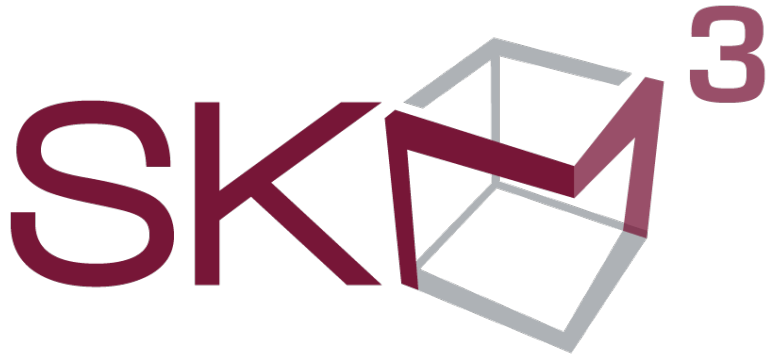


AI for Advancing Scientific Research

Angelo Salatino

KMi, The Open University, UK

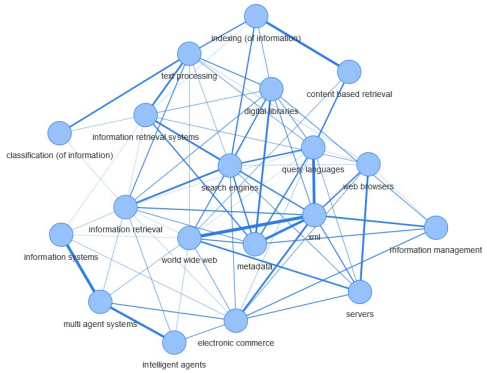
𝕏 **@angelosalatino**

Knowledge Media Institute

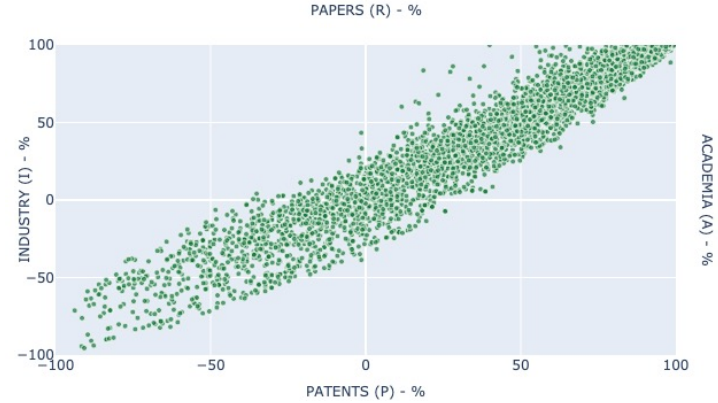
The Open University

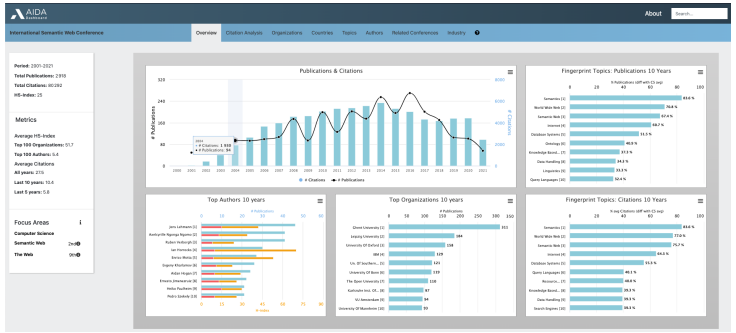
Scholarly Knowledge Modelling (SKM) Team

- SKM develops **innovative AI solution** leveraging large-scale data mining, semantic technologies, machine learning, and visual analytics for making sense of scholarly data and forecasting research dynamics.



- The SKM team has an extensive list of collaborators, which include (but are not limited to):
  - Leading European **universities** (e.g., FIZ Karlsruhe, Bologna, Cagliari, Paris Sorbonne, Seville, Milano Bicocca, Oxford)
  - **Intergovernmental organizations**, such as OECD
  - **Research institutes** (FBK and GESIS)
  - **Commercial organizations**, including Digital Science, Linkalab and one of the top international scientific publishers: Springer Nature.



