



# PROJECT AD





# OUR TEAM



**MennaTullah Yasser**



**Habiba Mohammad**



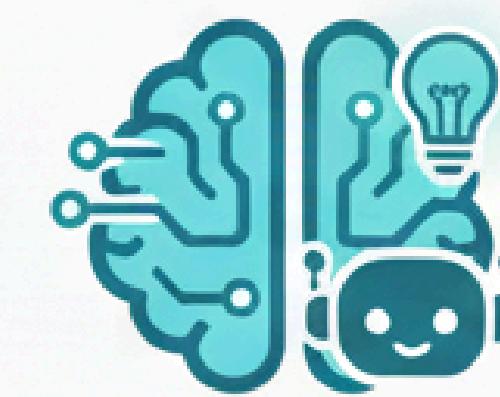
**Aya Abdelmoniem**



**Assem Ali**



**Ahmed Elnashar**

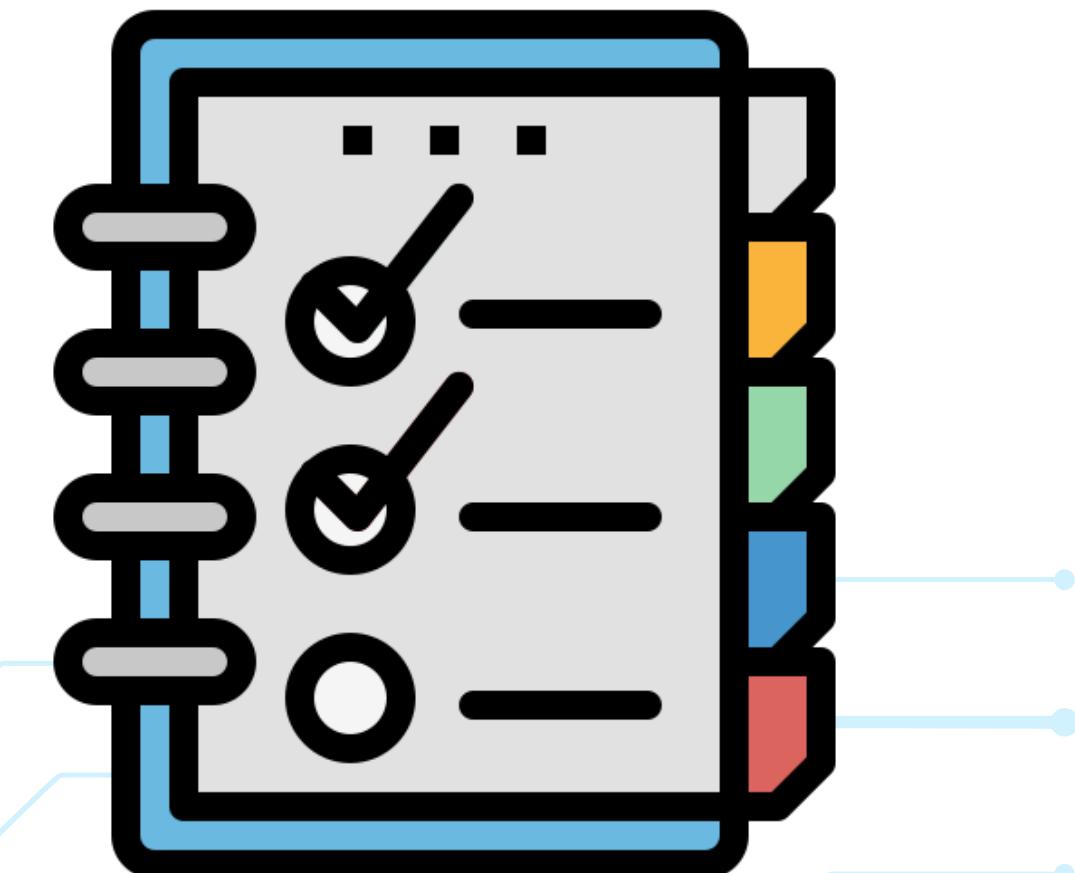


**FAHEM**  
AI Learning Assistant

[www.Fahem.com](http://www.Fahem.com)

# AGENDA

- 1 Business Value
- 2 Challenges
- 3 Workflow
- 4 Technical Stack
- 5 Future Work
- 6 Questions?





# BUSINESS VALUE



# BUSINESS VALUE



Transforms static educational content into adaptive, multimodal learning experiences tailored to every learner's style

