



Agentic AI

M1: Introduction to agentic workflows



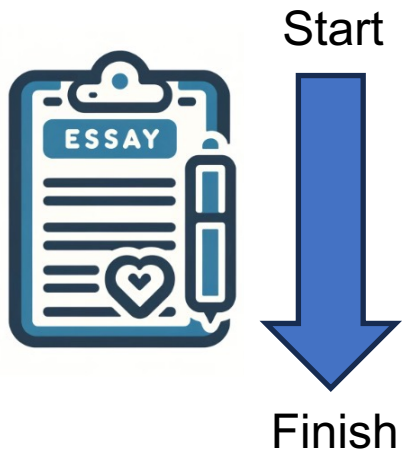
Introduction to Agentic Workflows

What is Agentic AI?

Agentic AI

Non-agentic workflow (zero-shot):

Please type out an essay on topic X from start to finish in one go, without using backspace.



Agentic workflow:

Write an essay outline on topic X

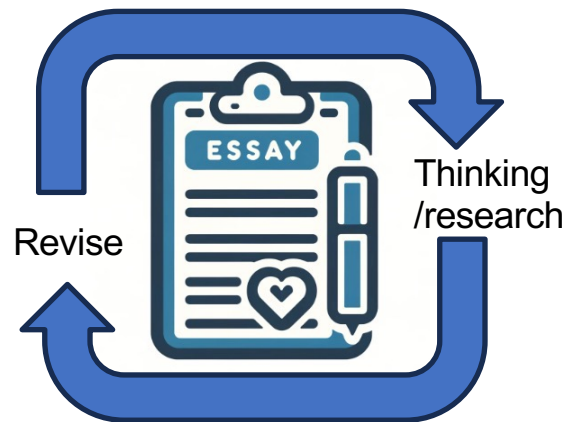
Do you need any web research?

Write a first draft.

Consider what parts need revision or more research.

Revise your draft.

....



Agentic AI workflows

An agentic AI workflow is a process where an LLM-based app executes multiple steps to complete a task.

Essay-writing example:

Write an essay outline on topic X

LLM

Do you need any web research?

LLM

+

web search

Write a first draft.

LLM

Consider what parts need revision or more research.

LLM

+

request
human review

Revise your draft.

LLM



Introduction to Agentic Workflows

Degrees of Agenticness



Sequoia Ascent, March 2024



....

Rather than arguing over which work to include or exclude as being a true agent, we can acknowledge that there are different degrees to which systems can be agentic.

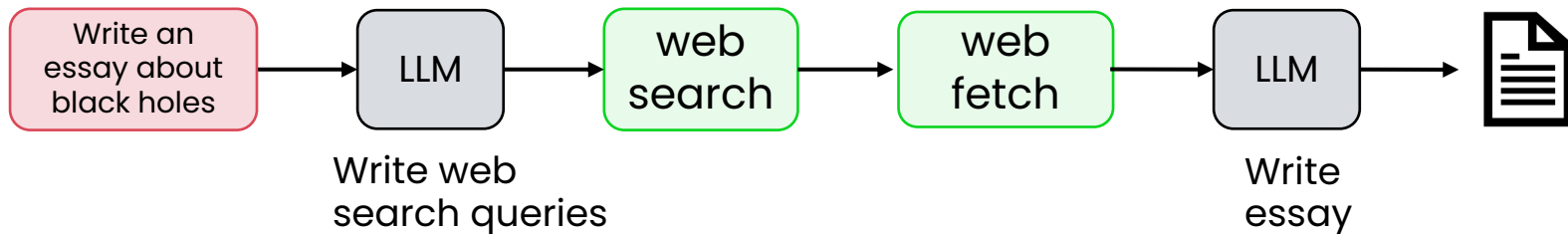
....