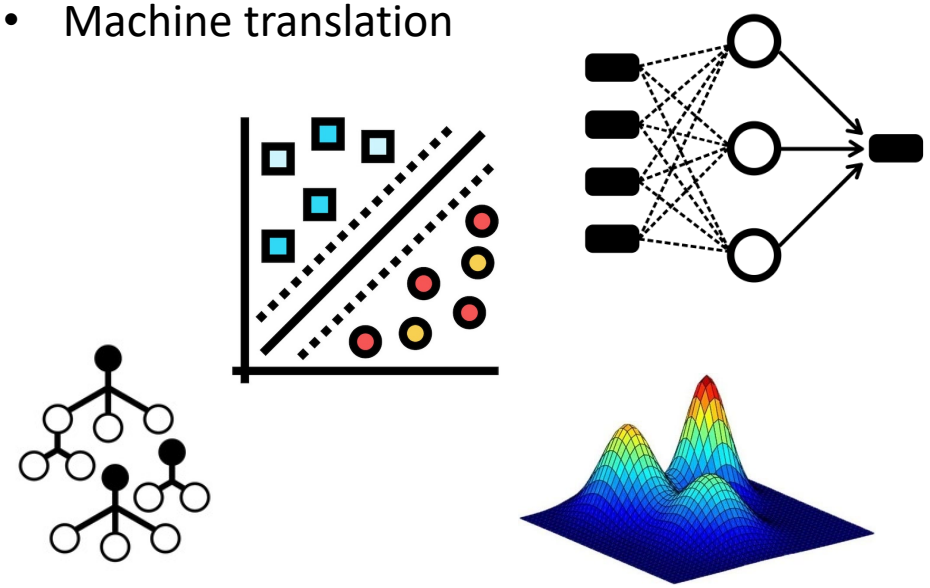


How has AI supported the Research Ecosystem so far?

Past decades

AI has been supporting research with Machine Learning and Deep Learning models:

- Product recommendations
- Medical diagnosis
- Fraud detection
- Face recognition
- Sentiment Analysis
- Machine translation



Past few years

AI has been supporting research with co-pilots:

- Question Answering
- Summarising, proofreading
- Literature reviews
- Code development
- Drafting content

ChatGPT

Microsoft Copilot

LLaMA by Meta

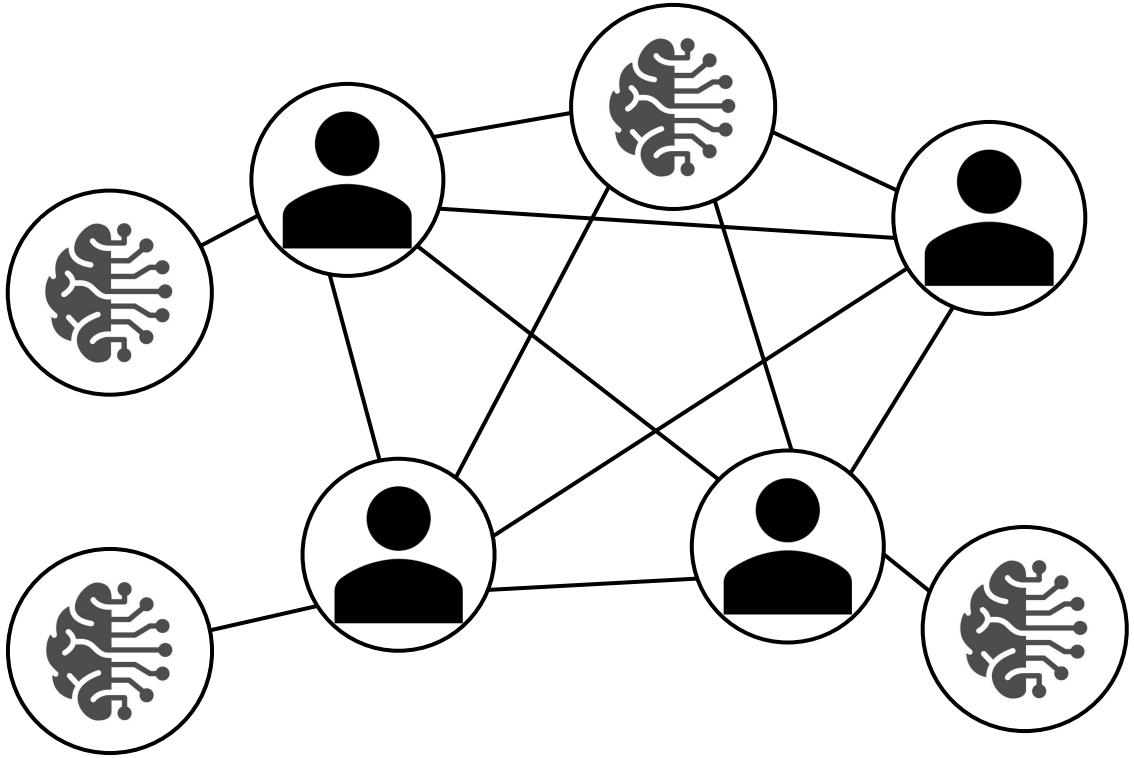
Gemini

Vision: Extending Human Performance in Research

- **Next decade?**

AI as Smart Personal Research Assistant:

- Daily digests on recently published literature tailored to researcher's interest
- High-quality literature reviews
- Hypothesis generation and prioritisation
- Packaging data for Open Science
- Drafting impact cases
- Assisting peer review



Agenda

Automatic Literature Review Generation

Hypothesis Generation

Diversity of Expertise

Scientific Question Answering

Agenda

Automatic Literature Review Generation

Hypothesis Generation

Diversity of Expertise

Scientific Question Answering

Automatic Generation of Literature Review

Are there AI tools that can generate a draft literature review based on a research question/topic?

