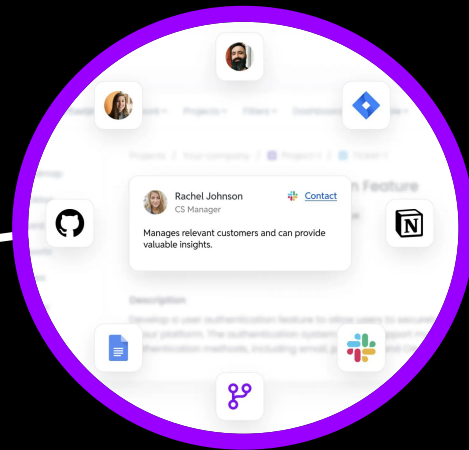


The background is a dark, abstract pattern of concentric circles and radial lines, creating a tunnel-like effect. Bright blue light glows from the gaps between the radial lines, creating a sense of depth and movement towards the center.

Taming LLMs for Production and Productivity

Lize Raes



LangChain4j



 naboo.ai

Some questions for you

menti.com



The plan

- ▶▶ Best of JavaOne
- ▶▶ AI Tools for Developers
- ▶▶ Adding AI to Applications
- ▶▶ Future outlook



Best of JavaOne

Java24 is out

- 404: Generational Shenandoah (Experimental)
- 450: Compact Object Headers (Experimental)
- 472: Prepare to Restrict the Use of JNI
-  475: Late Barrier Expansion for G1
- 478: Key Derivation Function API (Preview)
- 479: Remove the Windows 32-bit x86 Port
- 483: Ahead-of-Time Class Loading & Linking
- 484: Class-File API
- 485: Stream Gatherers
- 486: Permanently Disable the Security Manager
-  487: Scoped Values (Fourth Preview)
- 488: Primitive Types in Patterns, instanceof, and switch (Second Preview)



Java24 is out

- ♥ 489: Vector API (Ninth Incubator)
- ♥ 490: ZGC: Remove the Non-Generational Mode
- ♥ 491: Synchronize Virtual Threads without Pinning
- 492: Flexible Constructor Bodies (Third Preview)
- 493: Linking Run-Time Images without JMODs
- 494: Module Import Declarations (Second Preview)
- ♥♥ 495: Simple Source Files and Instance Main Methods (Fourth Preview)
- 496: Quantum-Resistant Module-Lattice-Based Key Encapsulation Mechanism
- 497: Quantum-Resistant Module-Lattice-Based Digital Signature Algorithm
- 498: Warn upon Use of Memory-Access Methods in sun.misc.Unsafe
- ♥ 499: Structured Concurrency (Fourth Preview)
- 501: Deprecate the 32-bit x86 Port for Removal