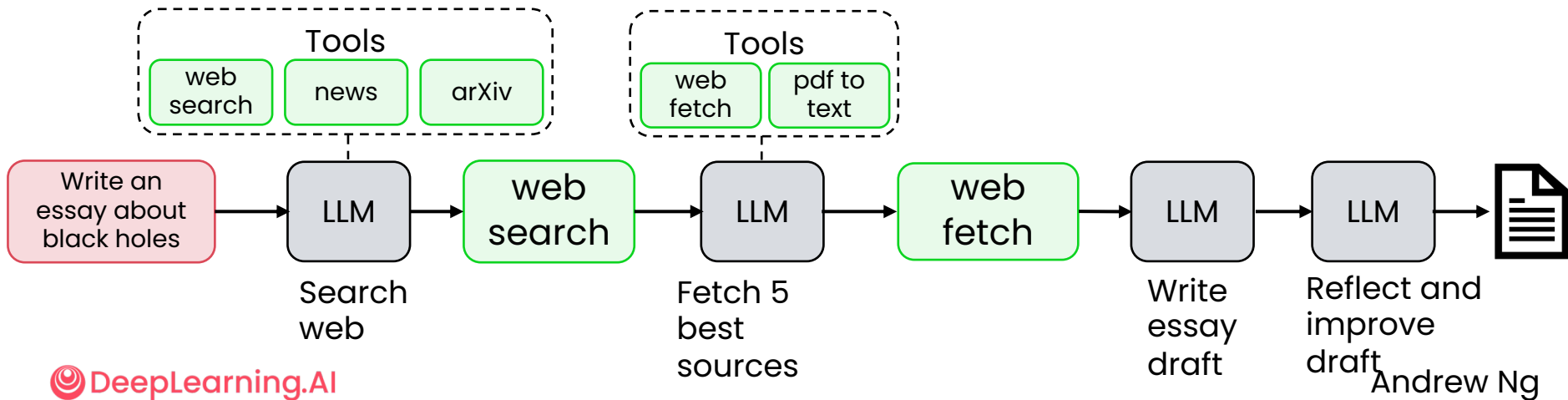
X (twitter) post , June 2024

Degrees of autonomy

Less autonomous

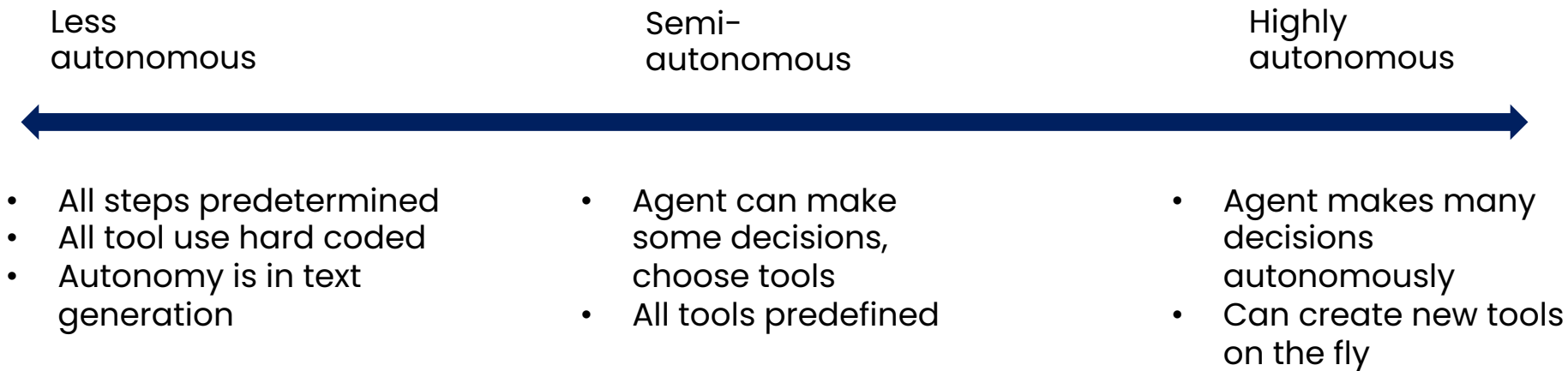


More autonomous



Degrees of “agenticness”

