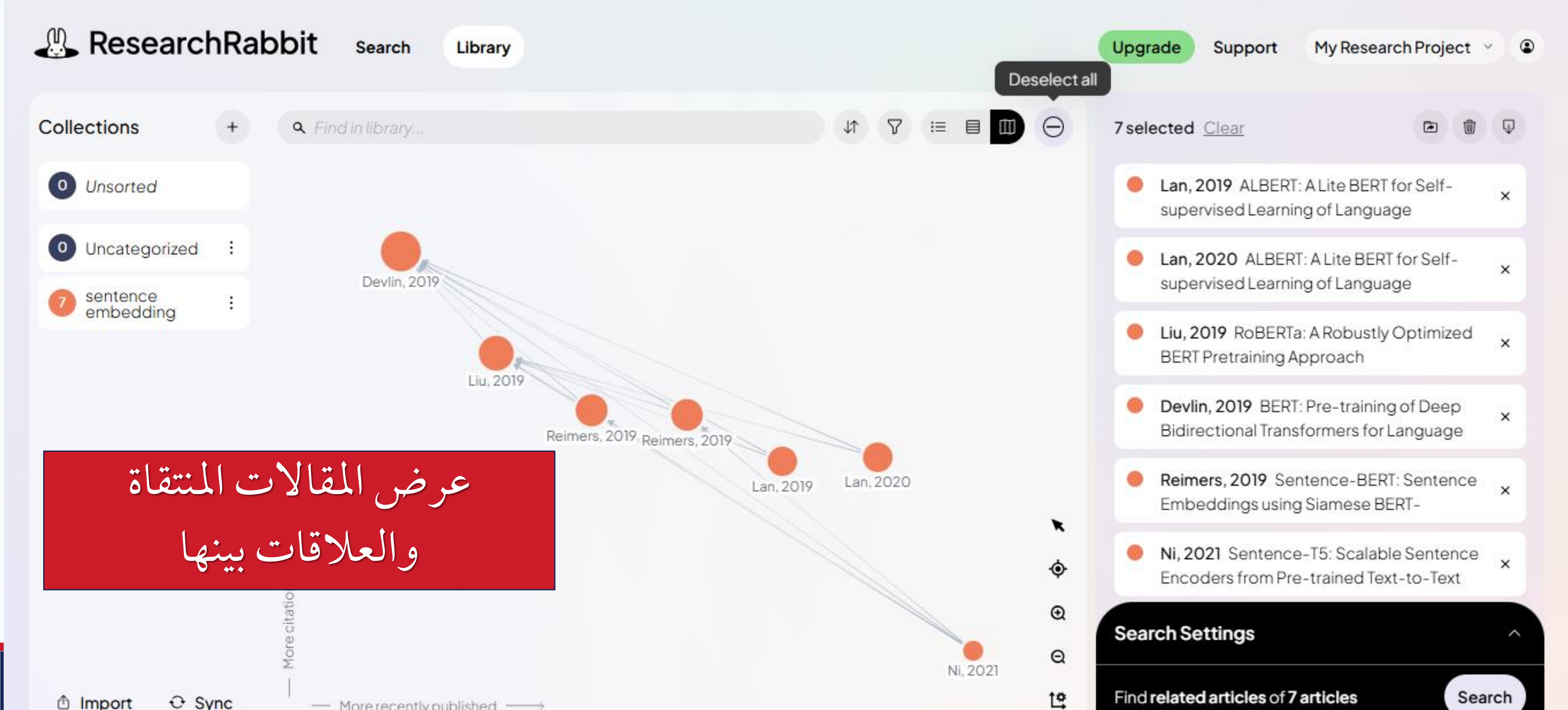
ResearchRabbit.ai

يعرض العلاقات بين المقالات المستنتجة
ويتيح إمكانية إخيار ما يناسب لاضافته

ResearchRabbit.ai

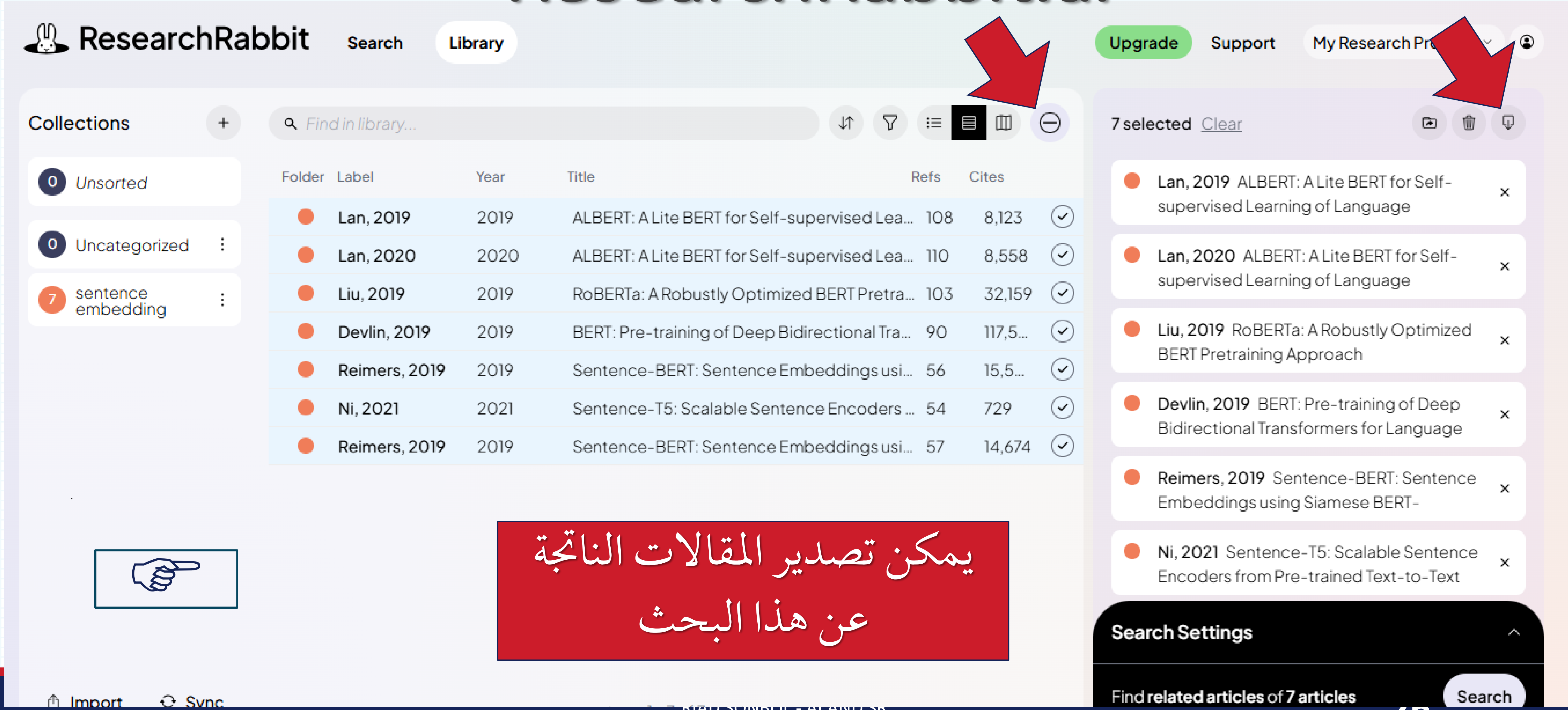


ResearchRabbit.ai



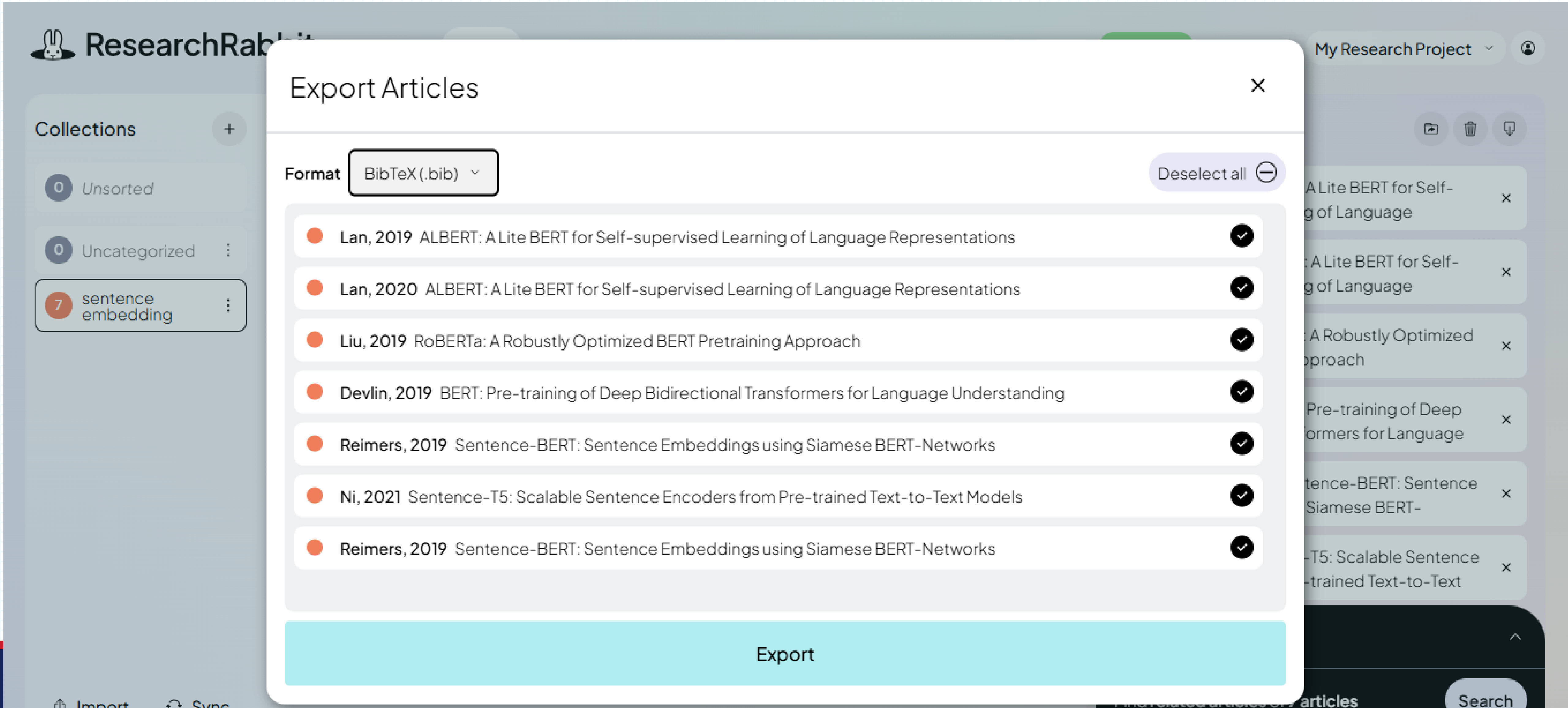
عرض المقالات المنتقاة
والعلاقات بينها

ResearchRabbit.ai



يمكن تصدير المقالات الناتجة
عن هذا البحث

ResearchRabbit.ai



ResearchRabbit.ai

```
% Exported from ResearchRabbit (https://researchrabbitapp.com/)

@article{albert_lan_2019,
    title = {ALBERT: A Lite BERT for Self-supervised Learning of Language Representations},
    author = {Lan, Zhen and Chen, Mingda and Goodman, Sebastian and Gimpel, Kevin and Sharma, Piyush and Soricut, Radu},