What are the main research gaps in the field of explainable AI?

Explainable AI (XAI) has emerged as a vital research domain aimed at enhancing the transparency and interpretability of artificial intelligence systems (Adadi & Berrada, 2018). While substantial progress has been made …

Bolanos et al. (2023) Artificial Intelligence for Literature Reviews: Opportunities and Challenges. https://arxiv.org/abs/2402.08565

Automatic Generation of Literature Review

21 leading Systematic Literature Review (SLR) tools

Abstrackr, Colandr, Covidence, EPPI-Reviewer, Rayyan, SWIFT-Active Screener, DistillerSR, SWIFT-Review, SysRev, Nested Knowledge, RobotAnalyst, LitSuggest, Pitts.ai, Research Screener, ASReview, PICOPortal, FAST2, Dextr, ExaCT, **Iris.ai**, **RobotReviewer/RobotSearch**

11 LLM-based tools aiding LR

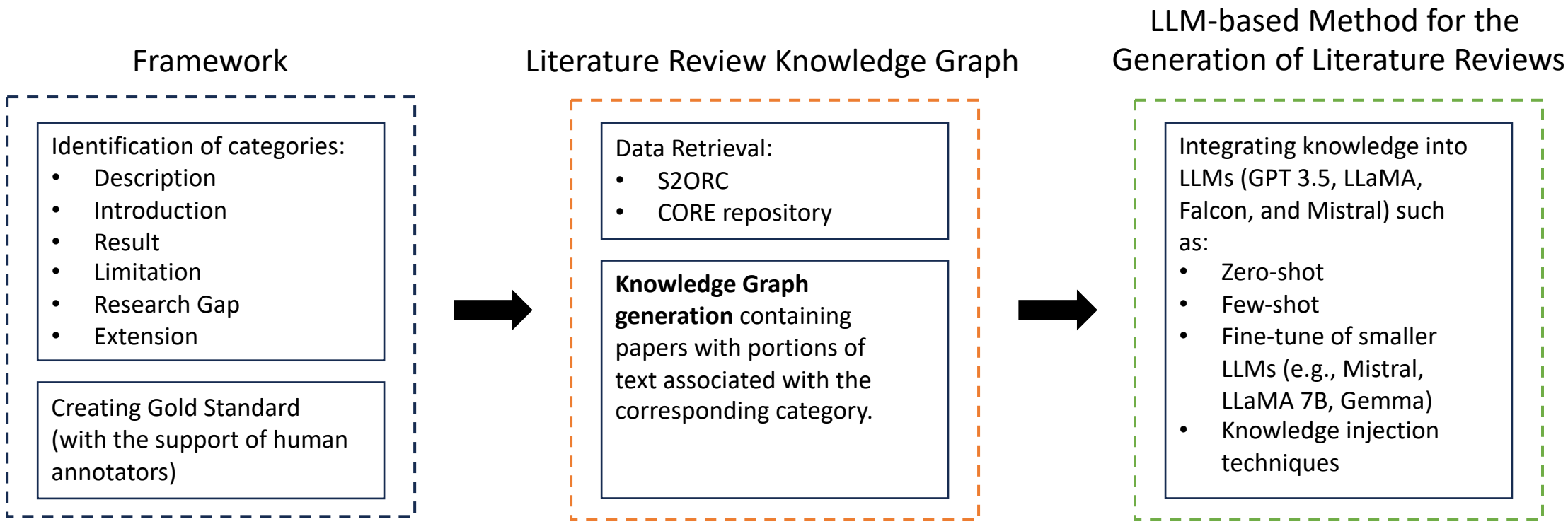
Scite, Elicit, Consensus, EvidenceHunt, MirrorThink, Perplexity, Scispace, Jenni.ai, ResearchBuddies, Silatus, Textero.ai

The generated literature reviews lack the structure and presentation style found in literature reviews written by researchers.

Bolanos et al. (2024) Artificial Intelligence for Literature Reviews: Opportunities and Challenges. https://arxiv.org/abs/2402.08565

Automatic Generation of Literature Review

We are designing a new **AI framework** for **automatically generating literature reviews** that integrates information from multiple sources and exploits an advanced characterization of scientific concepts and research questions.