Agentic AI can be less or more autonomous

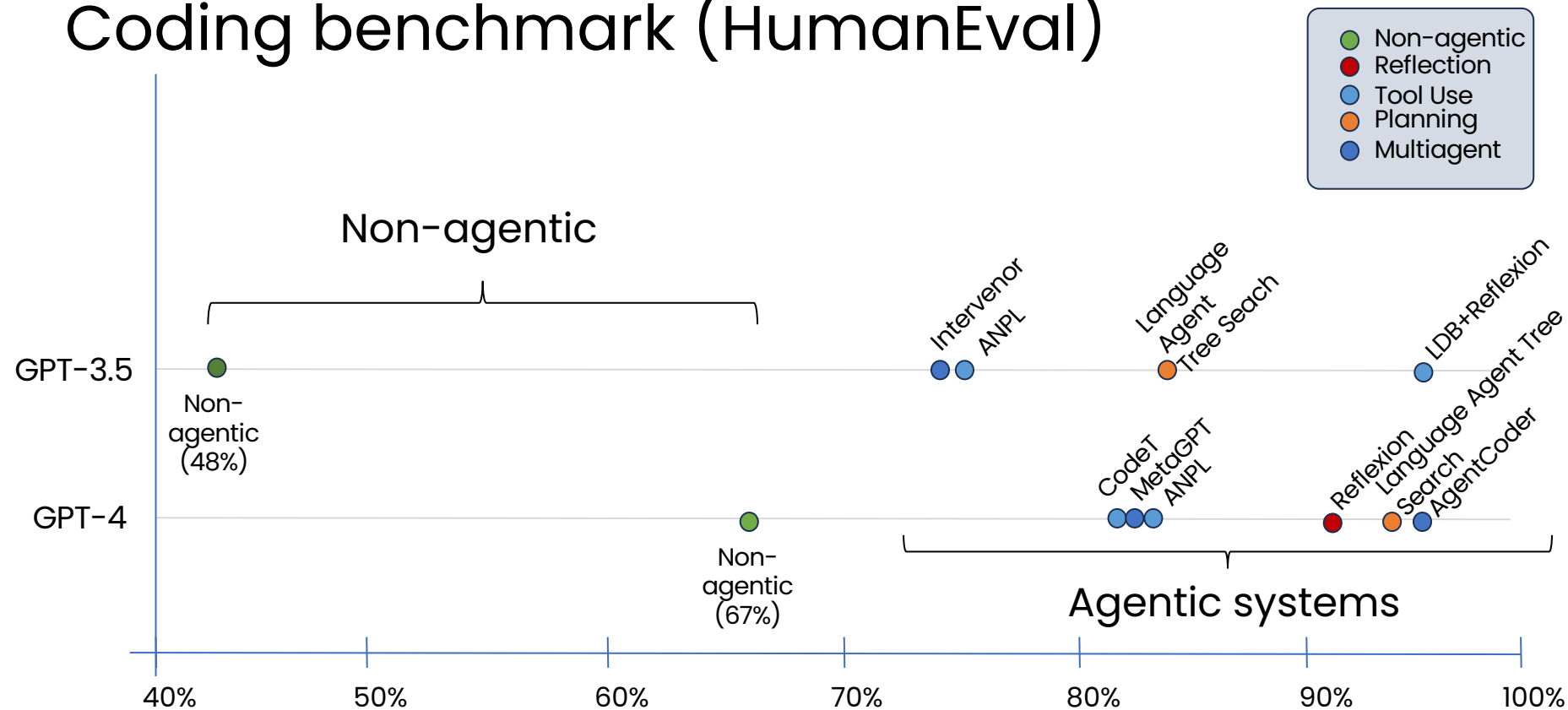




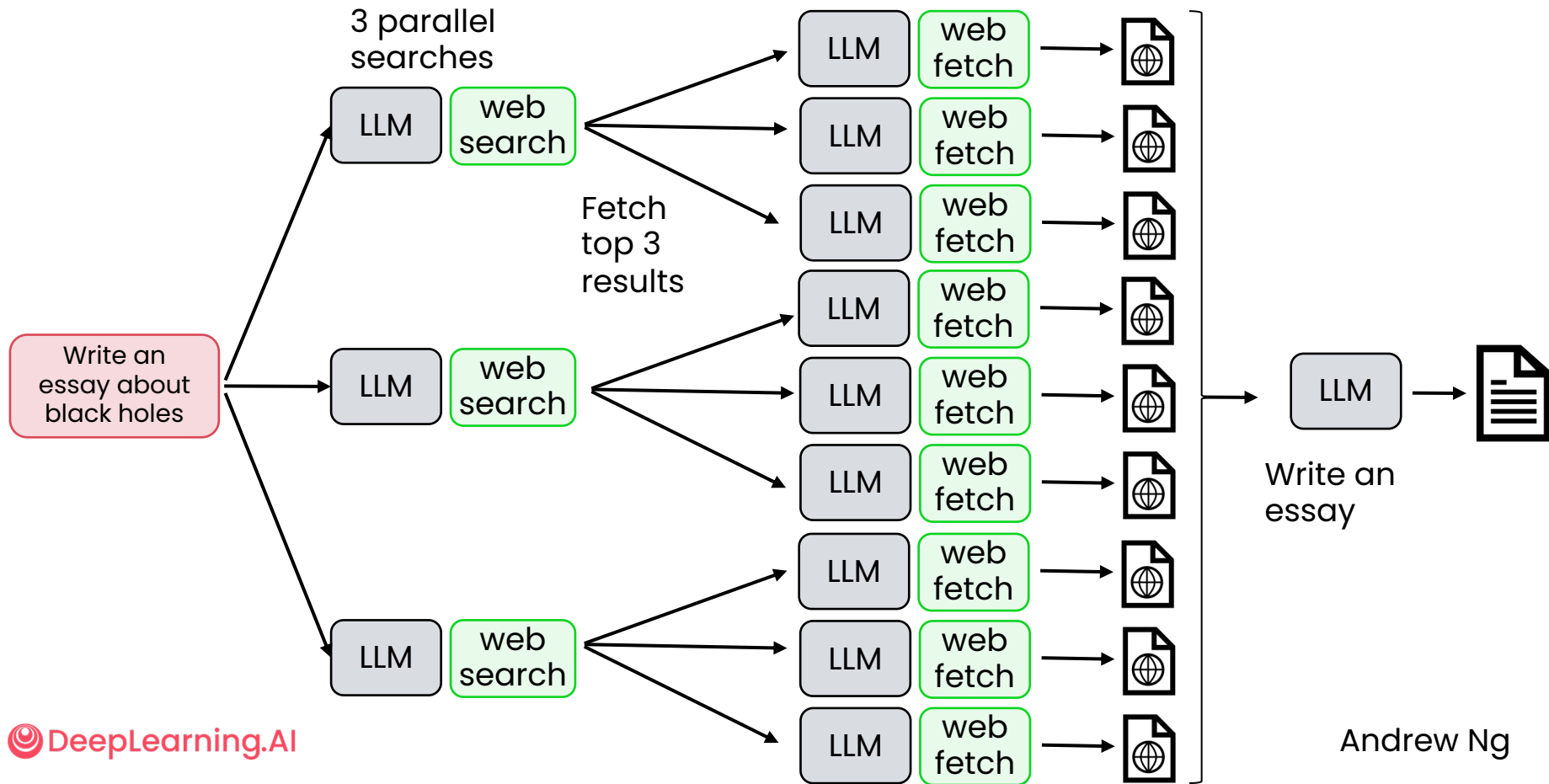