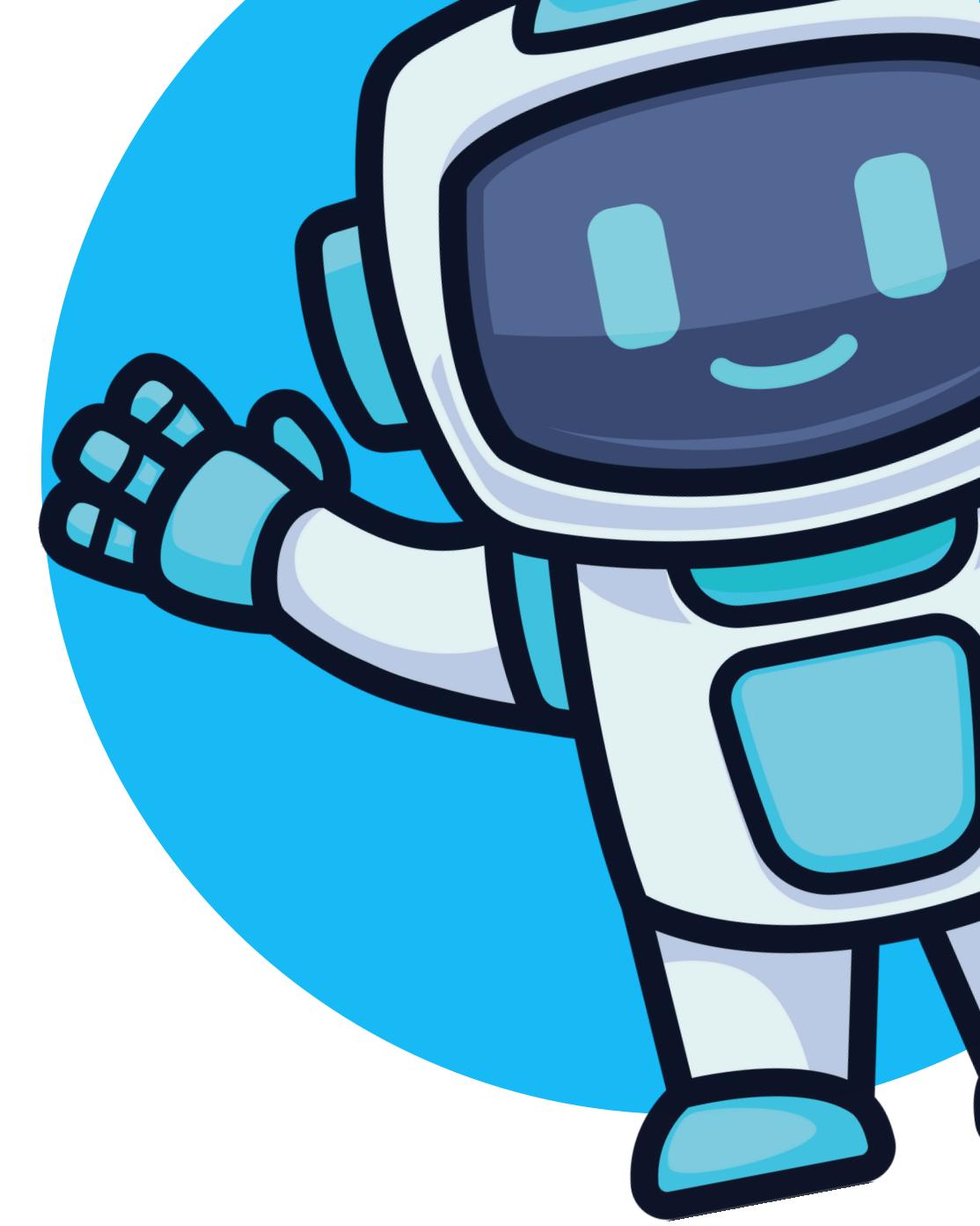
E-Learning Platforms

Learners





# CHALLENGES



# CHALLENGES

1

**Computational Resources Constraints**

2

**Reliance on Third-Party APIs**

3

**Limited Time**

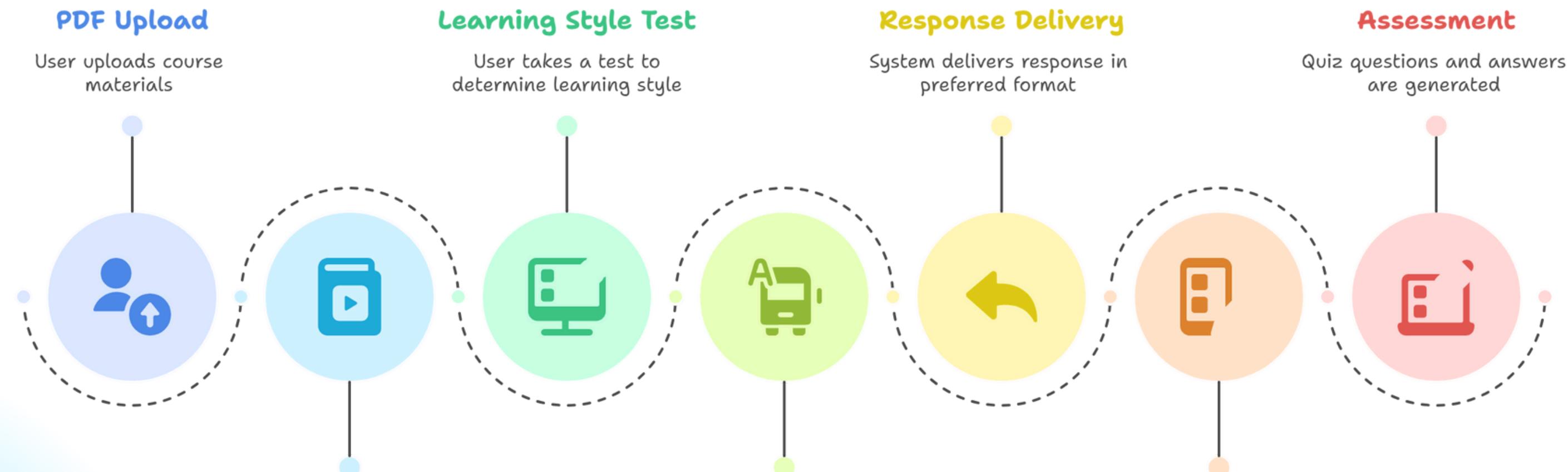
4

**Cost of Deployment Servers**

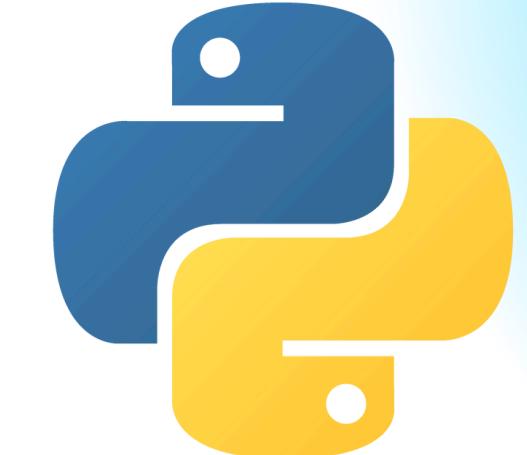


