AI Research Assistant

Andrew Peterson*

June 2023

Abstract

We conceive of a range of tasks on which AI can aid human-based research. Rather than displacing human researchers, who are far superior in planning, analysis, and evaluation, we identify key tasks on which AI can best act as an assistant to human researchers. To do so, we begin by identifying areas of comparative advantage of AI, then develop tools to improve on these tasks and integrate them into a research pipeline. The exact usage will naturally depend on the type of research being conducted, but we develop a range of research tools that help human researchers brainstorm new approaches, critically evaluate existing research, and identify and process possible sources of information. Finally, we develop tools for evaluating the quality of these outputs with an eye towards continuous improvement.

1 Introduction

How can researchers best leverage artificial intelligence to improve their scientific workflow? While some have already dreamed of fully-automated AI researchers, who conduct the full scientific process from literature review to hypothesis generation and even testing based on on-demand laboratories, we believe this is not the best way currently to leverage AI technologies. Instead, we attempt to identify key comparative advantages of AI systems such as GPT-4, and develop tools to improve those processes, while preserving the highest planning, reasoning, analytic, supervision, and evaluation tasks for human researchers.

1.1 What is the comparative advantages of GPT-4 versus human researchers?

Tasks on which humans dominate GPT-4 (i.e. GPT-4 might be much cheaper, but results are so much poorer that it's better to have a human involved):

- Planning (GPT-4 basically incapable of planning acc. to Bubeck et al. [2023])
- Reasoning, analysis (assuming a researcher well-versed in their field)
- Evaluation of (intermediate and) final outputs

Tasks on which a GPT-4 based AI system (see below for more info) has a comparative advantage (i.e. much lower cost in time, $)

- summarizing texts, gathering similar papers,
- idea generation - GPT-4 can quickly “brainstorm" dozens of possible ideas, critiques, ...

Tasks on which GPT-4 has an absolute advantage over humans:

- breadth of knowledge: GPT-4 knows a vast array of topics, even if its knowledge is shallower in any given field than a specialized human
- Checking grammar and spelling, translating simple code between different computer languages (again, vast knowledge of different programming languages)
- Transcribing audio

*Université de Poitiers

2 Existing AI Research Tools and Resources

- SemanticScholar Various tools, including paper lookup based on semantic embedding
- Elicit.org
- Consensus
- Nomic.ai - embedding of 21 million PubMed papers, tools for visualizing embeddings
- Iris.AI “Search, Filter, Analyze, Extract, Summarize, Automate, Report”
- “AutoResearcher" Automated literature review, summarize results from SemanticScholar retrieval
- Chat with academic papers
- Google Scholar
- Microsoft Academic
- There are more specialized approaches to AI theorem solving, such as Davies et al. [2021], Lample et al. [2022], and for discovering algorithms, such as Fawzi et al. [2022].
- R Discovery: literature discovery from keywords/ tags.

3 AI Research Assistant Tasks

3.1 Unlikely Insights - AI Connections and Analogical Reasoning

In contrast with the above approaches, our focus is not to write a literature review in place of the researcher, but to make surprising connections to diverse sources, pull the abstracts, and present them in a quickly digestible format so that a research can mine them for ideas. We presume that out of 100 such proposals, only 1 or 2 might be of interest to a researcher, but if we are succeeding, those few ideas would not have occurred to the researcher through their normal process.

Input: Research question
Optional Input: Additional prompt info (example: ask for historical examples
Output: Revised text

Based on the research question, ask GPT-4 to generate a list of analogies or avenues for related research;
foreach *topic in analogies_list* **do**
    Generate keyword combinations from the topic description;
    Get papers using SemanticScholar;
    Sort papers based on similarity, citation counts;
    Parse these results;
    Generate an outline for the output based on the research question and titles of the articles;
    Ask GPT-4 to order the articles according to the outline;
    Ask GPT-4 to re-draft the final text from the abstracts and article titles organized according to the outline;
end
Compile results to pdf from a LaTeXdocument;

Algorithm 1: Analogical Reasoning-Based Literature Insights Algorithm

See the Appendix for extracts from examples using this approach to elucidate ideas about AI disruption and approaches to solving mathematical theorems.