Agenda

Automatic Literature Review Generation

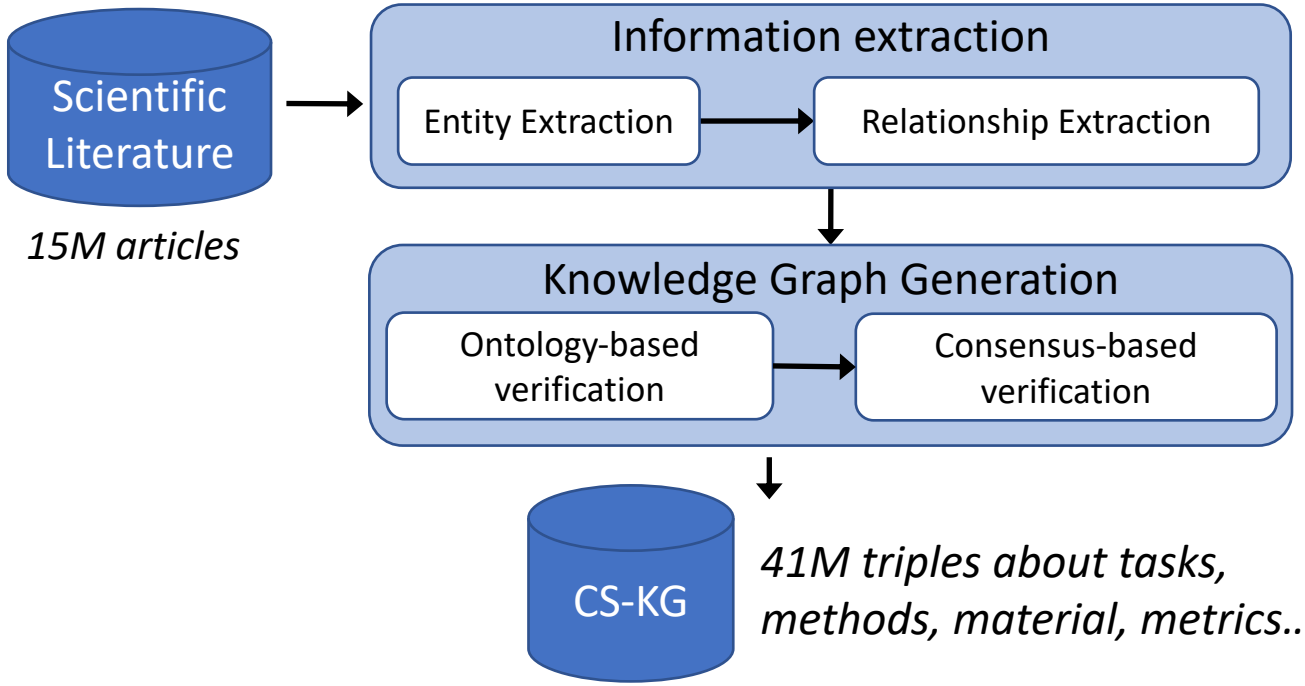
Hypothesis Generation

Diversity of Expertise

Scientific Question Answering

AI System for Hypothesis Generation - Architecture

A novel approach which exploits **link prediction models, impact features**, and **large language models** to formulate original research hypotheses and relevant research plans.



