

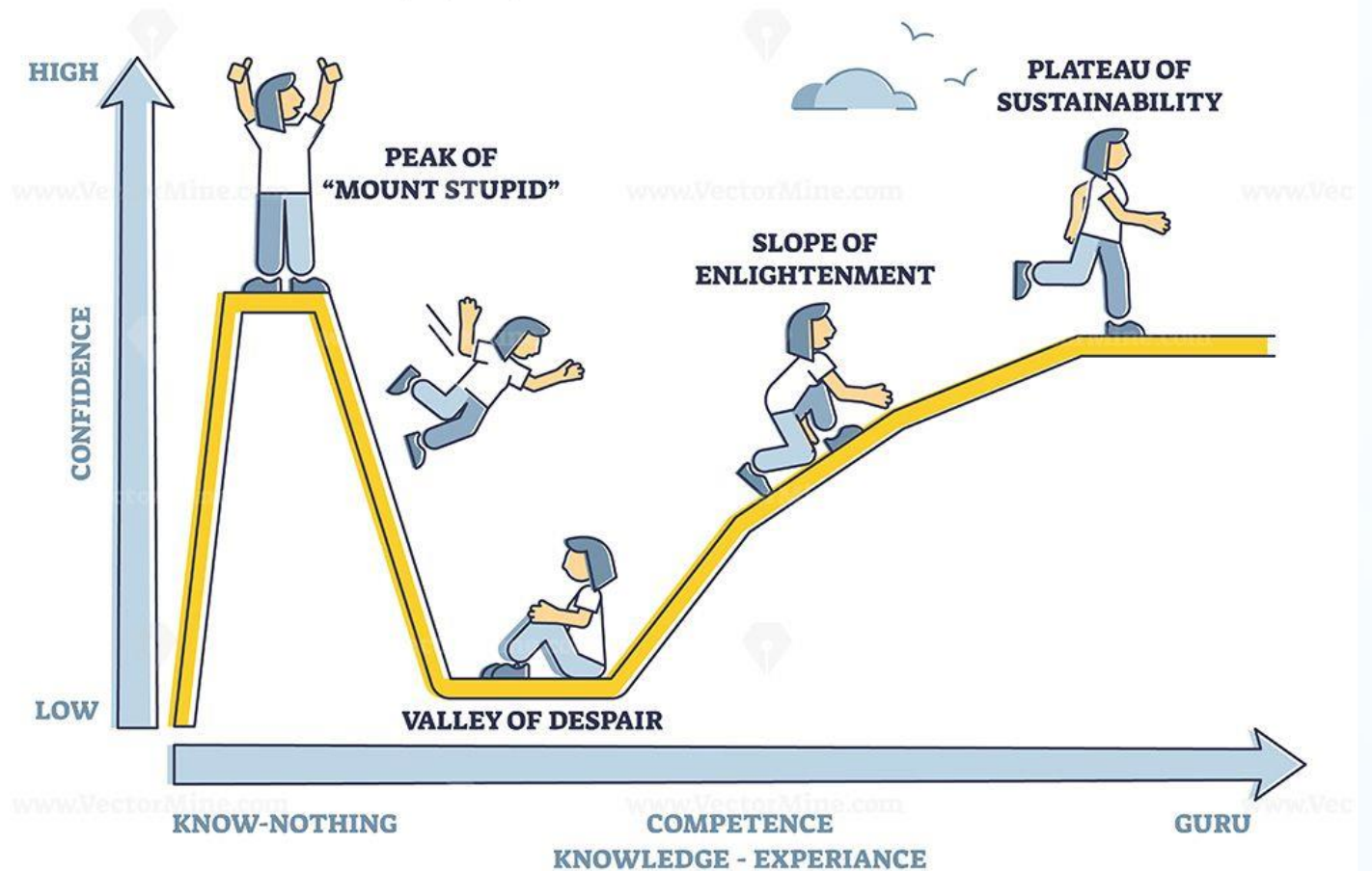
# HIPPO RAG

SIMPLY EXPLAINED

# GOALS

Explain simply but  
not over simplified

## DUNNING KRUGER EFFECT



# GOALS

Lifting barrier for newcomers to the AI land who wish they had GPU and instructions for running SOTA models.