Introduction to Agentic Workflows

Benefits of agentic AI

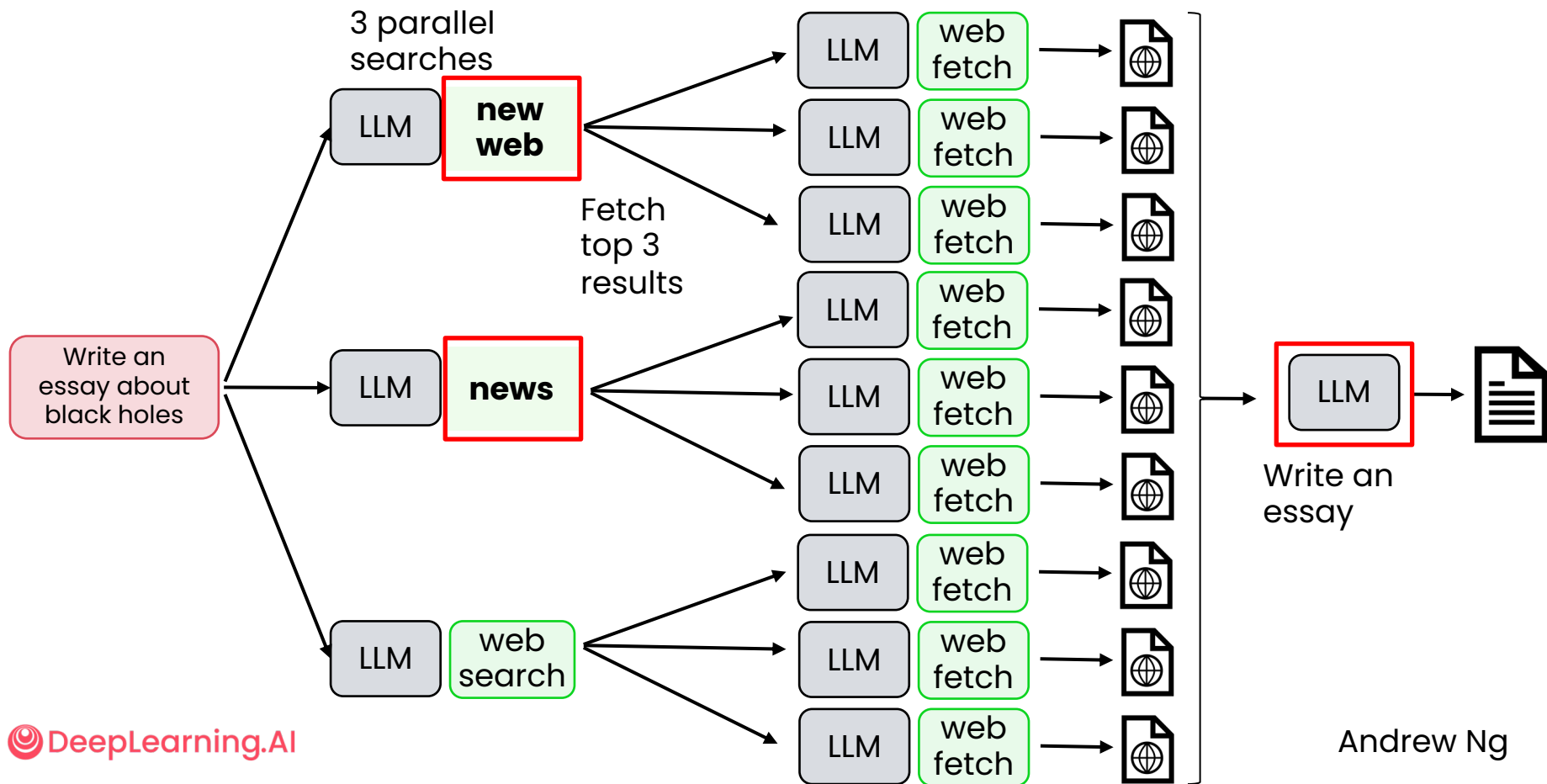
Coding benchmark (HumanEval)