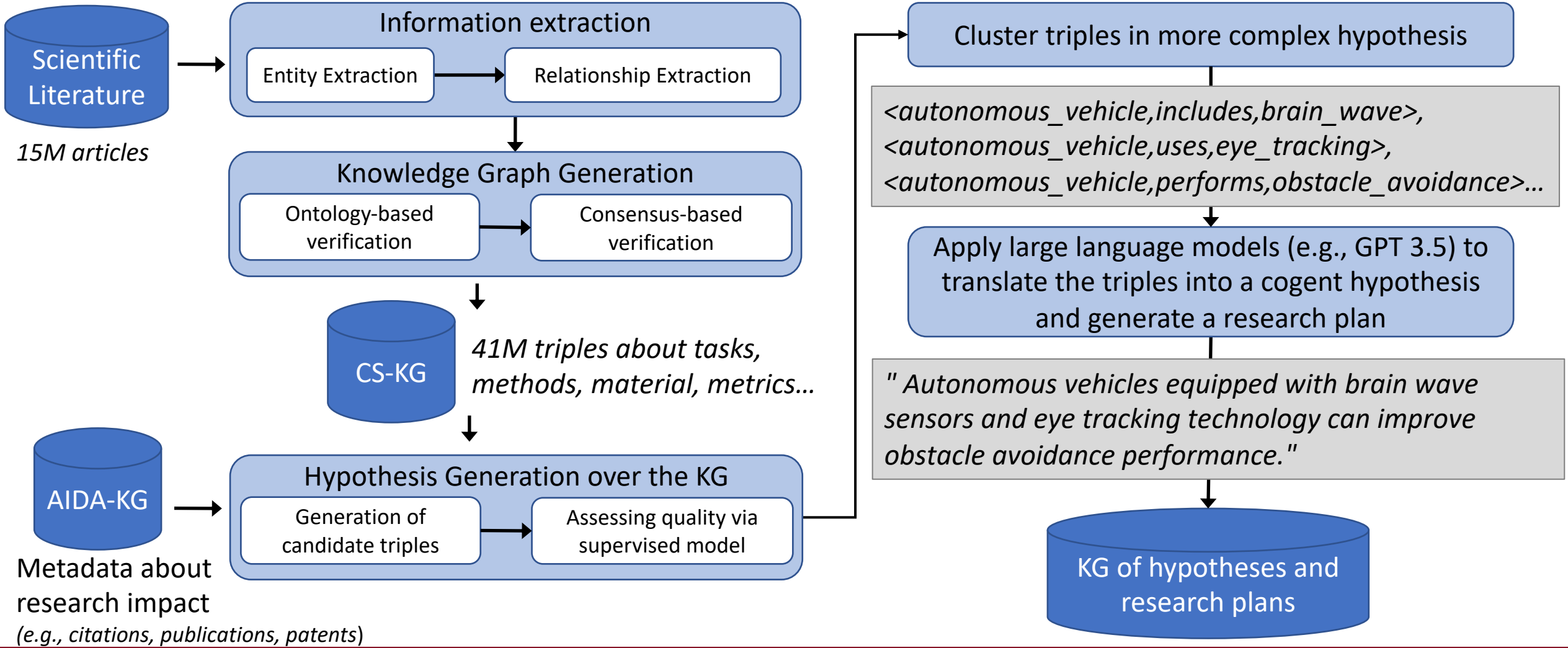
AIDA-KG

Metadata about research impact
(e.g., citations, publications, patents)

Borrego et al. (2024) Research Hypothesis Generation over Scientific Knowledge Graphs. Submitted to Knowledge-Based System

AI System for Hypothesis Generation - Architecture

A novel approach which exploits **link prediction models, impact features**, and **large language models** to formulate original research hypotheses and relevant research plans.



Hypothesis Generation – An Example

Hypothesis: Autonomous vehicles equipped with brain wave sensors and eye-tracking technology can improve obstacle avoidance performance.

Discussion: Autonomous vehicles are becoming increasingly popular, and their safety is a crucial concern. Obstacle avoidance is a critical task for autonomous vehicles, and it requires the integration of various technologies. One possible solution is to equip autonomous vehicles with brain wave sensors and eye-tracking technology. Brain wave sensors can detect the driver's cognitive state and predict their intentions, while eye tracking technology can monitor the driver's gaze and attention. By combining these technologies, it is possible to improve obstacle avoidance performance and enhance the safety of autonomous vehicles.

Agenda

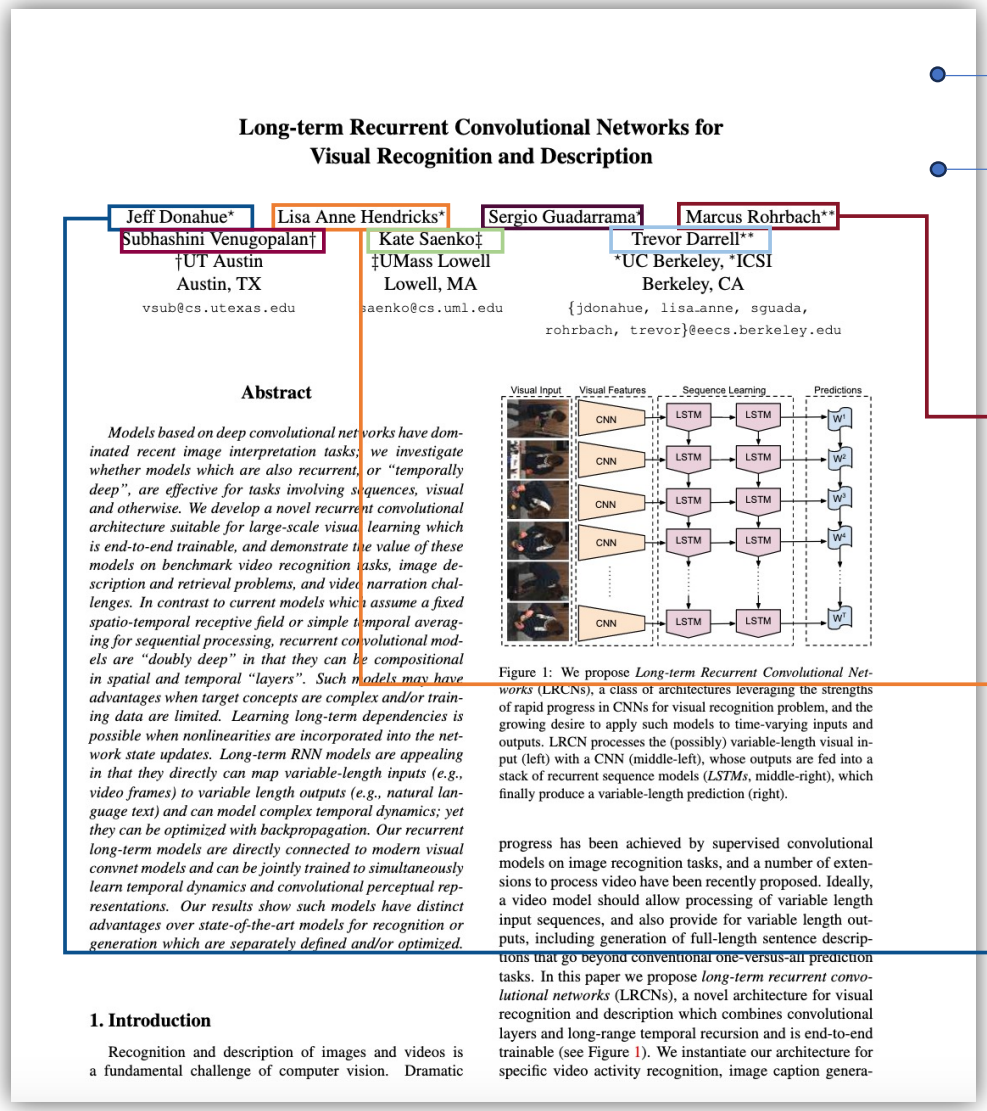
Automatic Literature Review Generation

Hypothesis Generation

Diversity of Expertise

Scientific Question Answering

Our approach: a glimpse



Long-term Recurrent Convolutional Networks for Visual Recognition and Description