# EXPLAINING THE PAPER

# PAPER ANALYSIS FRAME WORK

## PROBLEM(S) ADDRESSED

(What was the problem that the author(s) were trying to solve?)

## MOTIVATION

(Why were previous approaches insufficient?)

## PROPOSED SOLUTION

(What did they propose?)

## EVALUATION & RESULTS

(How was the solution evaluated? What were the results?)

## REFLECTIONS ON LEARNING

(What did you learn from reading the article? Learning can be positive {"I never knew that ..."} or negative {"I learned how not to ..."} or both. Aim for 3-5 items)

## RESEARCH IDEAS OR DIRECTIONS (GRADUATE STUDENT)

(What are some future research directions of the work? Often authors will indicate these, but come up with one or two of your own ideas.)



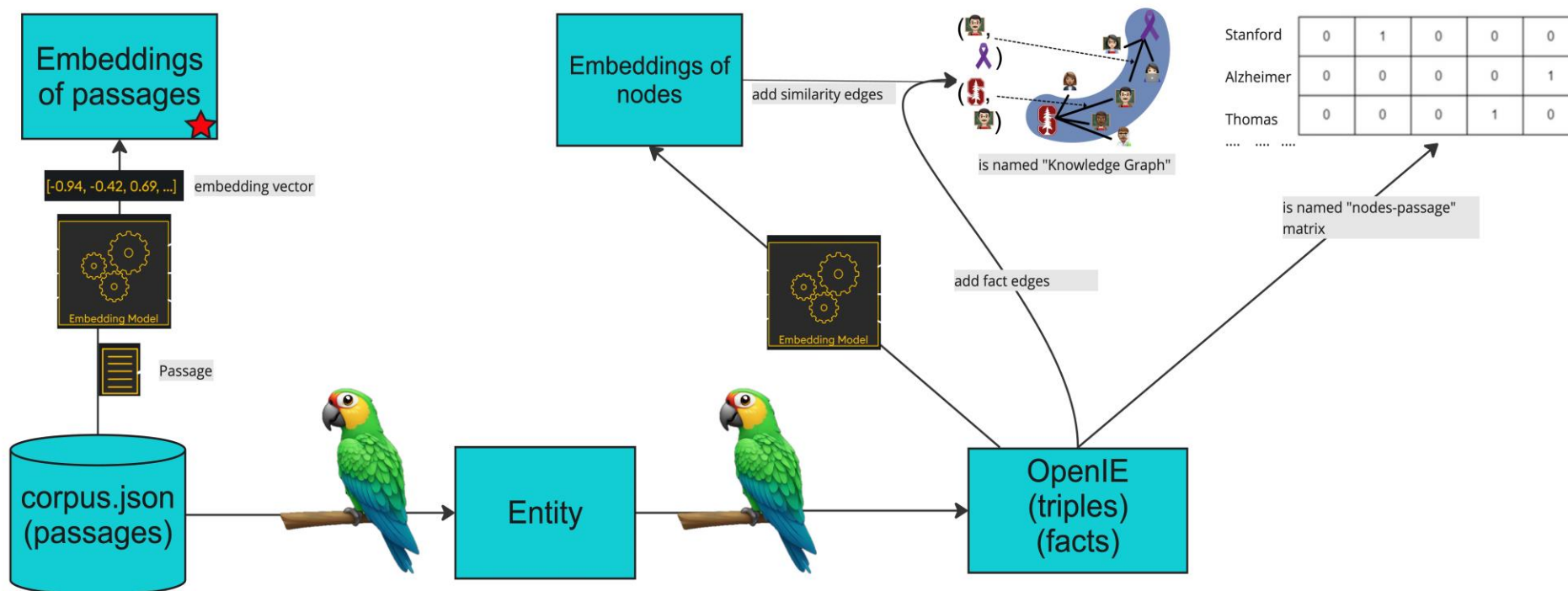
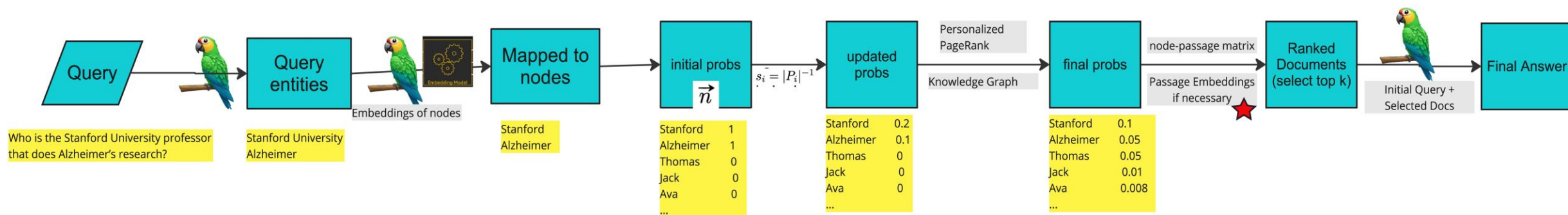
# PROBLEM(S) ADDRESSED