    journal = {International Conference on Learning Representations},
    year = {2019},
    researchRabbitId = {7fd55f38-2ef3-45a1-9b4d-6ec4e6ae6e11}
}


@article{albert_lan_2020,
    title = {ALBERT: A Lite BERT for Self-supervised Learning of Language Representations},
    author = {Lan, Zhenzhong and Chen, Mingda and Goodman, Sebastian and Gimpel, Kevin and Sharma, Piyush and Soricut, Radu},
    journal = {arXiv (Cornell University)},
    year = {2020},
    researchRabbitId = {d741e1f5-6690-442d-839a-ea143e8d9785}
}


@article{roberta_liu_2019,
    title = {RoBERTa: A Robustly Optimized BERT Pretraining Approach},
    author = {Liu, Yinhan and Ott, Myle and Goyal, Naman and Du, Jingfei and Joshi, Mandar and Chen, Danqi and Levy, Omer
    and Lewis, Mike and Zettlemoyer, Luke and Stoyanov, Veselin},
    journal = {arXiv: Computation and Language},
    year = {2019},
    researchRabbitId = {afa6dc8b-c0e0-4e81-805c-c92e19e28a8c}
}


@article{bert_devlin_2019,
    title = {BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding},
    doi = {10.18653/V1/N19-1423},
    author = {Devlin, Jacob and Chang, Ming□Wei and Lee, Kenton and Toutanova, Kristina},
    journal = {North American Chapter of the Association for Computational Linguistics},
    year = {2019},
    researchRabbitId = {b66e85cf-145c-4bcf-8753-c6443a87a0f9}
```