Parallelization for speed

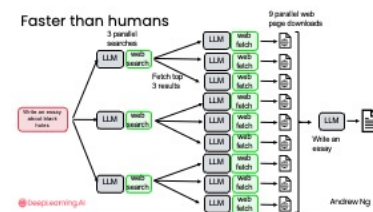
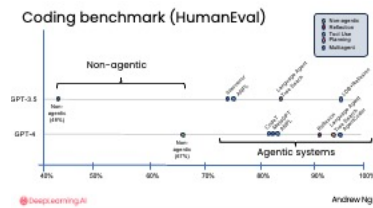


Modular: Add/swap out components



Key benefits of agentic workflows

- Much better performance
- Faster than humans because of parallelization
- Modular: can add or update tools, swap out models





Introduction to Agentic Workflows

Agentic AI applications

Example: Invoice processing workflow

TechFlow Solutions LLC

890 Juniper Drive
San Mateo, CA 94401
Phone: (415) 555-7890
Email: billing@techflowsol.com

Due Date: August 20, 2025

Invoice Date: August 6, 2025

Description	Qty	Unit Price	Line Total
Consulting - Systems Integration (hrs)	20	\$150.00	\$3,000.00

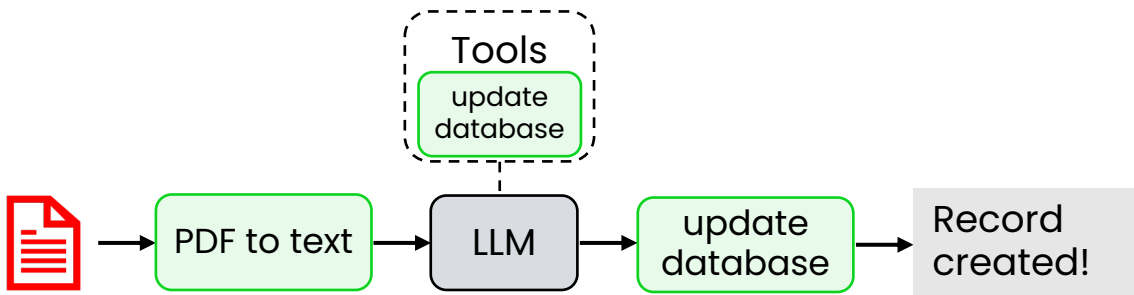
Total Due: \$3,000.00

4 required fields:

Biller
Biller address
Amount due
Due date

Steps:

1. Identify required fields
2. Record in database



Example: Responding to customer email

From: Susan Jones
Subject: Wrong item shipped

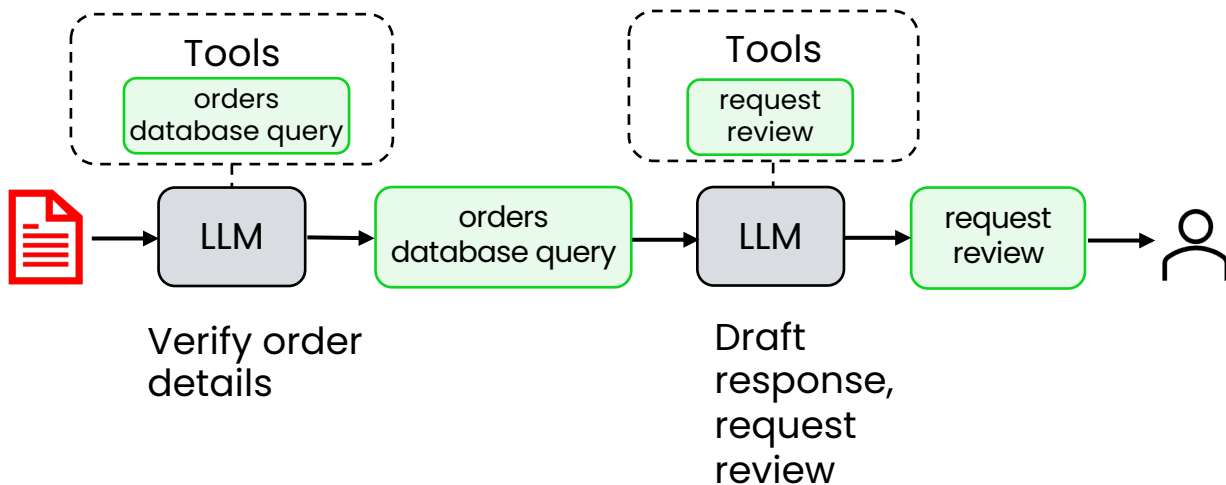
I ordered a blue KitchenPro blender (Order #8847) but received a red toaster instead.

I need the blender for my daughter's birthday party this weekend. Can you help?

Susan

Steps:

1. Extract key information
2. Find relevant customer records
3. Draft response for human review



More challenging: Customer service agent

Do you have any black jeans or blue jeans?