impressive accomplishments, large language models (LLMs), even with retrieval-augmented generation (RAG), still struggle to efficiently and effectively integrate a large amount of new experiences after pre-training. In this work, we introduce HippoRAG, a novel retrieval framework inspired by the hippocampal indexing theory of human long-term memory to enable deeper and more efficient knowledge integration over new experiences. HippoRAG synergistically orchestrates LLMs, knowledge graphs, and the Deepened Deep-Depth algorithm to unify the different

# MOTIVATION

However, current RAG methods are still unable to help LLMs perform tasks that require integrating new knowledge across passage boundaries since each new passage is encoded in isolation. Many

# PROPOSED SOLUTION





# EVALUATION & RESULTS

- 1) Are they cherry picking in their examples?
- 2) Is their approach statistically significant?
- 3) Is their approach practically significant?
- 4) Are they sweeping the limitations under the rug?
- 5) Are there any confounding factors?
- 6) Are their baseline(s) the most recent?
- 7) Where is the creativity/novelty?

# REFLECTIONS ON LEARNING

## Evaluation



Sacrosanct chain

The evaluation metric should be in-line with the learning task, the loss function should be in-line with the evaluation metric, and the model should be designed to minimize the loss.

Important Sentence Selection / F1 score metric / BinacyCrossEnropy / LLM with classification head

An analogy in real-life that the chain is (probably) not well connected :

Does a person have a quality life? / Happiness / number of sad pics on their socials / Count function  
Connected? Connected? ☒ Connected

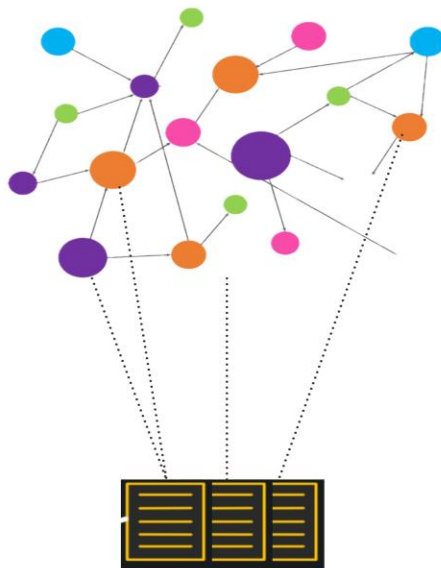
- 1) Despite the fact that we cannot always define the task that we want to learn (like AGI), there are several intermediary steps that might introduce extra gaps that prevent us from fully learning the task.
- 2) Benchmarks are stagnant, especially in text summarization field.