3.2 Mining Research for New Ideas

Here again we aim to leverage the ability of AI to generate a range of possible avenues for new research, which a human would then look through to identify those worth pursuing further. Instead of first having GPT brainstorm avenues of research, then looking up and summarizing related academic papers on those topics, here the approach is to start with one or more key academic papers, then use them as a foundation on which to

- Identify stated and/ or unstated premises in the paper, that might have substantive bearing on the conclusions
- Comparing and contrasting two or more papers, with a view to identifying how to resolve their conflicting claims through additional research
- Analyzing papers to identify anomalies, abnormal data
- Analyzing a paper and suggesting future research directions
- Asking GPT to analyze a paper from a specific perspective (e.g. "What objections might an [X-type] person have to this research?")

(forthcoming)

3.3 Research Agent

We have currently developed a streamlit demo which uses Grobid-converted XML files from the pdf. Functions are defined for navigating the xml structure, extracting tables, etc. Final step (optional) turn these functions into tools for a LangChain agent.

(partially completed)

3.4 Critique Agent

Any researcher can put the text of academic paper into ChatGPT and ask for critiques. What is lacking however, is:

- A base of tested prompts for specific research applications
- The ability to work with long texts that don't fit into the context window
- Tools for operating with differently formatted elements of the paper (tables, figures, citations, algorithms, etc) - (based on the existing codebase for Section 3.3)

(forthcoming)

4 Evaluation

As noted above, the best evaluation for these outputs is the considered judgement of an expert researcher in a given field, which can be costly. One approach is simply to release the demo and get feedback from researchers. In addition, however, we develop a pipeline for evaluating outputs that can be (semi-) automated:

Approach 1: Human Gold Standard. Since GPT-4 knowledge ends in 2021, we can 'backtest' the utility of our approaches by having it attempt to generate research outputs from 2022 or 2023. For example, the Kahn–Kalai conjecture was known to GPT-4 but a proof was not provided until 2022 [Park and Pham, 2022]. We can then evaluate whether the approaches proposed above might have been useful in generating such a proof (with the caveat that there are often different approaches to prove a given result). Similarly, when the goal is a review article, one can find a review article published in 2023 (without a preprint in the training data) and compare an AI-generated literature review to this standard. Beyond the question of quality, this can be used to identify different ways to structure the analysis or write-up.

Approach 2: Comparative GPT evaluations When a human-based gold-standard does not exist, two or more alternative approaches can be evaluated by using GPT-4 to evaluate the quality, for example using an ELO-based rating system.

(forthcoming)

References

Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*, 2023.

Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, András Juhász, Marc Lackenby, Geordie Williamson, Demis Hassabis, and Pushmeet Kohli. Advancing mathematics by guiding human intuition with AI. *Nature*, 600(7887):70–74, December 2021. ISSN 1476-4687. doi:10.1038/s41586-021-04086-x. URL `https://doi.org/10.1038/s41586-021-04086-x`.

Guillaume Lample, Timothee Lacroix, Marie-Anne Lachaux, Aurelien Rodriguez, Amaury Hayat, Thibaut Lavril, Gabriel Ebner, and Xavier Martinet. Hypertree proof search for neural theorem proving. *Advances in Neural Information Processing Systems*, 35:26337–26349, 2022.

Alhussein Fawzi, Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Francisco J R Ruiz, Julian Schrittwieser, Grzegorz Swirszcz, et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature*, 610(7930):47–53, 2022.

Jinyoung Park and Huy Tuan Pham. A proof of the kahn-kalai conjecture. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 636–639. IEEE, 2022.

J. Huebschmann. Poisson cohomology and quantization. *Journal fr die reine und angewandte Mathematik (Crelles Journal)*, 1990:113 – 57, 2013.