# USER WORKFLOW

## AI Learning Assistant System Architecture



# TECHNICAL TOOLS

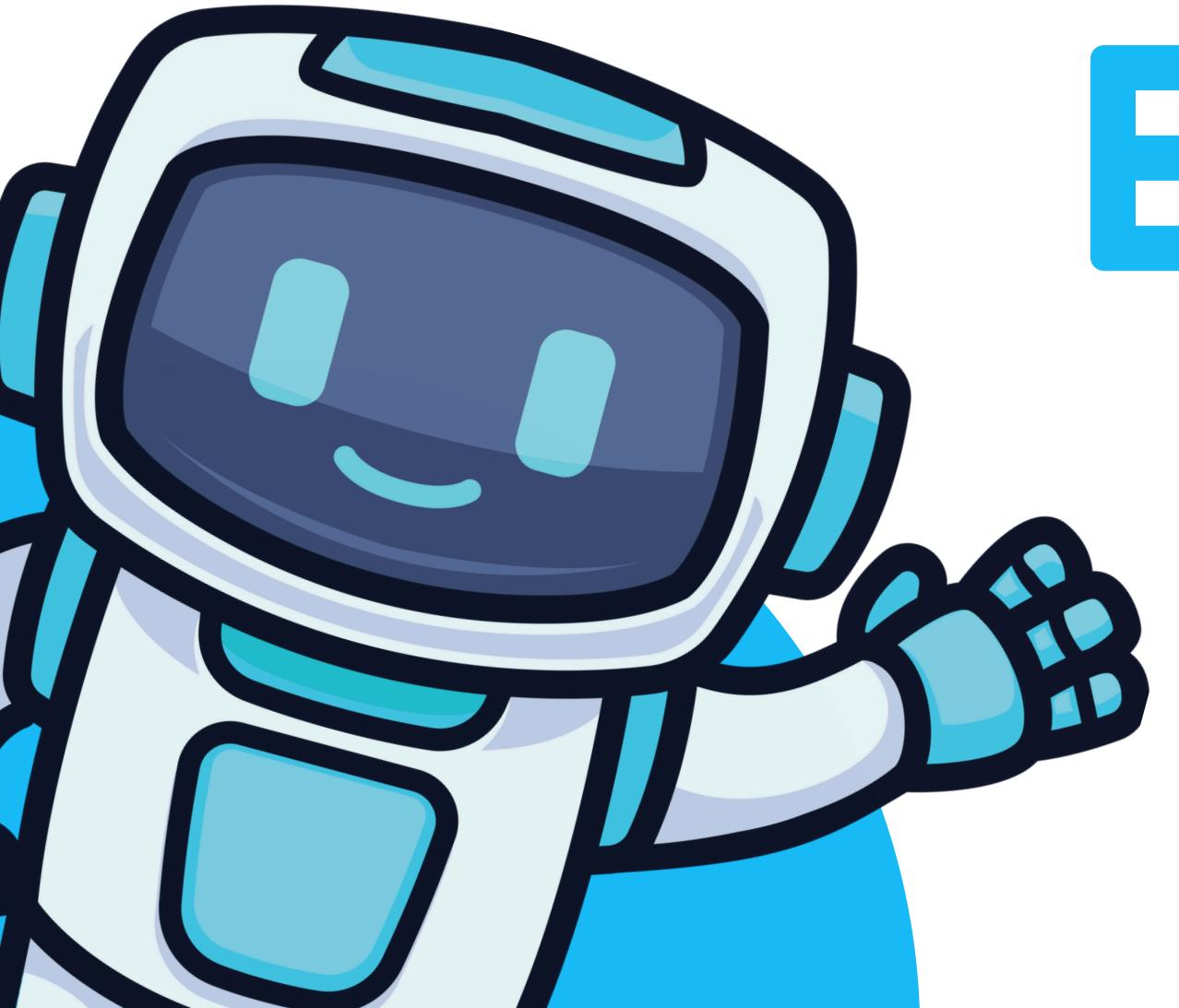


**ngrok**





# DOCUMENT EMBEDDING



# CHALLENGES

1

Dynamic input

2

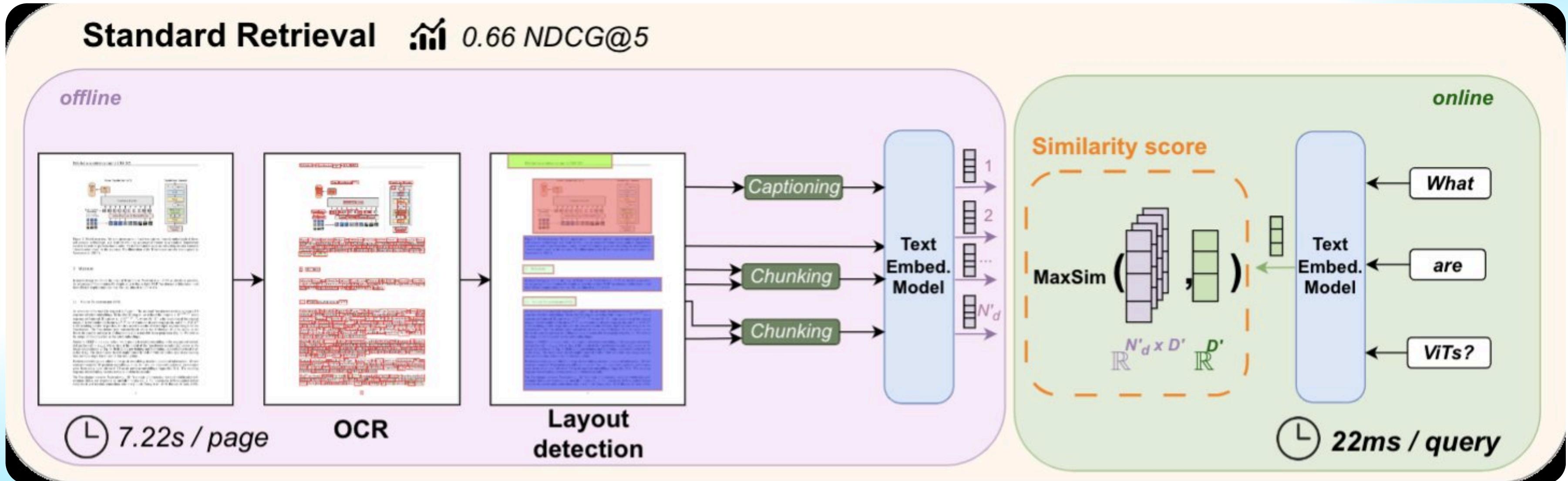
Embedding time

3

computational Power

