6CCSAHAI
Human AI Interaction

LECTURE 3 – LGT

Interface and Interaction Patterns for Human-AI Systems

[What do we want to get done today ?]

1. **Deep-dive into critical thinking theories**   [30 min interactive Q&A]
   (Bloom's taxonomy / Socratic questioning)

2. **LGT – getting grips with Langchain + Streamlit + Mistral:**   [40 min]
   - develop basic chatbot-tutor + evaluating against usability heuristics
     ("Knowledge" layer of Bloom taxonomy)

3. **SGT-task transition:**   [30 min + SGT]
   - extending your tutor with better LLM scaffolding
     (for a higher-up layer of Bloom's taxonomy).

And this is Aubrey!

(Hello!)

She will lead discussions of Socratic questioning/Bloom's taxonomy a bit later. :)

[theme of the day]

Prompt engineering … aka making LLMs do what you want

(especially if what you want is supporting Bloom's style questions)

Bloom's ladder ascends,
Each prompt a rung upward bound,
Wisdom, not answers.

With scaffolds and questions so bright,
The chatbot climbed Bloom's cognitive height,
"Don't give them the answer,
Make thinking the dancer!"
Now students analyze, compare, cite

[LGT task – work in pairs]

Your mission:
create a basic tutor-bot for Knowledge level of Bloom's framework

Step 1: Take <u>up to 20 minutes</u> to get a one-prompt tutoring chatbot to compile & run reliably.

Your goals are to:
- Allow the user to input a <u>paragraph of text / topic</u> that LLMs should know
- Your bot will ask appropriate <u>socratic questions</u> on the Knowledge layer of Bloom's framework.

<u>You will be sharing this bot with another team after the 15 minutes for testing!</u>

[LGT task – work in pairs / triads]

Your mission:
create a basic tutor-bot for Knowledge level of Bloom's framework

Step 2: Spend <u>10 minutes</u> testing your colleagues' prototype.

As part of the testing, keep a short 'bullet-point' report of your findings to share with the other team. Try to ground your feedback around the 10 usability heuristics from the lecture.

Spend <u>10 minutes</u> discussing your findings with the other team (and listening to their comments!)

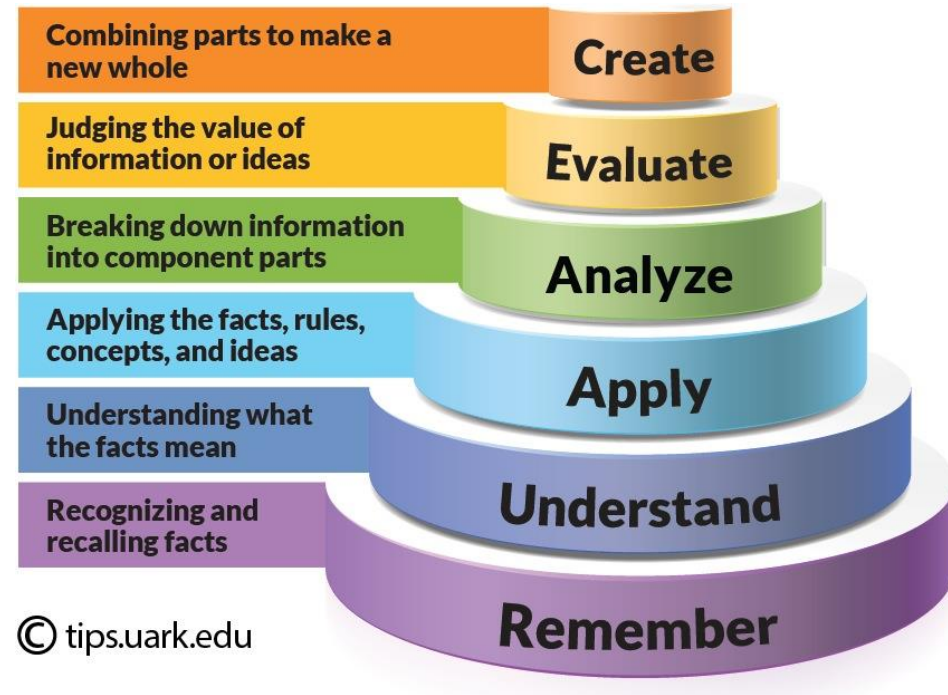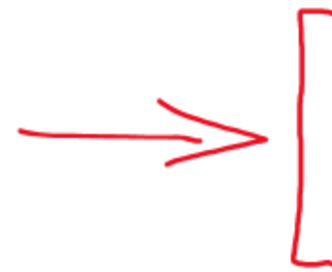# *designing* for **usability**

Jakob Nielsen's 10 usability heuristics of UI design

1. Visibility & Feedback

2. Match between system & the real world

3. User control & freedom

4. Consistency & standards

5. Safety: Prevent Errors

6. <u>Recognition</u> rather than <u>recall</u>

7. Efficiency, and flexibility

8. Aesthetics & minimalism

9. Recognise, Diagnose, Recover from Errors

10. Help and Documentation

[SGT task – work in pairs / triads]

**Mission: design an updated bot that tutors on a more complicated layer of Bloom's framework!**

Tutor me about <a selected topic> using questions based in <a selected Bloom's layer>.