J. May and K. Ponto. More concise algebraic topology: Localization, completion, and model categories. 2012.

J. Pardon. An algebraic approach to virtual fundamental cycles on moduli spaces of pseudo-holomorphic curves. *Geometry & Topology*, 20:779–1034, 2013.

M. Husek and J. Mill. Recent progress in general topology. 2002.

M. Hopkins. Algebraic topology and modular forms. *arXiv: Algebraic Topology*, 2002.

S. Popa. On a class of type ii_1 factors with betti numbers invariants. *Annals of Mathematics*, 163:809–899, 2002.

Yu Xie, Guang zhen Wang, P. Box, S. Achen, Miranda Brown, Cindy Glovinsky, Katherine King, Yang Jiang, A. Thornton, and L. Young-DeMarco. Chinese people's beliefs about the relationship between economic development and social inequality. 2008.

R. Gowland, Anwen C Caffell, S. Newman, A. Levene, Malin, and Holst. Title: Broken childhoods: Rural and urban non-adult health during the industrial revolution in northern england (eighteenth-nineteenth centuries). authors: *rebecca. 2018.

H. Balzer. Russia and china in the global economy. *Demokratizatsiya*, 16:37–47, 2008.

S. Church. The eurasian silk road: Its historical roots and the chinese imagination. *Cambridge Journal of Eurasian Studies*, 2, 2018.

The following illustrate single page extracts from larger documents generated on different topics.

A Example: Crossing Mathematical Fields

Here is the source research question that served as a prompt: "The inscribed square problem, also known as the square peg problem or the Toeplitz' conjecture, is unsolved: Does every plane simple closed curve contain all four vertices of some square? This is true if the curve is convex or piecewise smooth and in other special cases. Please provide suggestions for diverse approaches from other fields of mathematics that might help solve this problem."

A.1 Topological invariants and algebraic topology

A.1.1 Foundations of Topological Invariants and Algebraic Topology

In the study of algebraic topology, various papers have contributed to the development of the field. Elmendorf et al. (2007) discusses rings, modules, and algebras in stable homotopy theory, focusing on the category of $\mathbb{L}$-spectra and related topics. This can be mathematically expressed as:

$$\mathbb{L} = \bigoplus_{n \in \mathbb{Z}} \mathbb{L}_n$$

where $\mathbb{L}_n$ represents the n-th homotopy group of the $\mathbb{L}$-spectrum.

Huebschmann [2013] introduces Poisson cohomology and quantization, constructing an (R,A)-Lie algebra structure on the A-module of Kähler differentials of A. May and Ponto [2012] provides a concise treatment of algebraic topology, covering topics such as localization and completion of topological spaces, model categories, and Hopf algebras. Pardon [2013] presents an algebraic approach to virtual fundamental cycles on moduli spaces of pseudo-holomorphic curves, introducing the notion of an implicit atlas on a moduli space. Husek and Mill [2002] reviews recent progress in general topology, discussing various topics in topological dynamics, Banach spaces, metrizable spaces, and more.

A.1.2 Applications of Topological Methods in Algebraic Geometry and Group Theory

Hopkins [2002] explores the connection between algebraic topology and modular forms, discussing the construction of topological modular forms and their applications in various areas of mathematics, including geometry, mathematical physics, number theory, representation theory, and topology.

A.1.3 Topological Invariants in Braid Theory and Type II_1 Factors

Popa [2002] investigates a class of type II_1 factors with Betti numbers invariants, proving that a type II_1 factor M can have at most one Cartan subalgebra A satisfying a combination of rigidity and compact approximation properties. This can be mathematically expressed as:

$$\forall M \text{ (Type II}_1 \text{ Factors)}, \ \exists! A \text{ (Cartan Subalgebras of } M) : \text{Rigidity}(A) \ \wedge \ \text{CompactApprox}(A)$$

In the above equation, $\exists!$ denotes there exists exactly one, $\wedge$ denotes logical AND, $\text{Rigidity}(A)$ represents the rigidity property of the Cartan subalgebra A, and $\text{CompactApprox}(A)$ represents the compact approximation property of A.