# STANDARD RETRIEVAL

Standard Retrieval  0.66 NDCG@5



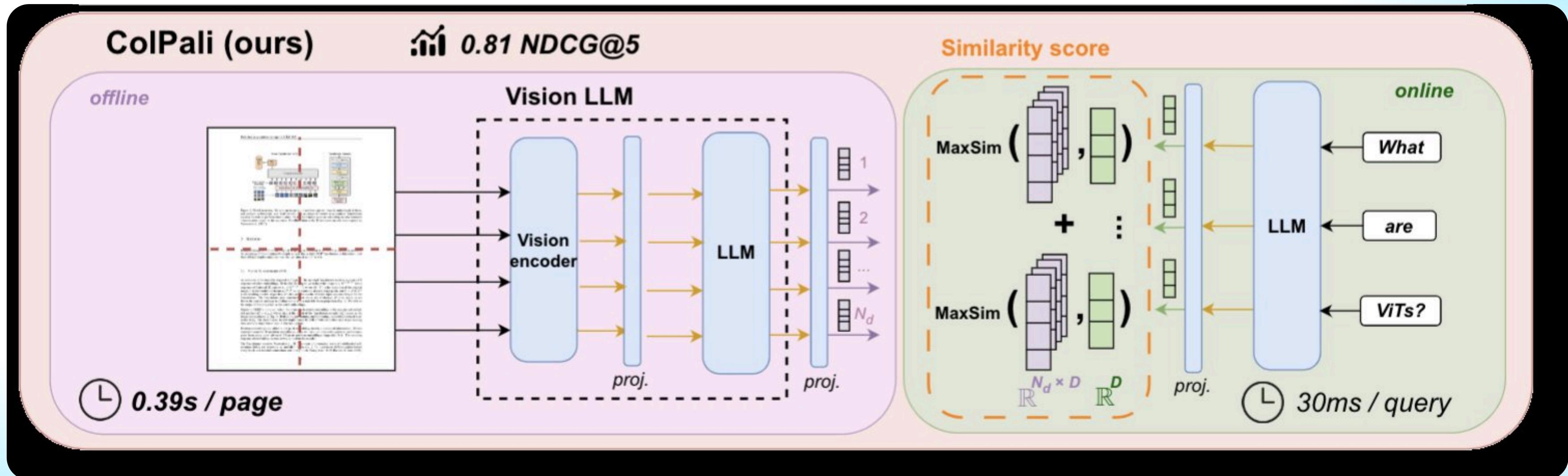


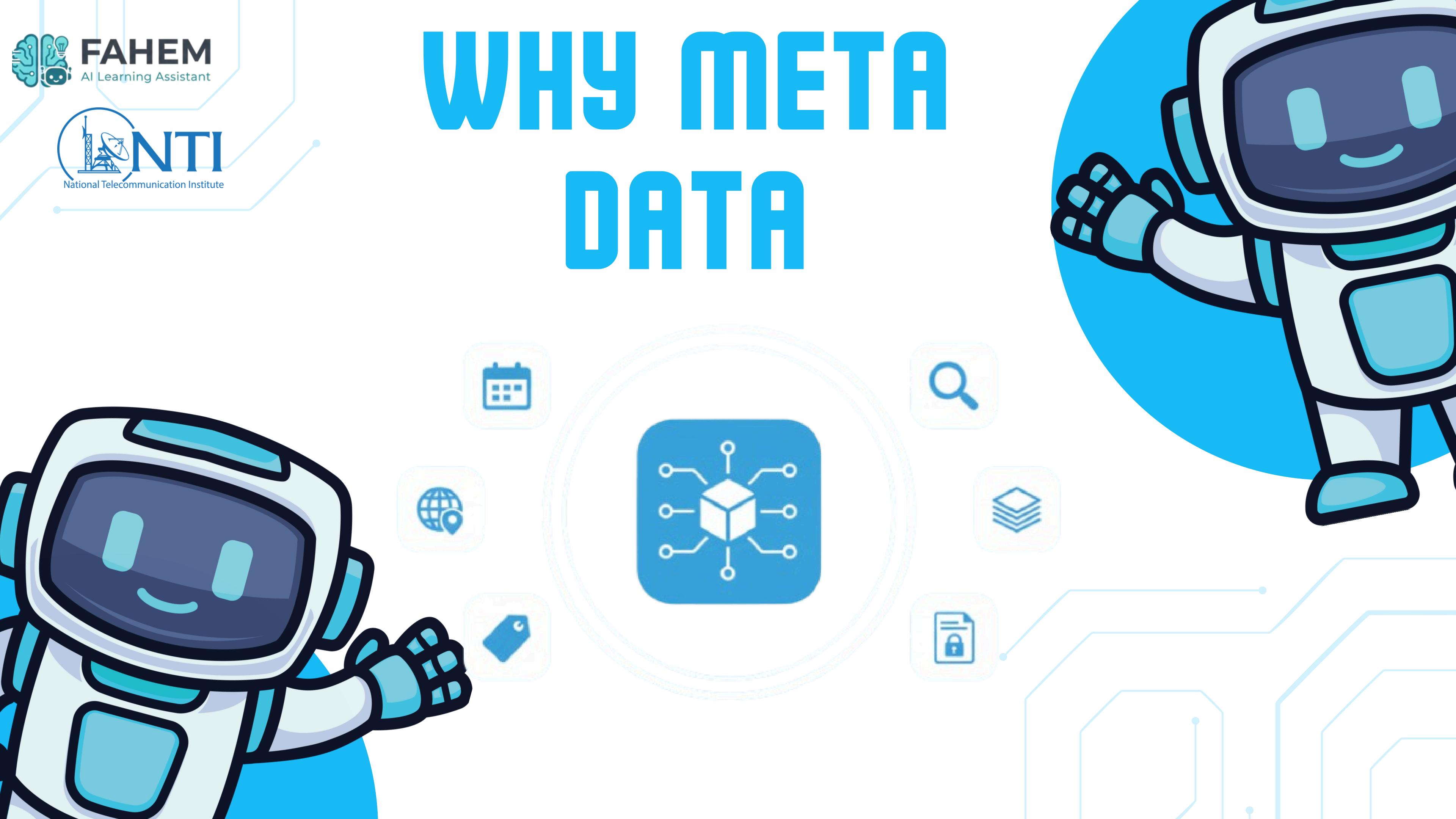
**USING COMPLEX RETRIEVAL SYSTEMS  
THAT RELY ON OCR, DOCUMENT LAYOUT  
RECOGNITION, CHUNKING STRATEGIES,  
FIGURE CAPTIONING AND POWERFUL  
TEXT EMBEDDING MODELS**