الذكاء الصنعي ومهام البحث العلمي

تحديد مشكلة البحث وصياغتها

مراجعة الأدبيات

تصميم المنهجية أو الحل المقترح

جمع البيانات وتنفيذ الاختبارات

تفسير النتائج ومناقشتها

الكتابة والنشر العلمي

اكتشاف فجوات المعرفة وصياغة أسئلة بحثية دقيقة

استخراج المراجع ذات الصلة بسرعة وتجميعها وتلخيصها وتحليلها بكفاءة

المساعدة في الكتابة العلمية باللغة الإنكليزية

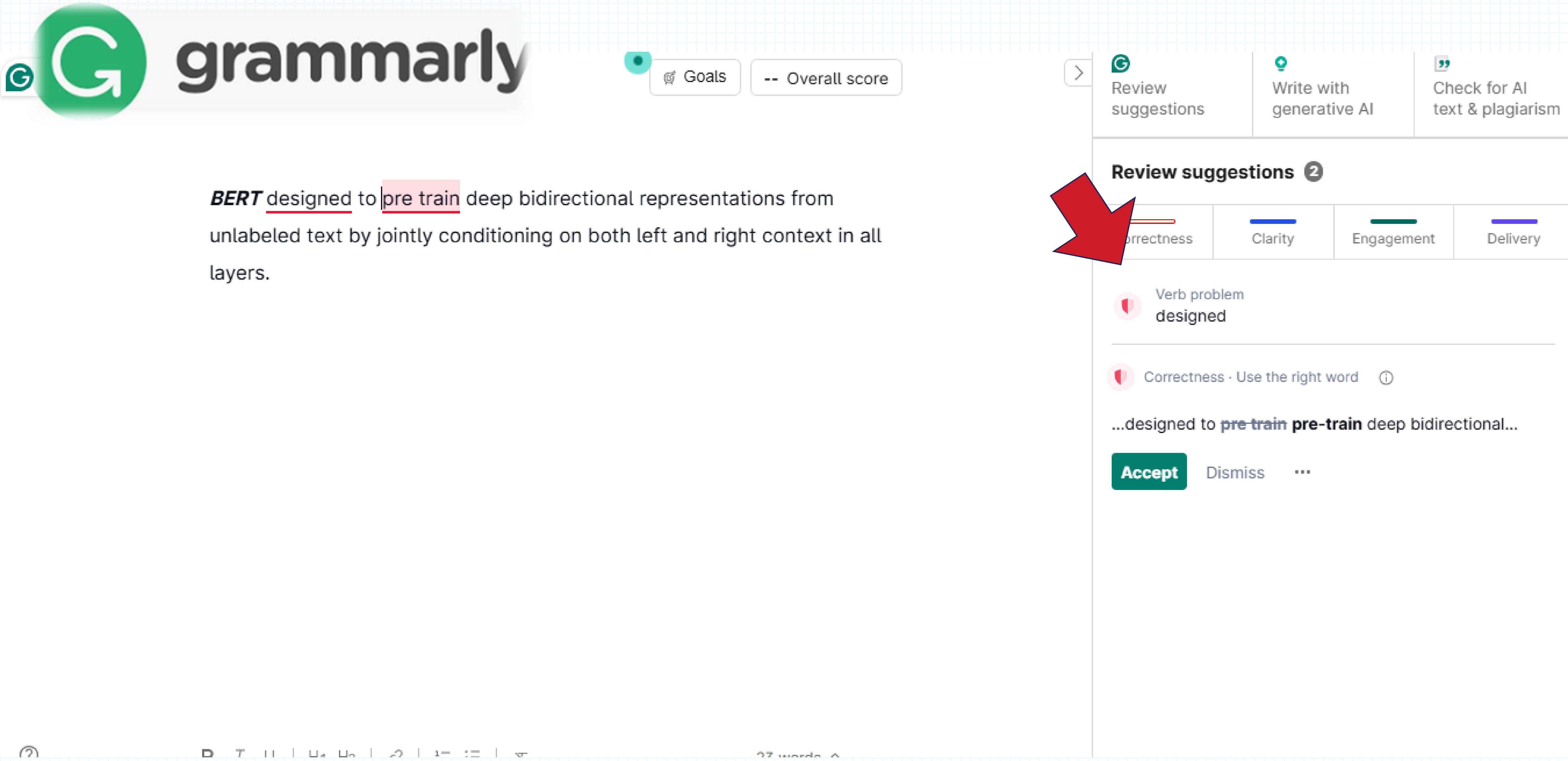
الذكاء الصنعي ومهام البحث العلمي

Paperpal
by Editage

grammarly

TRINKA
by enago

QuillBot



grammarly
Goals
-- Overall score
Review suggestions
Write with generative AI
Check for AI text & plagiarism
BERT designed to pre train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers.
Review suggestions 2
Correctness
Clarity
Engagement
Delivery
Verb problem
designed
Correctness · Use the right word
...designed to ~~pre train~~ **pre-train** deep bidirectional...
Accept
Dismiss