B Example: Social and Political Impacts of Technology Innovation

B.1 GPT identified the following historical examples to explore

1. The Industrial Revolution: Study how the transition from agrarian societies to industrialized ones led to job displacement, economic inequality, and shifts in power dynamics.
2. The Green Revolution: Examine the effects of the agricultural transformation in the mid-20th century on global food security, economic distribution, and employment changes.
3. The Gutenberg Press: Explore how the invention of the printing press disrupted traditional power structures, spread knowledge, and influenced the sense of individualism.
4. Japanese Edo Period: Investigate how Japan's isolationist policies during the Edo period (1603-1868) affected economic distribution, social stratification, and technological advancements.
5. The Silk Road Trade Network: Analyze the exchange of goods, ideas, and cultural practices along the Silk Road, and how it influenced the power dynamics and economic distribution in the participating regions.
6. The Inca Empire's Economic System: Research the Inca Empire's unique economic system based on reciprocity and labor taxation, and how it maintained social cohesion and resource distribution.
7. The Great Divergence: Examine the factors that led to the economic and technological disparities between Western Europe and East Asia during the 18th and 19th centuries.
8. The Space Race: Investigate the competition between the United States and the Soviet Union during the Cold War, and how it affected technological advancements, national pride, and global power dynamics.
9. The Internet Revolution: Study the widespread adoption of the internet and its effects on job displacement, economic distribution, and the power dynamics among nations and individuals.
10. The Cultural Revolution in China: Analyze how Mao Zedong's Cultural Revolution disrupted traditional power structures, social norms, and individual sense of meaning in Chinese society.
11. The Arab Spring: Investigate the role of social media and technology in the uprisings across the Arab world, and how it affected power dynamics, economic distribution, and individual empowerment.
12. The Meiji Restoration in Japan: Examine how Japan's rapid modernization and industrialization during the Meiji period (1868-1912) impacted job displacement, economic distribution, and national power dynamics.

B.2 Economic Factors: Trade, Industrialization, and Capitalism and Social inequality

The relationship between economic development and social inequality has long been hypothesized in both economics and sociology. In the case of China, rapid economic growth has been accompanied by a sharp rise in economic inequality Xie et al. [2008]. This has led to debates on whether economic development inevitably leads to increased inequality, as well as discussions on the perceived unfairness of inequality and its institutional mechanisms, such as corruption Xie et al. [2008].

During the Industrial Revolution in Northern England, urbanization and industrialization had detrimental effects on the health of children, particularly those living in urban environments and working in factories Gowland et al. [2018]. Contrary to expectations, a study comparing bioarchaeological evidence for non-adult health from contemporaneous urban and rural sites in the north of England found equal prevalence rates of metabolic and dental disease at both sites, but greater evidence of growth disruption and respiratory disease in the rural site Gowland et al. [2018]. This suggests that interpretations of rural/urban health during this period must take into account the consequences of social inequalities and economic migration.

In Russia and China, differing approaches to integration with the international economy have led to contrasting outcomes in terms of economic growth and development Balzer [2008]. China has embraced economic globalization and integration on a scale surpassing many other Asian countries, while Russia remains wary and peripheral. Russia's economy is open, but selling natural resources and arms generates few linkages leading to higher value-added production. China's integration is "thick," involving participation in technology chains and entire product cycles Balzer [2008].

The Eurasian Silk Road played a significant role in the economic development of China and its interactions with other regions Church [2018]. The Silk Road facilitated the exchange of goods, ideas, technologies, and religions across Eurasia, and its various routes connected China to Central Asia and beyond. Under Mongol rule, the route was at times an unbroken corridor between East and West, allowing for extensive travel and trade. When the Mongol empire broke up, travel overland was restricted again, which may have been why China took to the seas in the Ming dynasty. At present, China is building a New Silk Road to connect with the rest of the world in a more integrated way than ever before Church [2018].