**JUST EMBED THE IMAGE**

# COLPALI





# WHY META DATA

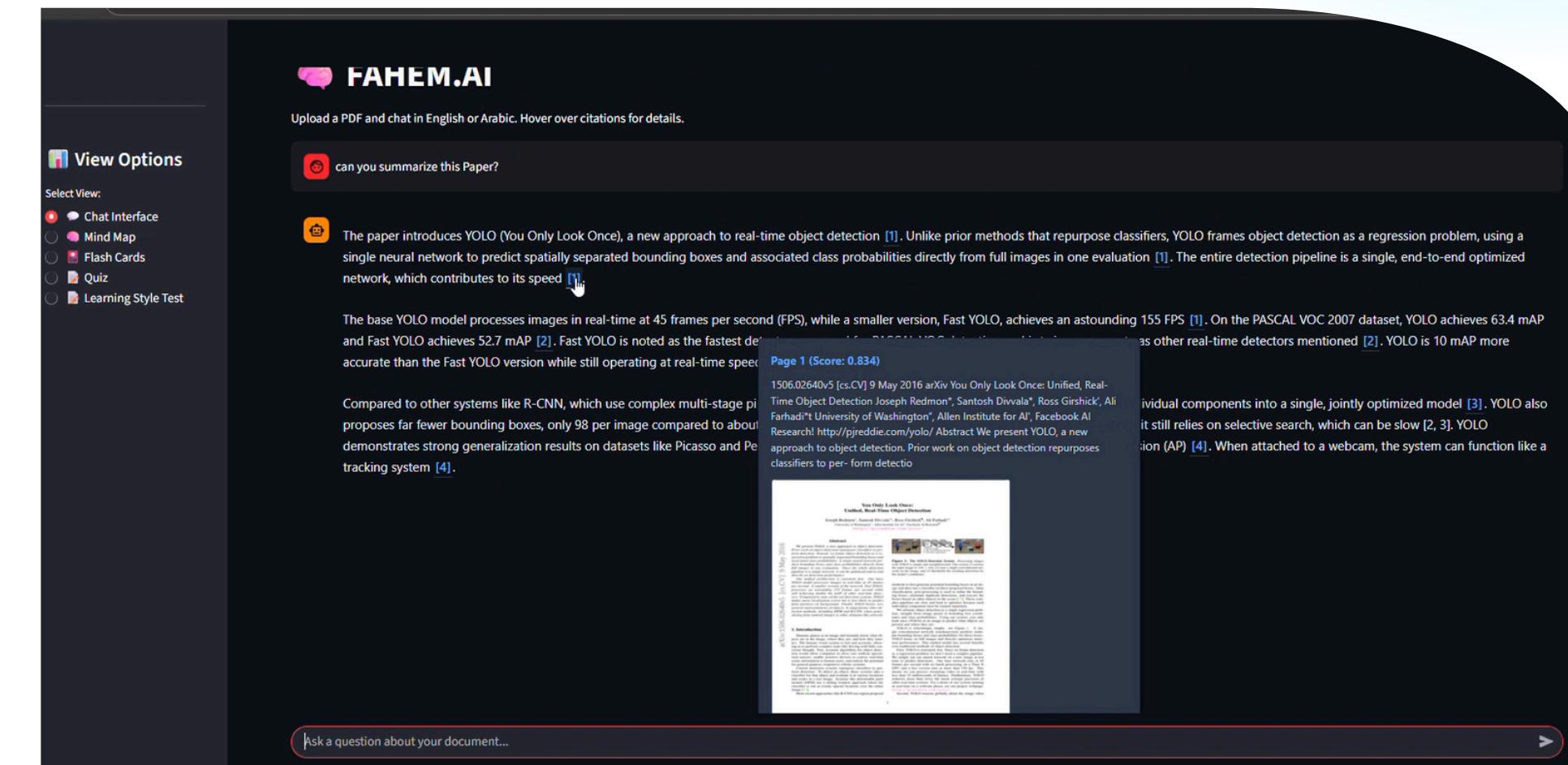
# METADATA INTEGRATION in RAG SYSTEM

📄 Colpali & OCR → extract text & page images from PDFs

📚 Qdrant Vector DB → store embeddings + metadata

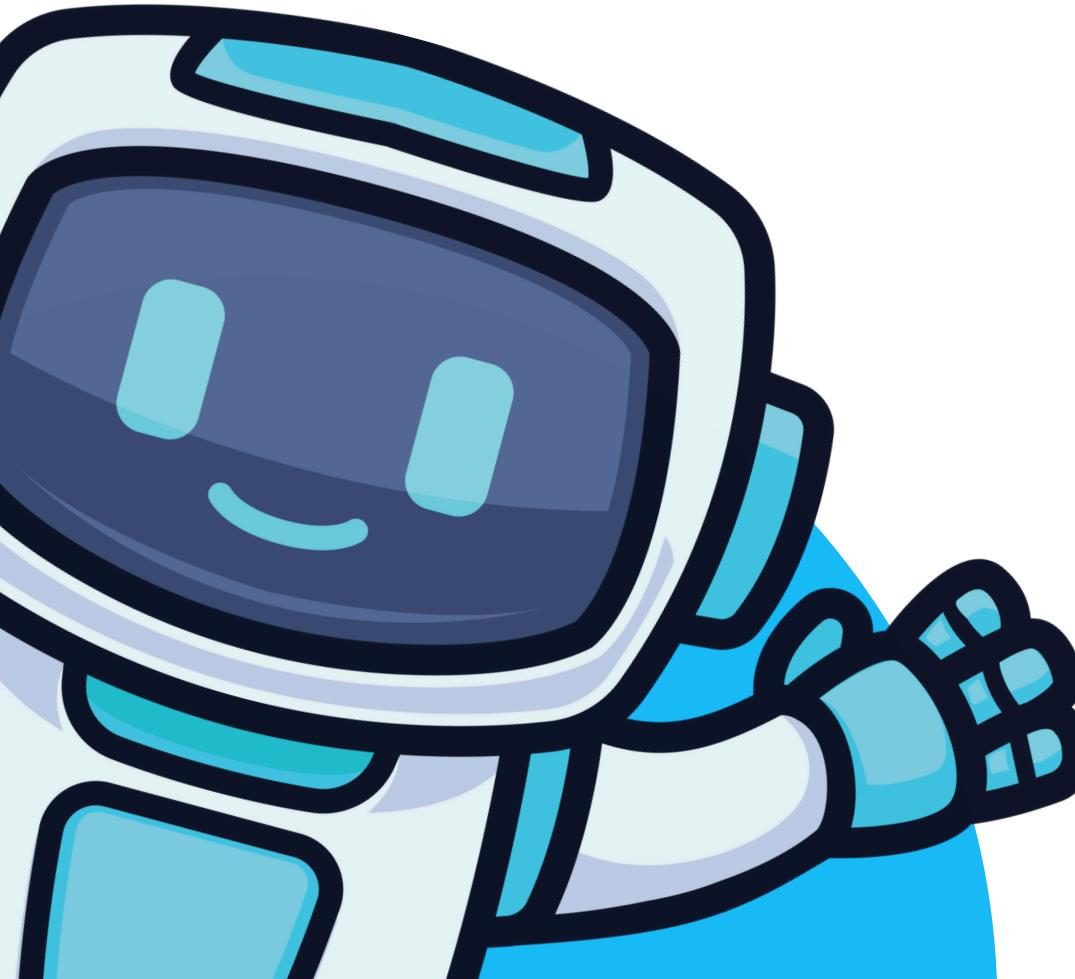
🧩 RAG pipeline → retrieval process

🧠 Gemini API → for answer generation





# AI AVATAR



# AI AVATAR

## PIPELINE STEPS:

🎧 INPUT: TTS-GENERATED  
AUDIO +AVATAR VIDEO

*Output: Natural, fully synchronized talking avatar.*

## PROCESS:

- *Wav2Lip aligns facial movements with audio phonemes.*
- *lipsync refines timing and handles frame-wise synchronization.*



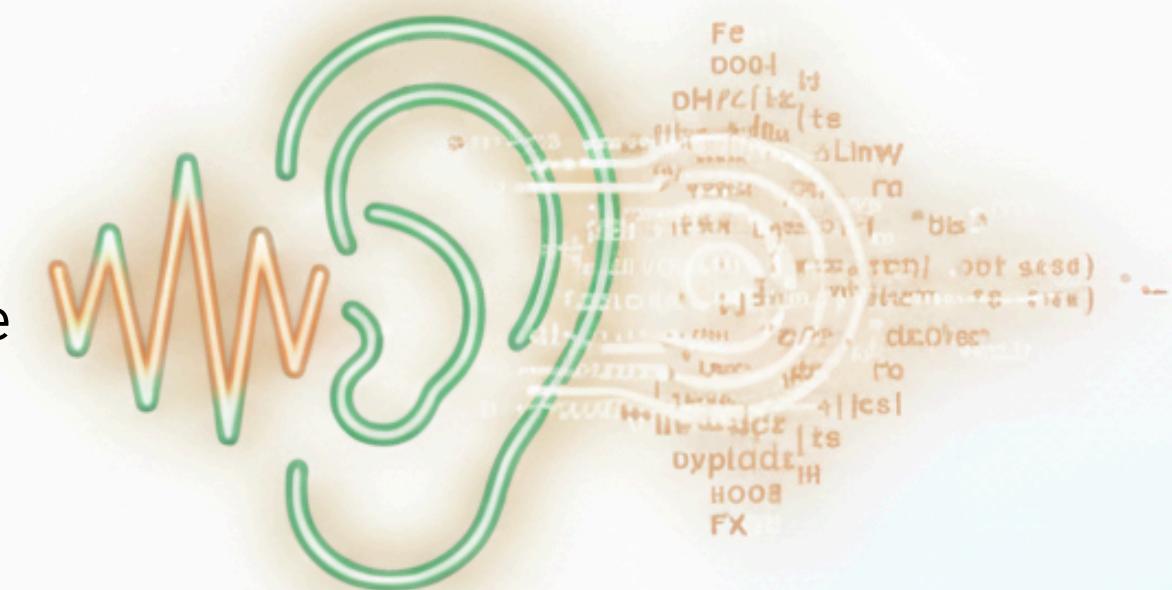


# TEXT TO SPEECH



# TEXT - TO - SPEECH

- Converts extracted text (summaries, explanations) into high-quality audio.
- Uses Gemini TTS for natural, human-like voice output for an immersive and easy listening experience.
- The generated audio is ready for real-time streaming or can be packaged as podcast episodes