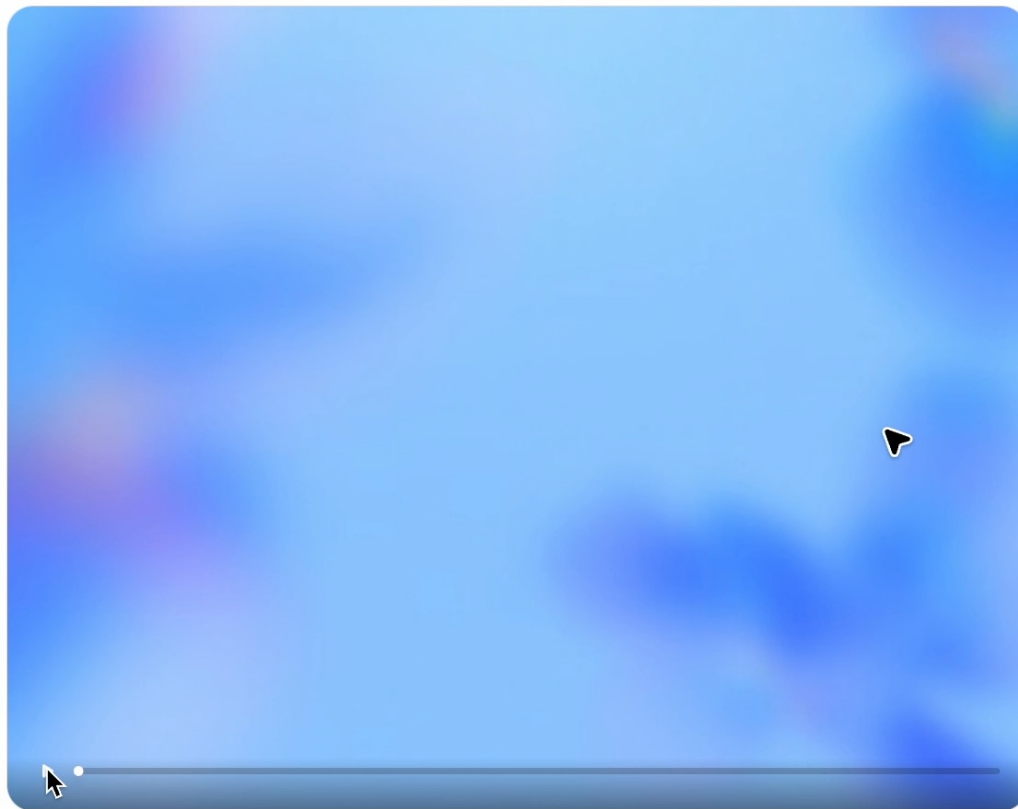
1. Check inventory for black jeans
2. Check inventory for blue jeans
3. Respond to customer

I'd like to return the beach towel I bought

1. Verify customer purchase
2. Check return policies
3. If return allowed = "yes", then:
 - a. Issue return packing slip
 - b. Set database record to "return pending"

Required steps
not known ahead
of time

Difficult: Visual computer use



ChatGPT Agent Mode
(OpenAI)

What tasks is agentic AI suited to?

Easier

Harder



Clear, step-by-step process

Steps not known ahead of time

Standard procedures to follow

Plan/solve as you go

Text assets only

Multimodal (sound, vision)



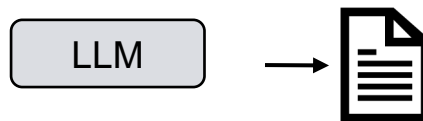
Introduction to Agentic Workflows

Task decomposition:
Identifying the steps in
a workflow

Example: Writing an Essay

Write an essay on topic X

1. Write an essay outline on topic X
2. Search web
3. ~~Write the essay~~ Write a first draft
4. Consider what parts need revision
5. Revise your draft



Recap: Breaking down essay writing into steps

Direct generation:

1. Write an essay on topic X

Not good enough...

3-step workflow:

1. Write an essay outline on topic X
2. Search web
3. Write the essay

Still not good enough...

5-step workflow:

1. Write an essay outline on topic X
2. Search web
3. Write the first draft
4. Consider what parts need revision
5. Revise your draft

Example: Responding to customer email

From: sjones9@email.com
Subject: Wrong item shipped

Hi, I ordered a blue KitchenPro blender (Order #8847) but received a red toaster instead.

I need the blender for my daughter's birthday party this weekend. Can you help?

Susan Jones

1. Extract key information LLM
2. Find relevant customer records LLM + orders database query
3. Write and send response LLM + send email

Example: Extracting information from invoice

TechFlow Solutions LLC

890 Juniper Drive
San Mateo, CA 94401
Phone: (415) 555-7890
Email: billing@techflowsol.com

Due Date: August 20, 2025

Invoice Date: August 6, 2025

Description	Qty	Unit Price	Line Total
Consulting - Systems Integration (hrs)	20	\$150.00	\$3,000.00
Total Due:			\$3,000.00

1. Find required information

LLM

LLM

+



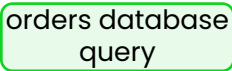

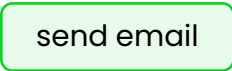
update
database

What building blocks do you have?