Paperpal
by Editage

Login

Get Started

Grammar Checker

Ensure high-quality academic writing in minutes

Paraphraser

Improve word choice, fluency, and academic tone

Research & Cite

Find & cite references from 250M+ research articles

AI Writing Assistant

Write 2x faster with contextual text suggestions

Chat with PDFs

Extract insights, summaries, and find related papers

Online Translator

Get accurate translation in 50+ languages

Plagiarism Checker

Avoid desk rejection with our trusted similarity check

Submission Checks

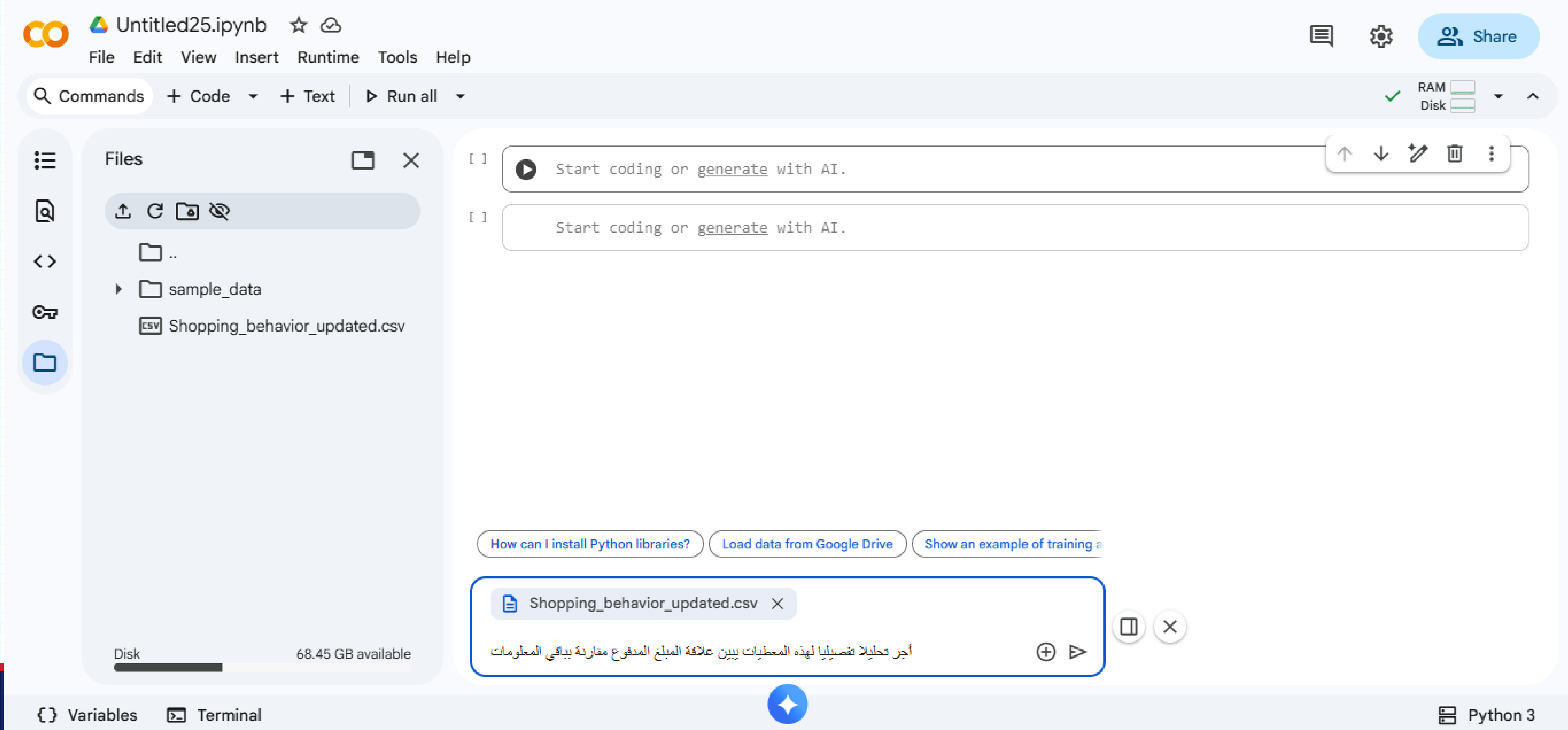
Perfect your manuscript with 30+ essential checks