Jeff Donahue* Lisa Anne Hendricks* Sergio Guadarrama† Marcus Rohrbach**
Subhashini Venugopalan† Kate Saenko‡ Trevor Darrell**
†UT Austin ‡UMass Lowell *UC Berkeley, *ICSI
Austin, TX Lowell, MA Berkeley, CA
vsub@cs.utexas.edu saenko@cs.uml.edu {jdonahue, lisa_anne, sguada, rohrbach, trevor}@eecs.berkeley.edu

Abstract

Models based on deep convolutional networks have dominated recent image interpretation tasks; we investigate whether models which are also recurrent, or "temporally deep", are effective for tasks involving sequences, visual and otherwise. We develop a novel recurrent convolutional architecture suitable for large-scale visual learning which is end-to-end trainable, and demonstrate the value of these models on benchmark video recognition tasks, image description and retrieval problems, and video narration challenges. In contrast to current models which assume a fixed spatio-temporal receptive field or simple temporal averaging for sequential processing, recurrent convolutional models are "doubly deep" in that they can be compositional in spatial and temporal "layers". Such models may have advantages when target concepts are complex and/or training data are limited. Learning long-term dependencies is possible when nonlinearities are incorporated into the network state updates. Long-term RNN models are appealing in that they directly can map variable-length inputs (e.g., video frames) to variable length outputs (e.g., natural language text) and can model complex temporal dynamics; yet they can be optimized with backpropagation. Our recurrent long-term models are directly connected to modern visual convnet models and can be jointly trained to simultaneously learn temporal dynamics and convolutional perceptual representations. Our results show such models have distinct advantages over state-of-the-art models for recognition or generation which are separately defined and/or optimized.

Figure 1: We propose *Long-term Recurrent Convolutional Networks* (LRCNs), a class of architectures leveraging the strengths of rapid progress in CNNs for visual recognition problem, and the growing desire to apply such models to time-varying inputs and outputs. LRCN processes the (possibly) variable-length visual input (left) with a CNN (middle-left), whose outputs are fed into a stack of recurrent sequence models (*LSTMs*, middle-right), which finally produce a variable-length prediction (right).

1. Introduction

Recognition and description of images and videos is a fundamental challenge of computer vision. Dramatic progress has been achieved by supervised convolutional models on image recognition tasks, and a number of extensions to process video have been recently proposed. Ideally, a video model should allow processing of variable length input sequences, and also provide for variable length outputs, including generation of full-length sentence descriptions that go beyond conventional one-versus-all prediction tasks. In this paper we propose *long-term recurrent convolutional networks* (LRCNs), a novel architecture for visual recognition and description which combines convolutional layers and long-range temporal recursion and is end-to-end trainable (see Figure 1). We instantiate our architecture for specific video activity recognition, image caption genera-

#citations 5yrs: 2652

Publication Year: 2015

Papers in the 5 years before collaboration

```
['multimedia systems', 'natural languages', 'image processing', 'machine learning', 'artificial intelligence', 'semantics', 'ontology', 'natural language processing', 'image coding', 'linguistics']
```

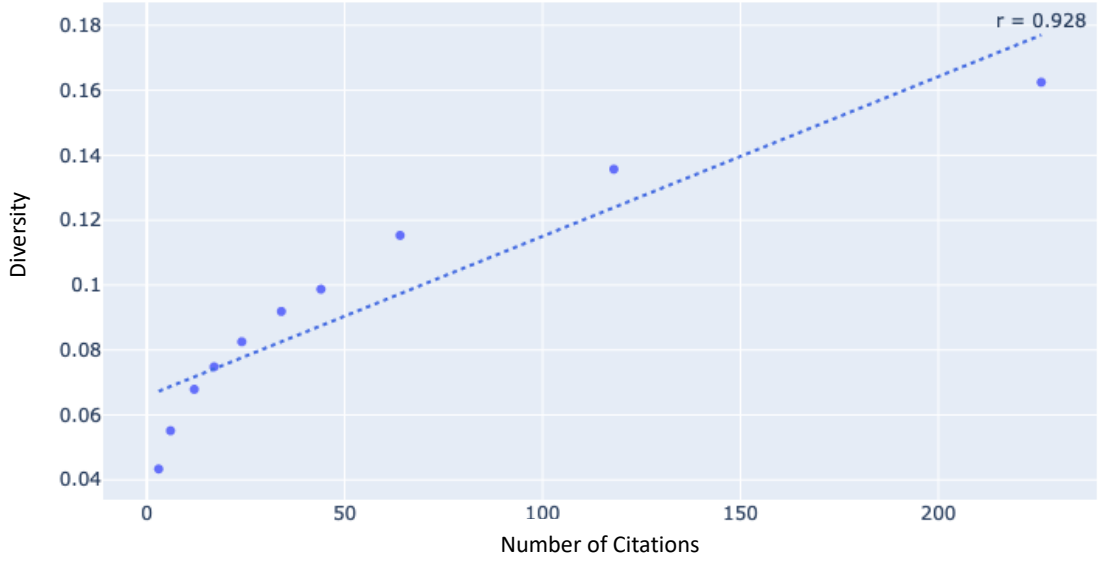
```
['convolutional neural networks', 'auto encoders', 'multilayer perceptrons', 'image analysis', 'image processing', 'machine learning', 'neural networks', 'artificial intelligence', 'network architecture', 'backpropagation algorithm']
```

```
['computer imaging and vision', 'object detection', 'image analysis', 'image processing', 'pattern recognition', 'machine learning', 'artificial intelligence', 'neural networks', 'object recognition', 'image segmentation']
```