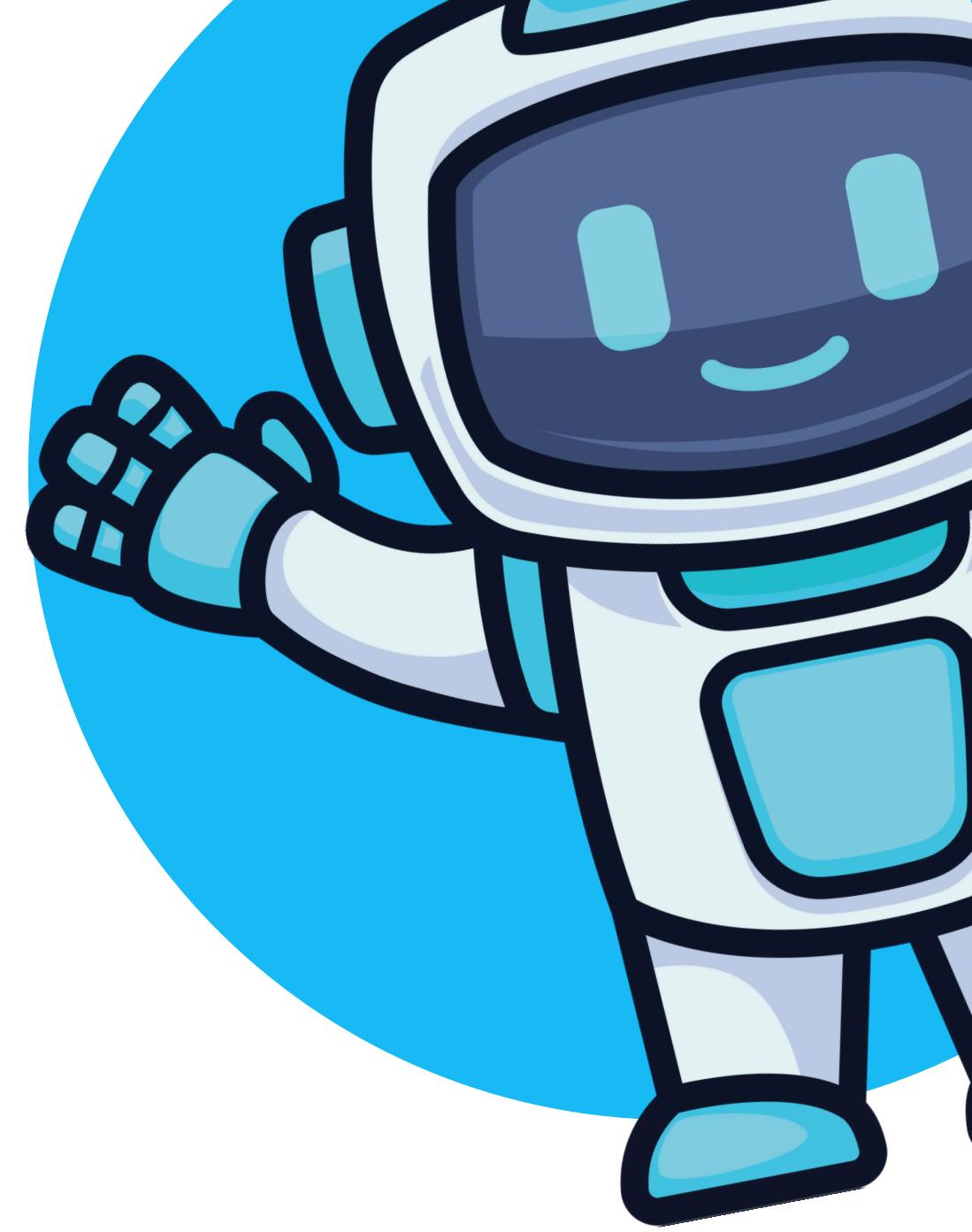
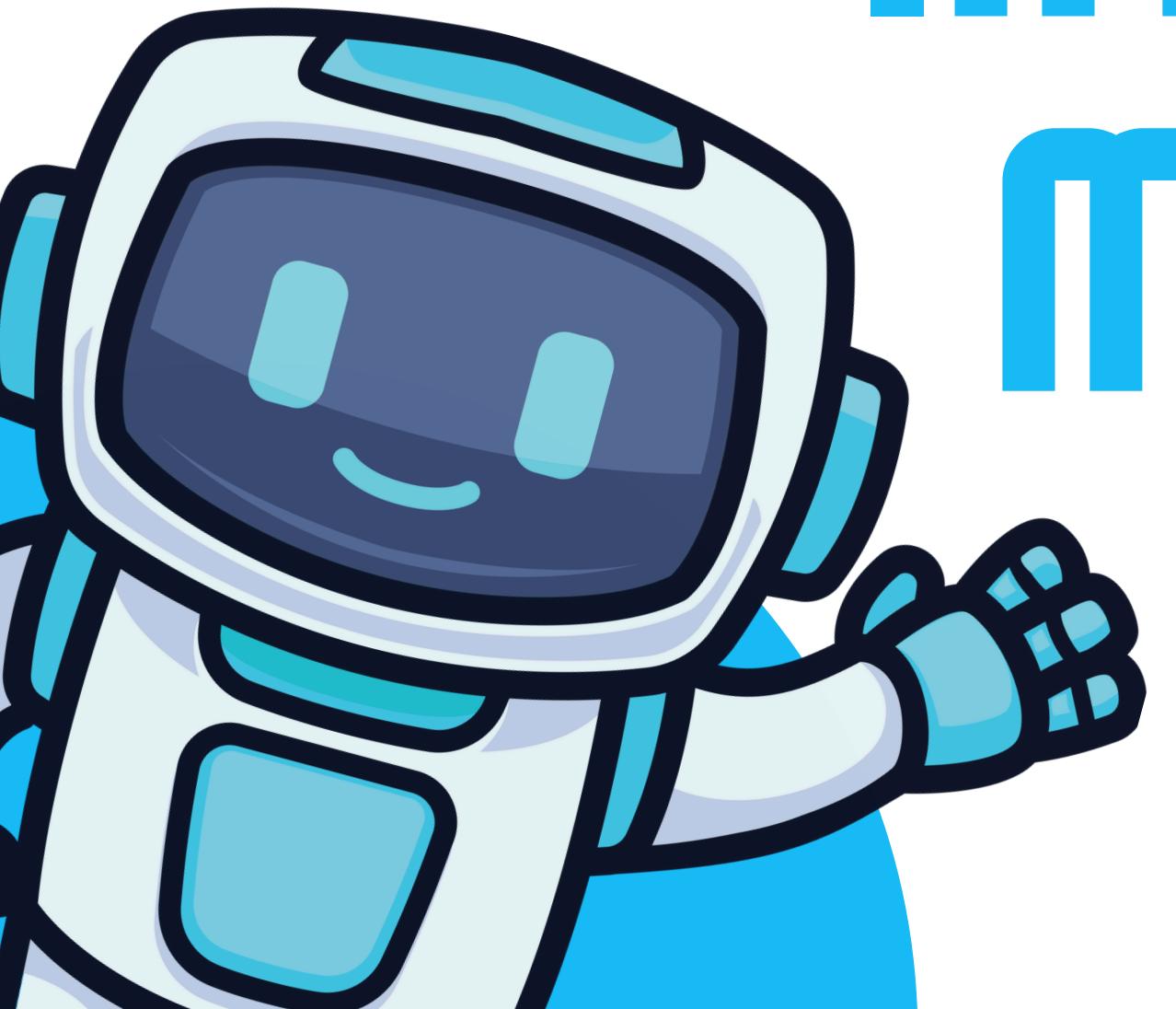