Combining parts to make a new whole — **Create**

Judging the value of information or ideas — **Evaluate**

Breaking down information into component parts — **Analyze**

Applying the facts, rules, concepts, and ideas — **Apply**

Understanding what the facts mean — **Understand**

Recognizing and recalling facts — **Remember**

© tips.uark.edu

[SGT task – work in pairs / triads]

Mission: design an updated bot that tutors on a more complicated layer of Bloom's framework!

We're now building a more complicated tool – you will want to think about how you want to scaffold the 'thinking' your bot does.
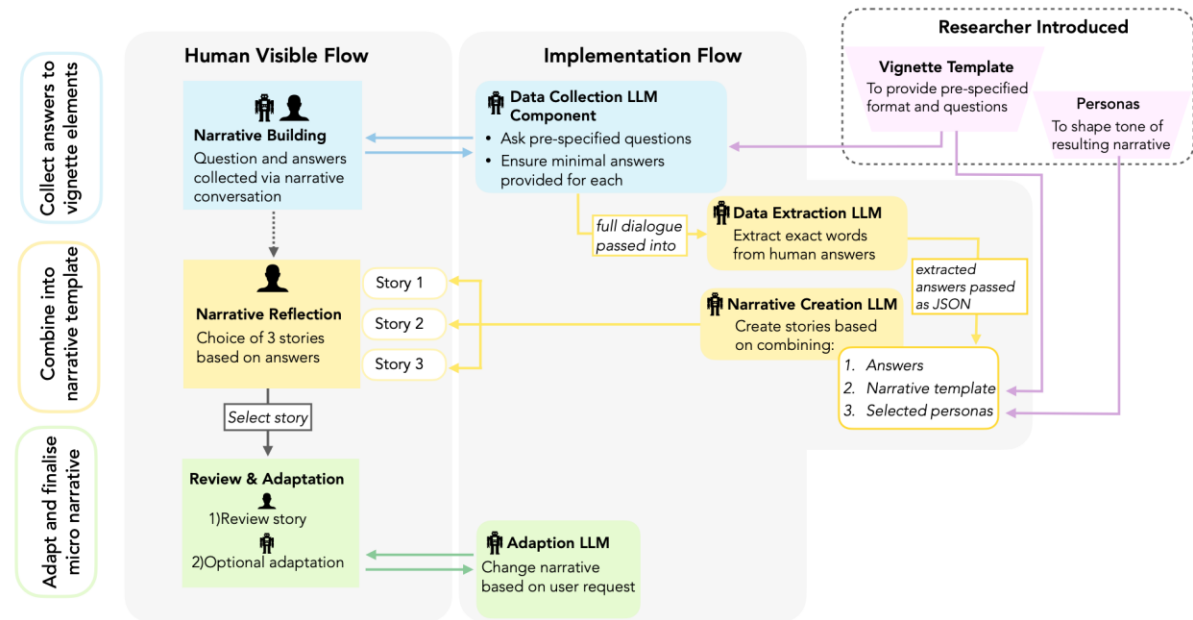
[SGT task – work in pairs / triads]

You can get as complicated as you like …

… but you need to <u>at least</u> include this **evaluator-optimizer loop** in your system

→ that checks if the system suggested question is 'appropriate'

Workflow: Evaluator-optimizer

In the evaluator-optimizer workflow, one LLM call generates a response while another provides evaluation and feedback in a loop.
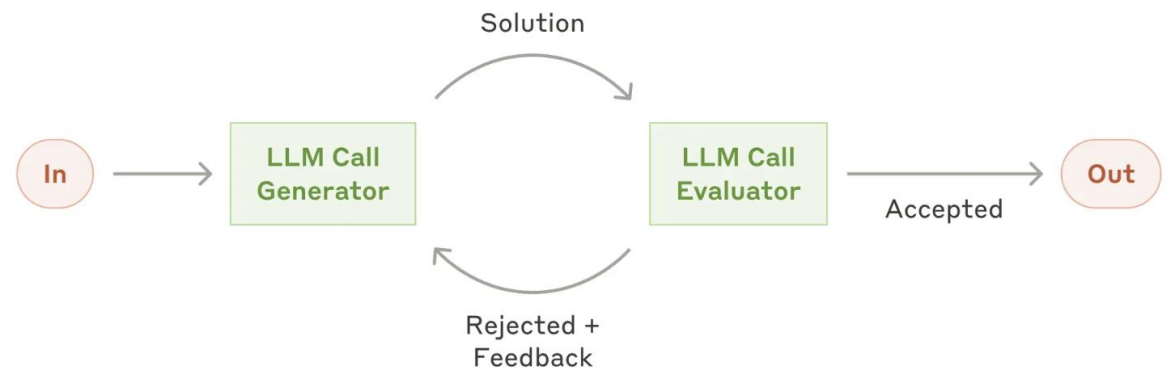
Building block: The augmented LLM