Author Expertise

Insights – Part I

- An analysis in Computer Science on ~1M papers and 4.4M authors
- Diversity and number of citations are highly correlated (Pearson's r=0.9, p<0.0001)



Salatino et al. (2023). *Diversity of Expertise is Key to Scientific Impact: a Large-Scale Analysis in the Field of Computer Science*. STI 2023

Insights – Part II

In this case, an AI system has been employed to model the expertise of authors.

This knowledge can be used to:

- develop new policies for improving the research enterprise and accelerate science
- build a successful research team

Agenda

Automatic Literature Review Generation

Hypothesis Generation

Diversity of Expertise

Scientific Question Answering

How to integrate LLM and Scientific Knowledge Graphs?

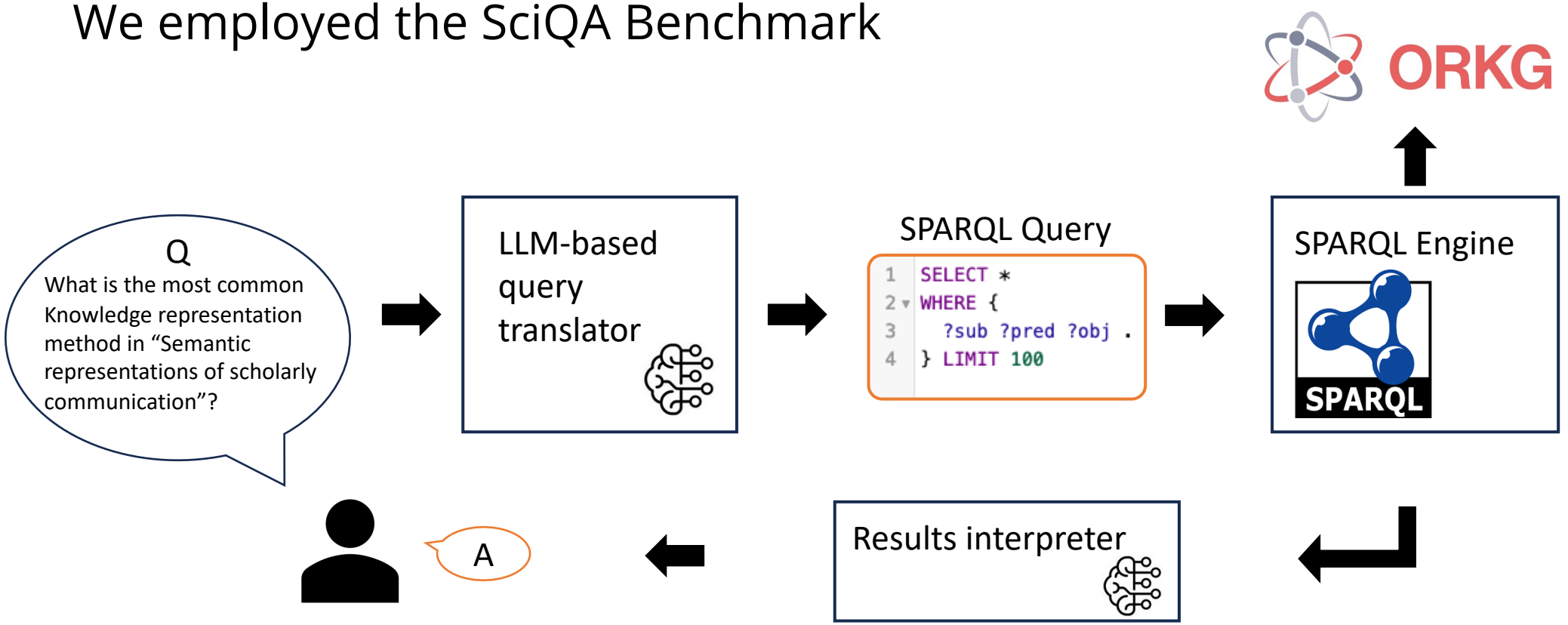
SKGs are complementary to LLMs. They can improve their performance and alleviate hallucinations.