## Gemini



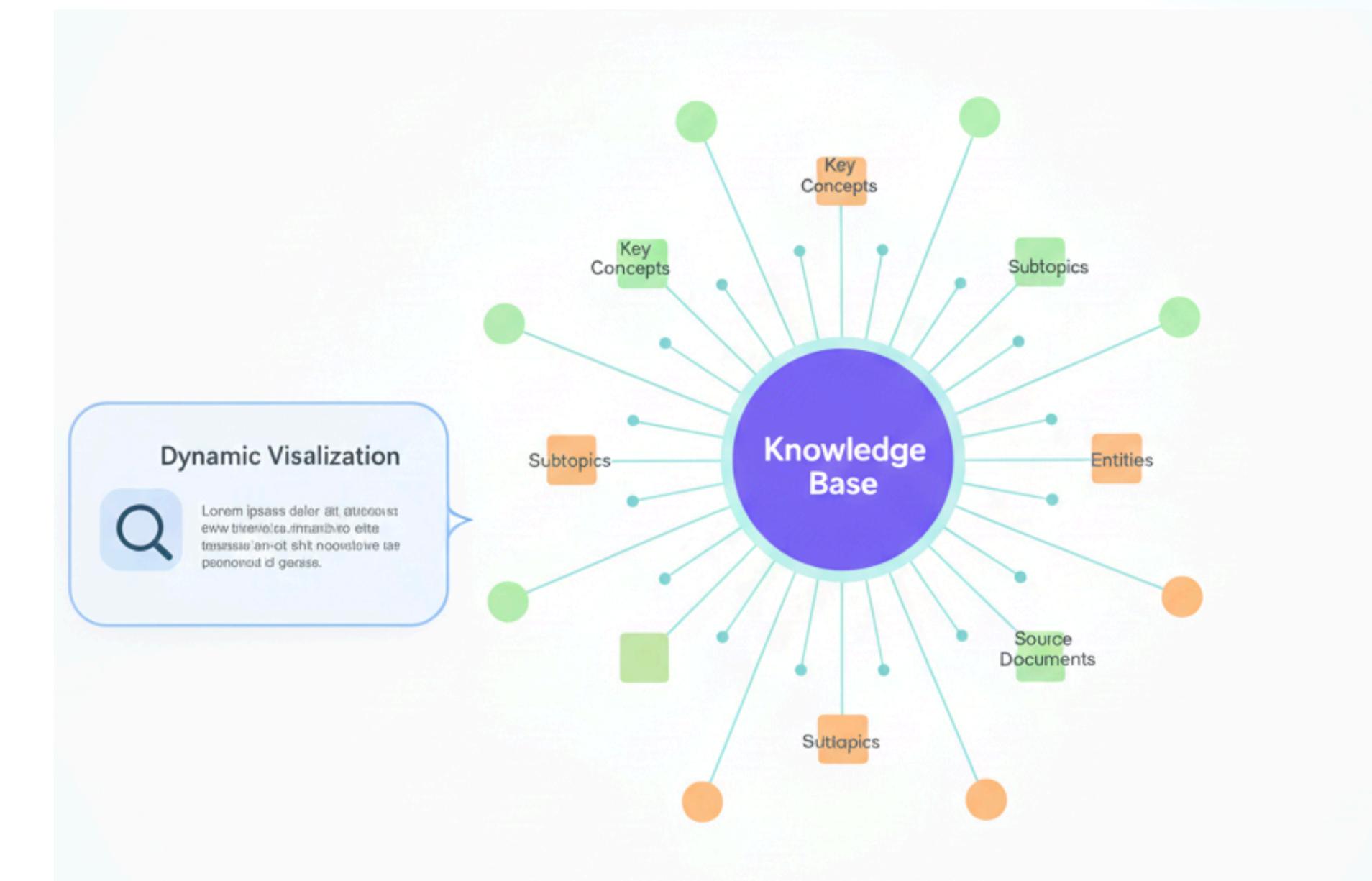


# INTERACTIVE MIND MAP



# INTERACTIVE MIND MAP

- ➡ Systematically analyzes document using LLM (Gemini) to build a hierarchical structure.
  
  
  
- ➡ Dynamic visualization: Interactive zoom, pan, and node expand/collapse.
  
  
  
- ➡ Direct link to source: Click a node to view the original OCR text and page image.
  
  
  
- ➡ Leverages Qdrant Vector DB for content analysis and retrieval.





# FLASHCARDS & QUIZ



# FLASHCARDS & QUIZ

- Generates instant MCQ and True/False quizzes from the document's content.
- Creates flip-style Flashcards (Question/Answer) for quick revision.
- **Intelligent Grading:** Gemini provides automated, context-based feedback for incorrect answers.
- Directly integrated with the RAG knowledge base for highly relevant content.





# AUTOMATIC KNOWLEDGE ROUTING





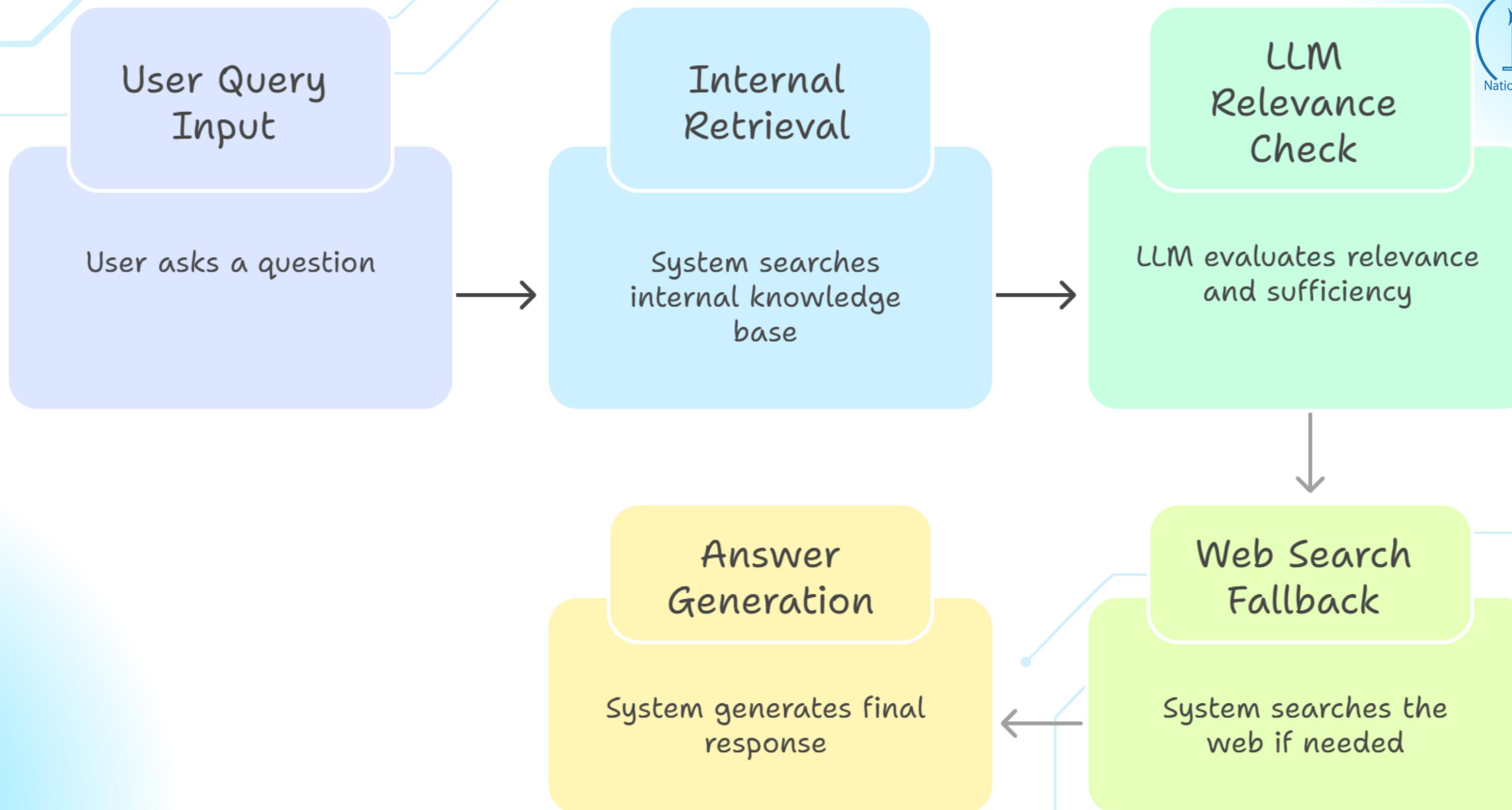
## RAG + Web Search Agent with Automatic Knowledge Routing

**External Knowledge**  
Performs live web searches for additional context



**Internal Knowledge**  
Uses embeddings and retrievers to search stored knowledge

# AI Query Processing Pipeline





# FUTURE WORK



# FUTURE WORK



## Platform Scalability

- Develop FAHEM as a multi-user scalable web app.
- Implement secure user accounts with login and registration.
- Add a database to store user profiles & learning preferences.
- Enable users to upload, save, and manage multiple documents.
- Overcome upload limits using data chunking and streaming.
- Use similarity function of Colpali



## Advanced Content Generation

- Introduce text-to-image generation for visual learning aids.
- Focus on generating Scalable Vector Graphics (SVG) for clear, editable diagrams and flowcharts.