The basic building block of agentic systems is an LLM enhanced with augmentations such as retrieval, tools, and memory. Our current models can actively use these capabilities—generating their own search queries, selecting appropriate tools, and determining what information to retain.
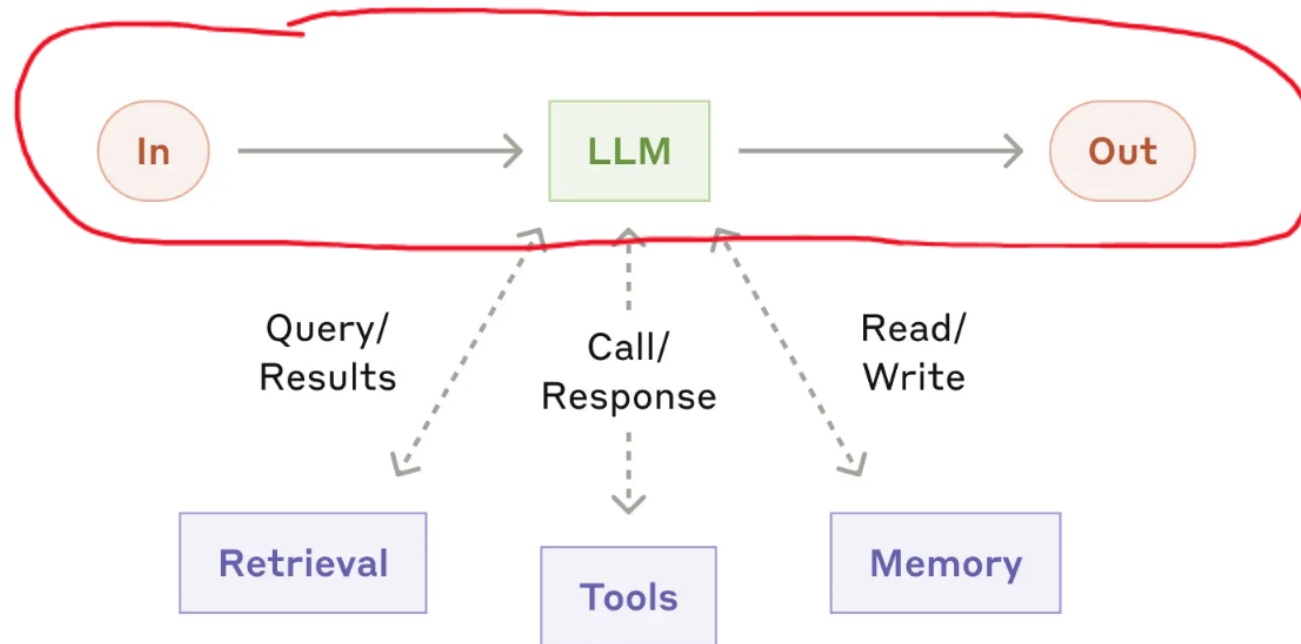
[SGT task – work in pairs / triads]

Note: you might find it helpful to **include one-shot / multi-shot examples** in your prompts.

The Bloom's verbs document will likely be a helpful starting point!

Critical thinking material

Below is the list of resources to look through prior to the LGTs:

🔗 MIT Critical thinking paper

🔗 Bloom's taxonomy (short video)

📄 Bloom's taxonomy -- verbs

Tutor me about <a selected topic> using questions based in <a selected Bloom's layer>.

Role: Worried Oxford tutor, trying to see if machines might be able to replace them

Output: Two files to submit

1. **Your single python file including all the prompts + code**

2. **One-page explanation of:**
   - The LLM-infrastructure framework + your prompts
   - Description of the 3 most difficult design challenges you faced today (whether you solved them or not!)

[SGT task – work in pairs / triads]

Suggested time progression:

- spent **30-40 minutes** on prototyping,

- then **10 minutes** on swapping your prototype with another team to get feedback

- **rest of time** to finalise & write up your SGT submission

Mentimeter