- SKGs can **support RAG** (Retrieval Augmented Generation) – feeding to the LLM only relevant and verifiable data (e.g., portions of articles)
- LLMs can be fine-tuned to **translate natural language to SPARQL** queries over the knowledge graphs (see for example *SciQA*)
- SKGs can be used to **detect hallucinations** (e.g., fake papers)
- LLMs can be augmented with various **knowledge injection techniques** to improve performance over several tasks.

Question Answering over Knowledge Graph

We conducted an in-depth analysis of LLMs for **scientific question answering** and different ways to optimize them.

We employed the SciQA Benchmark



Lehmann et al. (2024) Large Language Models for Scientific Question Answering: an Extensive Analysis of the SciQA Benchmark. Submitted to ESWC 2024.

Question Answering over Knowledge Graph

- ‘Solved’ SciQA, the top benchmark in the area with >97% F1
- Demonstrated excellent performance with fine-tuned version of small models (e.g.,T5)
- Now working on the creation of more challenging benchmarks

Strat.	C	Test name	S	T5-base	GPT2-large	Dolly-v2-3b	GPT-3.5-turbo
ZSL					0.0653 (0)	0.1087 (0)	0.2632 (0)
FT				**0.9751 (483)**	**0.9669 (430)**		
FSL	Similarity		1		0.2718 (0)	**0.8792** (167)	0.9368 (356)
			3		0.4051 (2)	0.8304 (182)	0.9667 (451)
			5			0.8242 (180)	0.9709 (464)
			7			0.8052 (181)	**0.9736 (475)**
	Ent.	Same_Templ	1		0.2029 (0)	0.5734 (2)	0.8988 (205)
		Ran_Templ	1		0.1421 (0)	0.4402 (0)	0.7016 (26)
		Low_Perp	1		0.2788 (0)	0.6757 (1)	
	Random		1		0.2005 (0)	0.5659 (27)	0.7362 (45)
			3		0.2187 (0)	0.5900 (31)	0.8259 (113)
			5			0.6242 (51)	0.8675 (165)
			7			0.6576 (69)	0.8905 (189)
	Diversity	Test A	3		0.2215 (0)	0.7000 (43)	0.9378 (315)
			5			0.6525 (39)	0.9428 (328)
			7			0.6729 (46)	0.9375 (313)
		Test B	3		0.2988 (1)	0.8025 (171)	0.9561 (412)
			5			0.8181 (201)	0.9566 (417)
			7			0.8261 **(212)**	0.9562 (422)

Lehmann et al. (2024) Large Language Models for Scientific Question Answering: an Extensive Analysis of the SciQA Benchmark. Submitted to ESWC 2024.

Conclusions & Future Work

- LLMs are shifting the paradigms of AI
- LLMs have unlocked a new level of opportunities

- **All that glitters is not gold!**

- Several challenges to address:
  - Quality of the responses
  - Quality of the skills
  - Ethical
  - Sustainability

Relevant Publications

Meloni et al. (2023) AIDA-Bot 2.0: Enhancing Conversational Agents with Knowledge Graphs for Analysing the Research Landscape. International Semantic Web Conference 2023, Athens, Greece.

Salatino et al. (2023) Diversity of Expertise is Key to Scientific Impact: a Large-Scale Analysis in the Field of Computer Science. International Conference on Science, Technology and Innovation Indicators 2023, Leiden, The Netherlands.

Cadeddu et al. (2023) Enhancing Scholarly Understanding: A Comparison of Knowledge Injection Strategies in Large Language Models. The Deep Learning for Knowledge Graphs (DL4KG@ISWC2023) Workshop, Athens, Greece.

Cadeddu et al. (2023) A Comparative Analysis of Knowledge Injection Strategies for Large Language Models in the Scholarly Domain. Submitted to Engineering Applications of Artificial Intelligence.

Borrego et al. (2023) Research Hypothesis Generation over Scientific Knowledge Graphs. Submitted to Knowledge Based Systems.

Meloni et al. (2023) Integrating Conversational Agents and Knowledge Graphs within the Scholarly Domain. IEEE Access.

Meloni et al. (2023) Enhancing Conversational Agents Reliability and Accuracy by Integrating Scientific Knowledge Graphs. Submitted to Expert System with Applications.

Buscaldi et al. (2024) Citation Prediction by Leveraging Transformers and Natural Language Processing Heuristics. Information Processing and Management.

Lehmann et al. (2024) Large Language Models for Scientific Question Answering: an Extensive Analysis of the SciQA Benchmark. ESWC 2024.

Bolanos et al. (2024) Artificial Intelligence for Literature Reviews: Opportunities and Challenges. https://arxiv.org/abs/2402.08565.

angelo.salatino@open.ac.uk

https://skm.kmi.open.ac.uk/