

# Amaan Shubaily Assignment

Summary For: "Re-Search: Integrating Search and Reasoning with Reinforcement Learning for Large Language Models"

This summary is based on the research paper titled **Re Search: Integrating Search and Reasoning with Reinforcement Learning for Large Language Models**. The paper talks about a way to improve how AI language models answer complex questions. Instead of just guessing or giving generic answers, the idea is to teach the model how to **search for information** and then **reason through it properly** before responding. To be honest, some parts of the paper were pretty technical but I tried my best to understand the main ideas.

From what I understood, in this system/model the AI model doesn't just rely on what it already knows it actually learns how to look stuff up (like searching online or in a knowledge base), and then uses that info to come up with a better answer. It also learns over time by getting rewards when it answers correctly. it's kind of like training the AI to think before speaking, instead of just blurting something out right away.

GRPO is a new way for the AI to learn. It stands for **Generalized Reaching Policy Optimization** which sounds fancy, but to me, it just means the model figures out what steps to take to get the info it needs. Instead of being shown what to do every time (like in regular training), it learns by doing — kind of like trial and error. If it searches smartly and gives the right answer, it gets a reward, and that helps it improve over time.

There are a few parts to the system:

**Search Module:** Finds extra info from outside sources when the model needs help.

**Reasoning Module:** Sorts through the info and uses it to try and answer the question.

**Learning Process (This is the GRPO):** The model gets better by trying things and getting feedback (gives rewards) when it does well.

The researchers tested on some challenging question-answering tasks. The model did better than the normal version of an ai model, especially on harder questions that needed extra thinking or more specific info. It was able to ask better questions while searching and use what it found more effectively by looking up everywhere

This model is very used in todays world it can be used for 1. Ai assistants that help with research 2. Can be used in customer support for which it look up things before answering the question 3. It is widely used in medical fiels and legal system to make human task more precise and easy

Some of the things that I found challenging and interesting at the same time was all the technical terms and stuff around the RL training model but i thought the idea of teaching a model to search like a human and then think through the answer was really interesting It made me realize how important it is for AI to not just sound smart but to actually **be accurate** and **have a process** behind its answers.

Overall this paper showed me that there's a lot more going on behind the scenes when it comes to improving how AI models think and answer even though I m not an expert, I got the main idea on how AI learn to **search better and think better** and that seems like a smart direction for the future of AI tools.

## **Small Presentation on the Research**



# ReSearch: A New Framework for AI Critical Thinking

ReSearch combines reasoning and external information search for complex questions. It overcomes limitations of relying on pre-existing knowledge.

The AI learns through trial and error. ReSearch is rewarded for correct answers, enhancing accuracy through constant information retrieval.

## Key Components of

### Search Module

Dynamically finds external information in real-time. Employs diverse data sources.

- Real-time search
- Diverse data sources
- Adapts to query demands

### Reasoning Module

Analyzes and synthesizes knowledge from multiple sources. Combines external information with internal knowledge.

- Analyzes and synthesizes
- Combines info
- Validates against sources

## Benefits and Applications



### Enhances AI decision-making

Uses real-world data for better decisions.



### Reduces reliance on biased data

Avoids outdated training data.



### Solves problems

Potential in medicine, finance, and engineering.



## Group discussion points :

- Should AI be allowed to look things up before answering?
- Can ai learn better by trying and failing like humans
- Does searching for info makes it more accurate
- How can ai know which sources to be trusted while doing the search
- Is it safe to let ai find answers on its own
- What if ai gets too good at finding answers ?
- Should we let ai use the internet all the time or limit it ?