Building block	Examples	Use cases
Models	LLMs	Text generation, tool use, information extraction
	Other AI models	PDF-to-text, text-to-speech, image analysis
Tools	API	Web search, get real-time data, send email, check calendar,....
	Information retrieval	Databases, Retrieval Augmented Generation (RAG)
	Code execution	Basic calculator, data analysis

Tips for task decomposition

Example: Responding to email

1. Extract key information 
2. Find relevant customer records  + 
3. Write and send response  + 

Questions to ask:

- Can LLM, tool, or both complete each step?
- If “no”, consider breaking down into further steps, or adding human-in-the-loop



Introduction to Agentic Workflows

Evaluating Agentic AI
(evals)

Look for low-quality outputs

From: sjones9@email.com
Subject: Wrong item shipped

Hi, I ordered a blue KitchenPro blender (Order #8847) but received a red toaster instead.

I need the blender for my daughter's birthday party this weekend. Can you help?

Susan Jones

1. Extract key information

LLM

2. Find relevant customer records

LLM

+

orders database
query

3. Write and send response

LLM

+

send email

"I'm glad you shopped with us – we're much better than our competitor CompCo!"

"Sure, we'll issue a refund. Unlike RivalCo, we make returns easy."

Add an evaluation to track the error

Customer
inquiry

We're better
than CompCo...

LLM response

```
if (competitor in response):  
    num_competitor_mentions += 1
```

List of competitors:

CompCo

RivalCo

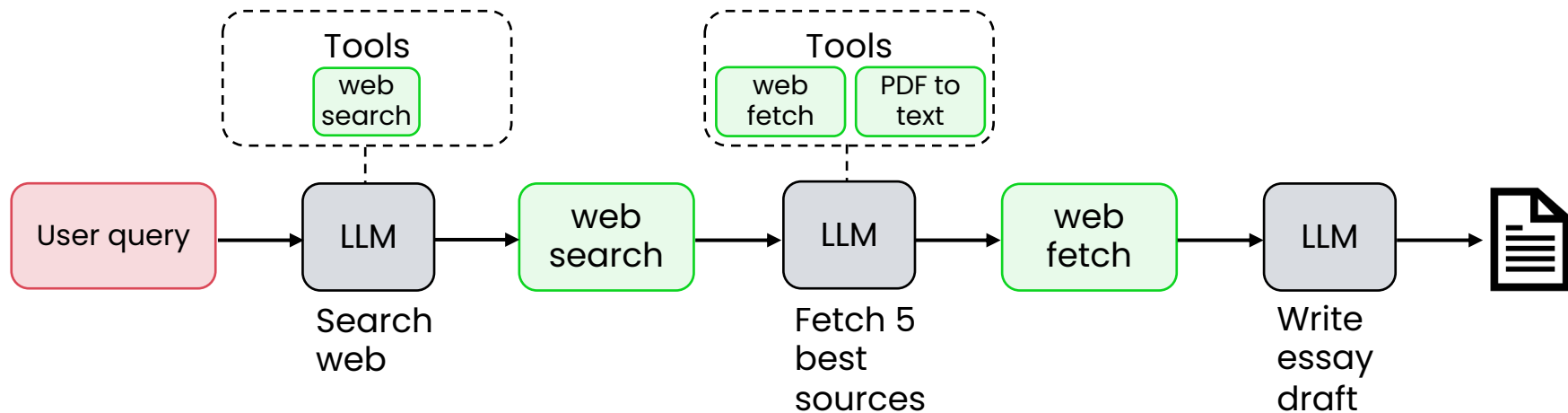
The Other Co

...

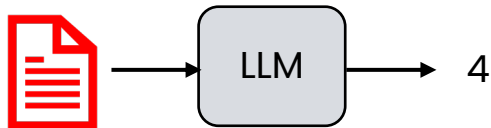
Objective error:

Use code to check for occurrences

Using LLM as a judge



Assign the following essay a quality score between 1 and 5, where 5 is the best: {essay}



Example prompt	LLM-assigned score
Black holes	3
Robotic harvesting	4

Evaluating Agentic AI

- Can evaluate using code (objective evals), or LLM-as-judge (subjective evals)
- Two types of evals: End-to-end and component-level
- Examine traces to perform error analysis
- Much more on evals and error analysis in Module 4!