الذكاء الصنعي... نظرة أعمق

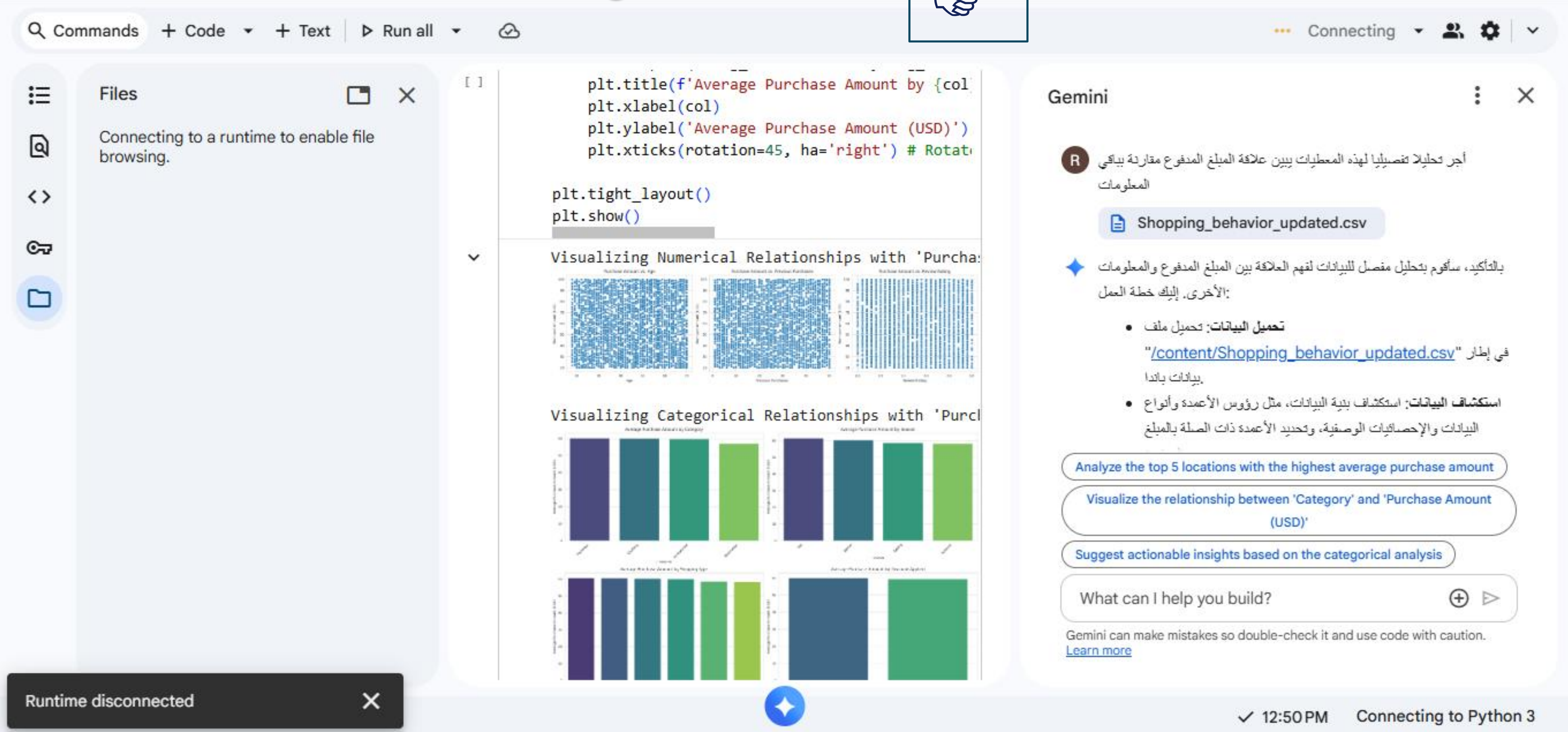
الذكاء الصنعي يساعد على الوصول لحلول متطورة