Task / **Benchmark** / Metric / Loss function / Model

# RESEARCH IDEAS OR DIRECTIONS

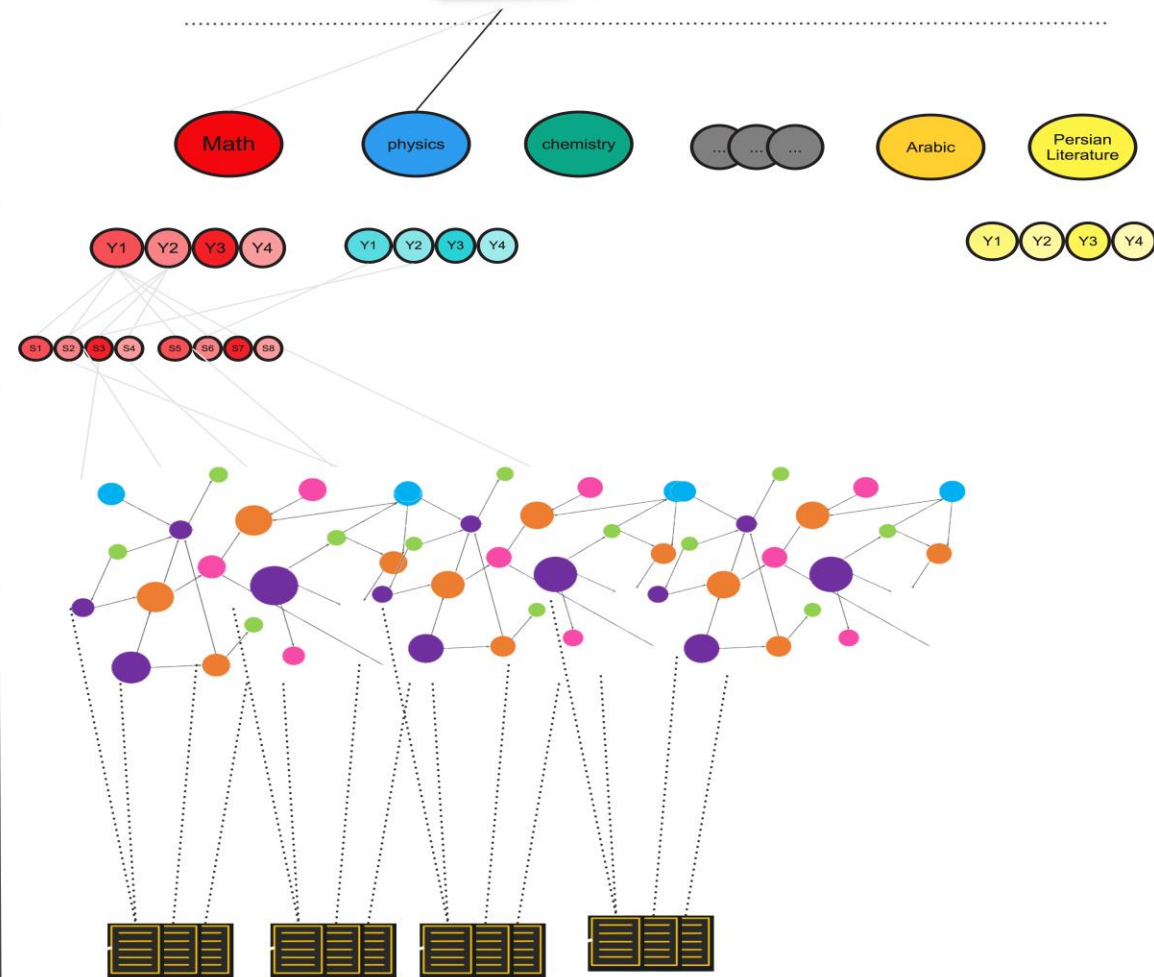
One final exam scenario

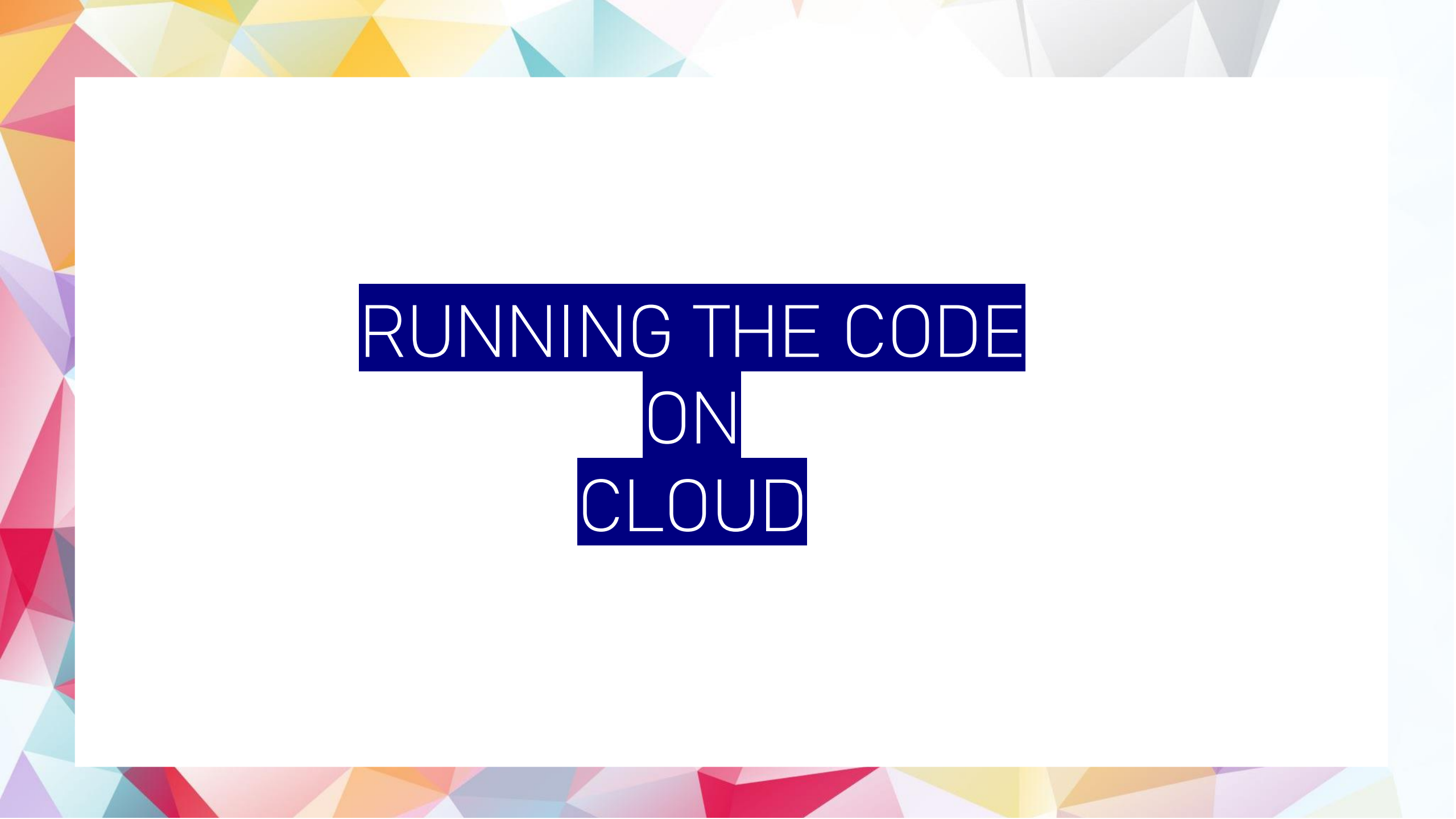
question



Entrance exam scenario

question





# RUNNING THE CODE ON CLOUD

Thank you for making it to this end.  
If you enjoyed, please **share with friends**,  
since sharing means caring.

Also please **smash the like bottom** and  
**subscribe to stay tuned** on the state-of-the-art  
models in the NLP world!