## Understanding Depth

- Enhance the quiz system with manual difficulty selection.
- Add open-ended quiz questions to assess deep understanding.
- Integrate a semantic similarity model to grade text answers intelligently.



## Interactive avatar will serve as a virtual tutor

- Display facial expressions and emotions synchronized with the spoken content
- React Agent



# PROJECT DEMO



## View Options

Select View:

- Chat Interface
- Mind Map
- Flash Cards
- Quiz
- Learning Style Test

Upload a PDF

Drag and drop file here

Limit 200MB per file • PDF

Browse files



Upload a PDF and chat in English or Arabic. Hover over citations for details.



Hello! Upload a PDF to start chatting. Ask me anything in English or Arabic!

Type your message...





yolo المختصر



يستخدم شبكة عصبية واحدة [\[1\]](#) مشاركة متغيرات (regression problem) هو نهج جديد للكشف الكائنات، حيث يعالج اكتشاف الكائنات كمشكلة انحدار (regression problem) حيث يعالج إصدار YOLO [2] في وقت الفحص، حيث يأخذ إصدار YOLO [3] الإحاطة واحتمالات الفئة من الصورة الكاملة.

### تقبل معبراته الرئيسية ما يلي

- \* ما يصل إلى الألفين الصور في الوقت الفحص، يصل إلى 45 إطارًا في الثانية، ويصل إلى 45 إطارًا في الثانية، حيث يأخذ إصدار YOLO [2] في وقت الفحص، حيث يأخذ إصدار YOLO [3] الإحاطة واحتمالات الفئة من الصورة [3].
- \* كل من YOLO v1 وYOLO v2 لا يأخذ الارتفاعات المروية [2]، في [3].
- \* فوق一切 طرق الكشف، مثل YOLO v3، يأخذ الارتفاعات المروية [3]، في [3].
- \* عند تطبيقه على مجالات جديدة أو مدخلات غير منتظمة الشاملة [3] يمكن تحسين خط أليس الكشف بالكامل [3].

عده الكائنات الصغيرة [1] يحقق YOLO ومع ذلك، لا يزال Fast YOLO 52.7 mAP دقة 45 mAP

### Page 1 (Score: 0.863)

1506.03640v5 [cs.CV] 9 May 2016 arXiv You Only Look Once: Unified, Real-Time Object Detection Joseph Redmon\*, Santosh Divvala\*, Ross Girshick\*, Ali Farhadi† University of Washington\*, Allen Institute for AI, Facebook AI Research! [Abstract](http://pjreddie.com/yolo/) We present YOLO, a new approach to object detection. Prior work on object detection repurposes classifiers to per-frame detection.

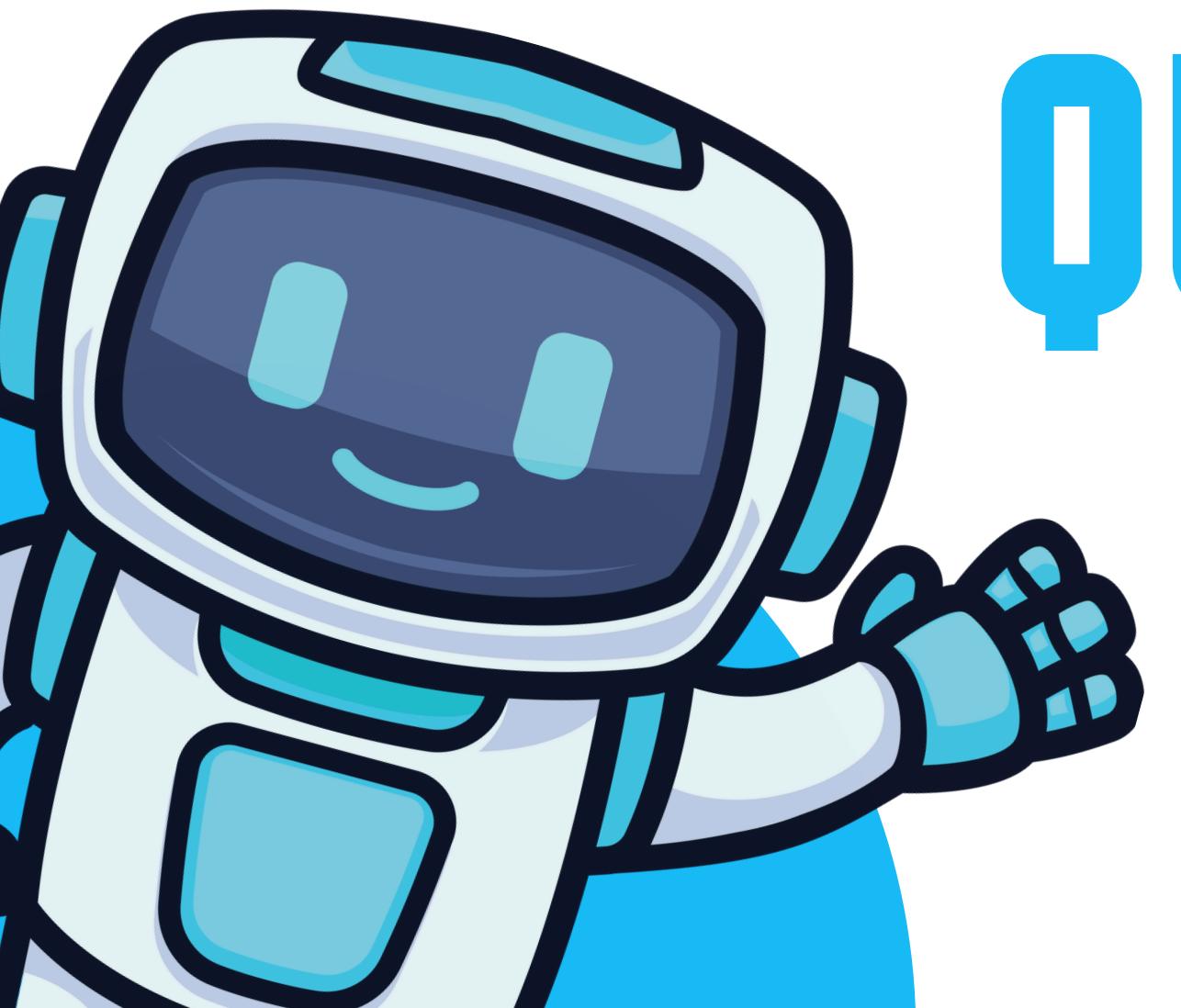


معدل YOLO 63.4 mAP 45 مختلف عن أنظمة

الصور بأكمامها [3]، مما يزيد من الدقة [2].  
في المقابل، فإن صور طبيعية واحدة على أعمال فنية [3].



# any questions?





# THANK YOU

[www.Fahem.com](http://www.Fahem.com)