- **معالجة اللغة الطبيعيةNLP :** تحليل البيانات وتلخيصها.
  - مثال: يستخدم طالب دكتوراه في التاريخ تقنيات معالجة اللغة الطبيعية لتحليل الخطب السياسية، بهدف تتبّع التحولات في الخطاب عبر الزمن.
- **تعلّم الآلة:** تحليل البيانات، التنبؤ، والنمذجة.
  - مثال: يستخدم طالب في علوم البيئة تقنيات التعلّم الآلي للتنبؤ بالتغيرات في كثافة الغابات اعتمادًا على صور الأقمار الصناعية.
- **اكتشاف الأنماط والرؤى Uncovering Insights and Patterns**
  استخدام الذكاء الاصطناعي لاستخلاص العلاقات الخفية واتجاهات البيانات.
- **المكتبات الإحصائية** Statistical Libraries مثل TensorFlow وPyTorch.

Google CoLab with Gemini



Google CoLab with Gemini



Google CoLab with Gemini

Commands | + Code | + Text | Run all | RAM | Disk

Subtask:

تحميل ملف "/content/Shopping_behavior_updated.csv" في إطار بيانات باندا.

Reasoning: First, I'll import the pandas library to enable data manipulation, then load the CSV file '/content/Shopping_behavior_updated.csv' into a DataFrame named 'df', and finally display the first few rows to confirm successful loading.

[1] 0s

```
import pandas as pd

df = pd.read_csv('/content/Shopping_behavior_updated.csv')

print("Data loaded successfully. Here are the first 5 rows:")
df.head()
```

```
Data loaded successfully. Here are the first 5 rows:
```

	Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Seas
0	1	55	Male	Blouse	Clothing	53	Kentucky	L	Gray	Win
1	2	19	Male	Sweater	Clothing	64	Maine	L	Maroon	Win
2	3	50	Male	Jeans	Clothing	73	Massachusetts	S	Maroon	Spri

Gemini

R هل يمكنك إجراء تحليل أعمق للوصول لخلاصات عميقة من هذه المعطيات؟

بالتأكيد، يمكننا إجراء تحليل أعمق للحصول على رؤى أكثر تفصيلاً. إليك خطة عمل مقترحة تركز على استكشاف التوزيعات والتفاعلات بين المتغيرات:

- **تحديد الميزات الفئوية الرئيسية للتحليل المتعمق**: مراجعة نتائج التحليل الفئوي السابق لتحديد الميزات التي أظهرت اختلافات ملحوظة في متوسط 'مبلغ الشراء (بالدولار الأمريكي)'. ستكون هذه الميزات هي محور تحليل التوزيع.
- **تصور توزيع مبلغ الشراء للميزات الفئوية الرئيسية**: إنشاء رسوم بيانية صندوقية (Box plots) أو رسوم بيانية كمانية (Violin plots) لتصور توزيع 'مبلغ الشراء (بالدولار الأمريكي)' للميزات الفئوية الرئيسية. سيظهر هذا ليس فقط المتوسط ولكن أيضًا الانتشار، الأرباع، والمخرجات المحتملة ضمن كل فئة، مما يوفر فهمًا أعمق من مجرد المتوسطات.
- **تحليل تأثيرات التفاعل بين المتغيرات الفئوية**: التحقيق في تأثيرات التفاعل المحتملة بين المتغيرات الفئوية الرئيسية. على سبيل المثال، تحليل

What can I help you build?

Gemini can make mistakes so double-check it and use code with caution. Learn more

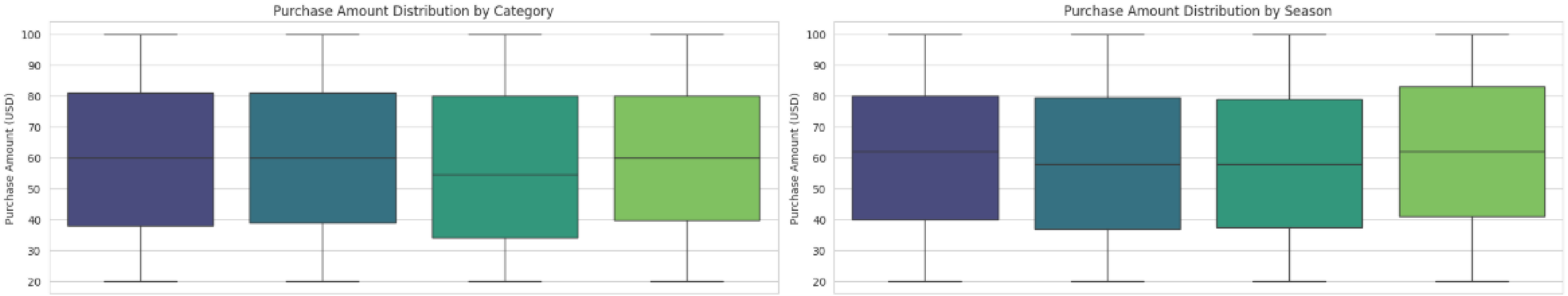
{} Variables | Terminal | Executing (5s) | Python 3