Introduction to Agentic Workflows

Agentic design patterns

Agentic Design Patterns

1. Reflection
2. Tool use
3. Planning
4. Multi-agent collaboration

Reflection



Please write code for {task}

`def do_task(x): ...`

`def do_task_v2(x):`

`def do_task_v3(x):`



Coder
Agent
(LLM)

Here's code intended for {task}:

```
def do_task (x):  
    ...
```

Check the code carefully for correctness, style and efficiency, and give constructive criticism for how to improve it.

There's a bug on line 5. Fix it by ...

It failed Unit Test 3. Try changing ...

Reflection



Please write code for {task}

`def do_task(x): ...`

`def do_task_v2(x):`

`def do_task_v3(x):`



Coder
Agent
(LLM)

Here's code intended for {task}:

```
def do_task(x):  
    ...
```

Check the code carefully for correctness, style and efficiency, and give constructive criticism for how to improve it.

There's a bug on line 5. Fix it by ...

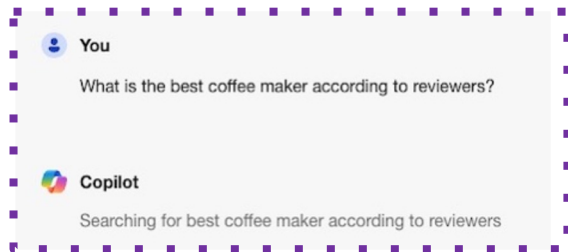
It failed Unit Test 3. Try changing ...



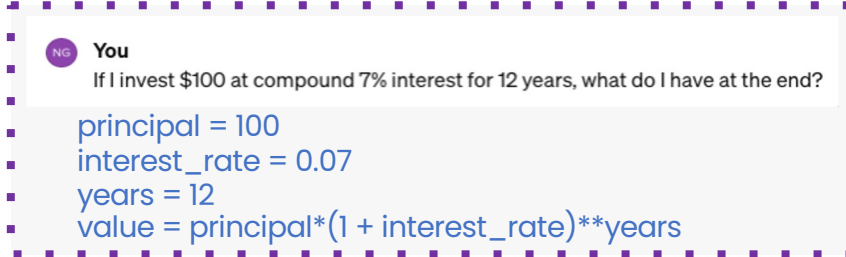
Critic
Agent
(LLM)

Tool use

Web search tool



Code execution tool



Analysis

- Code Execution
- Wolfram Alpha
- Bearly Code Interpreter

Information gathering

- Web search
- Wikipedia
- Database access

Productivity

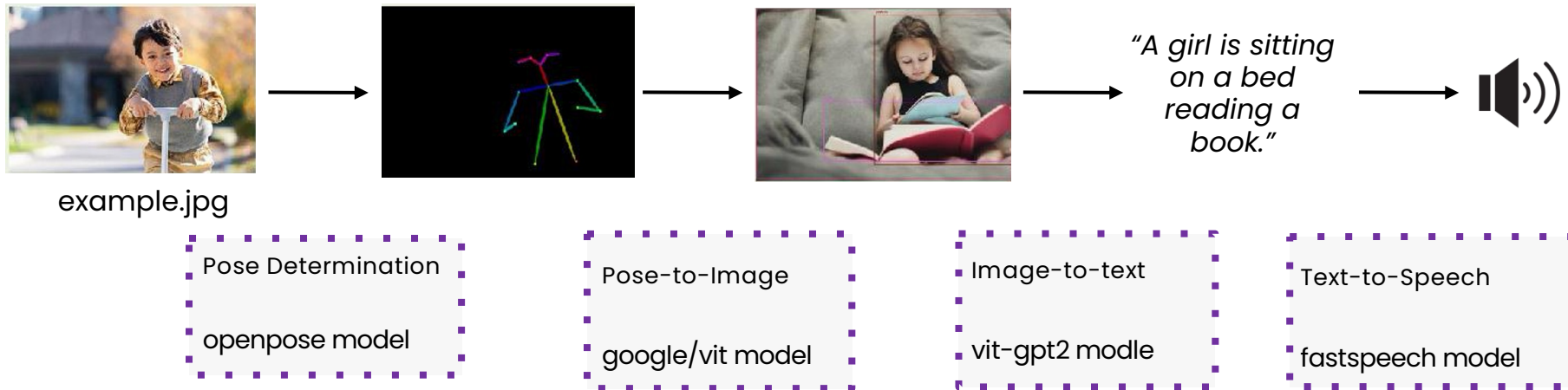
- Email
- Calendar
- Messaging

Images

- Image generation
- Image captioning
- OCR

Planning

Request: Please generate an image where a girl is reading a book, and her pose is the same as the boy in the image example.jpg, then please describe the new image with your voice.



[Example adapted from "HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face", Shen et al. (2023)]

Multi-agent workflows



Multiagent Debate

Task	Single agent	Multi-agent
Biographies	66.0%	73.8%
MMLU	63.9%	71.1%
Chess move	29.3%	45.2%

[From "Improving Factuality and Reasoning in Language Models through Multiagent Debate", Du et al., 2023]



DeepLearning.AI

End of M1