Google CoLab with Gemini

Gemini Example.ipynb

File Edit View Insert Runtime Tools Help

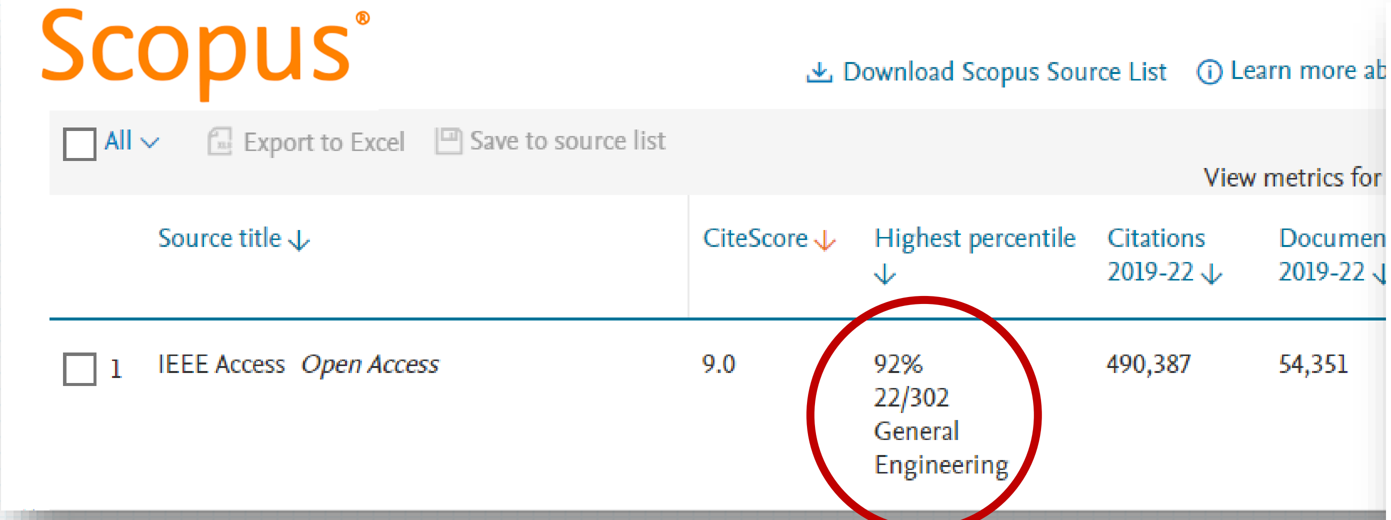
```
num_cols = 2

for i, col in enumerate(key_categorical_features):
    plt.subplot(num_rows, num_cols, i + 1) # Create subplot
    sns.boxplot(x=col, y='Purchase Amount (USD)', data=df, palette='viridis', hue=col, legend=False)
    plt.title(f'Purchase Amount Distribution by {col}')
    plt.xlabel(col)
    plt.ylabel('Purchase Amount (USD)')
    plt.xticks(rotation=45, ha='right') # Rotate x-axis labels for readability

plt.tight_layout() # Adjust layout to prevent overlapping
plt.show()
```



الذكاء الصنعي... المسؤولية الأكاديمية

المسؤولية تقع على الباحث

- يجب على الباحث التحقق من كل معلومة أو ادعاء أو مرجع قبل اعتمادها.
  - Published vs preprint
  - Scopus Ranking



المسؤولية تقع على الباحث

- لا تضع أبدًا بيانات حساسة أو سرّية أو غير منشورة في أدوات الذكاء الاصطناعي العامة (مثل النسخ المجانية من ChatGPT)
- فقد تُستخدم البيانات المُحمّلة لأغراض تدريب النماذج.

AI handles the "how" (data processing, analysis), but you provide the "why" (the research question, the interpretation).

استخدام الذكاء الصنعي في البحث العلمي

د. رياض سنبل
المعهد العالي للعلوم التطبيقية والتكنولوجيا