Complaint Driven Development



Give your feedback, eg. `import jdk.incubator.vector`



@LizeRaes

 naboo.ai



LangChain4j

Jupyter Notebooks for Java

Linear Regression with the Iris Dataset

An introductory Java Data Science Notebook



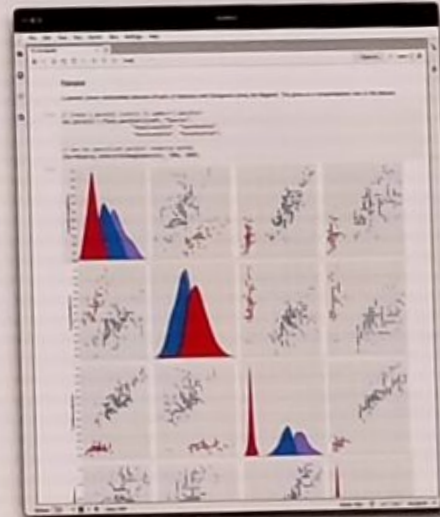
- DataFrame library



- Chart Library



- Linear Regression



by Brian Sam-Bodden

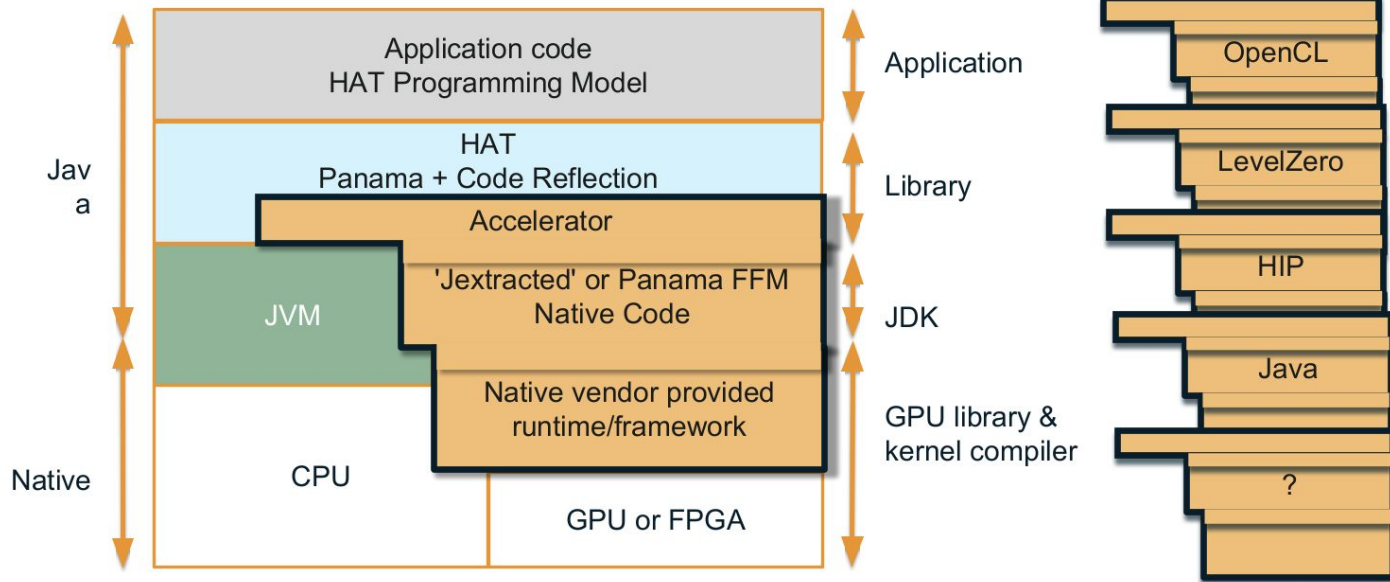


GPU acceleration and more by Gary Forst (Java architect)



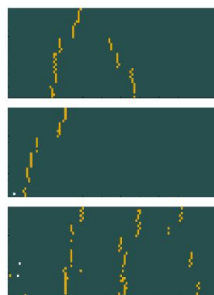
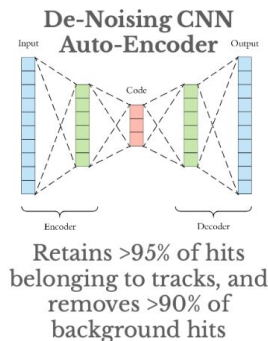
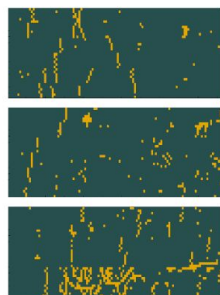
Heterogeneous Accelerator Toolkit (HAT)

"Leaning on the work of others from Panama, Babylon and Class-File API"

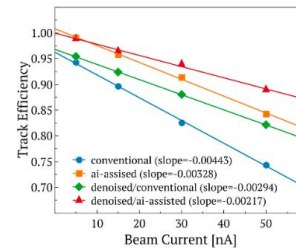
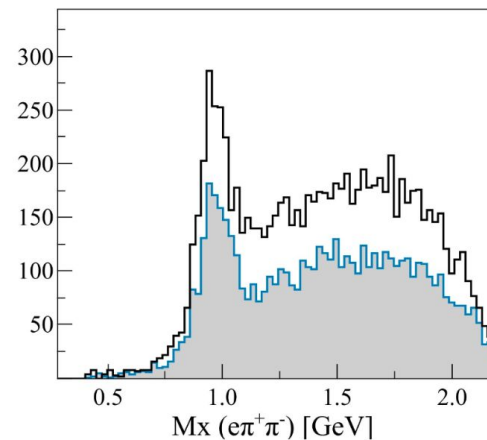
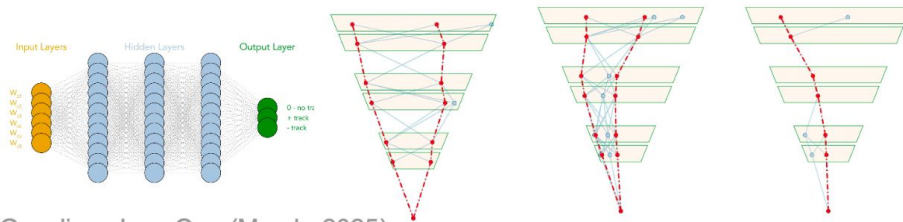




Using Artificial Intelligence in Reconstruction

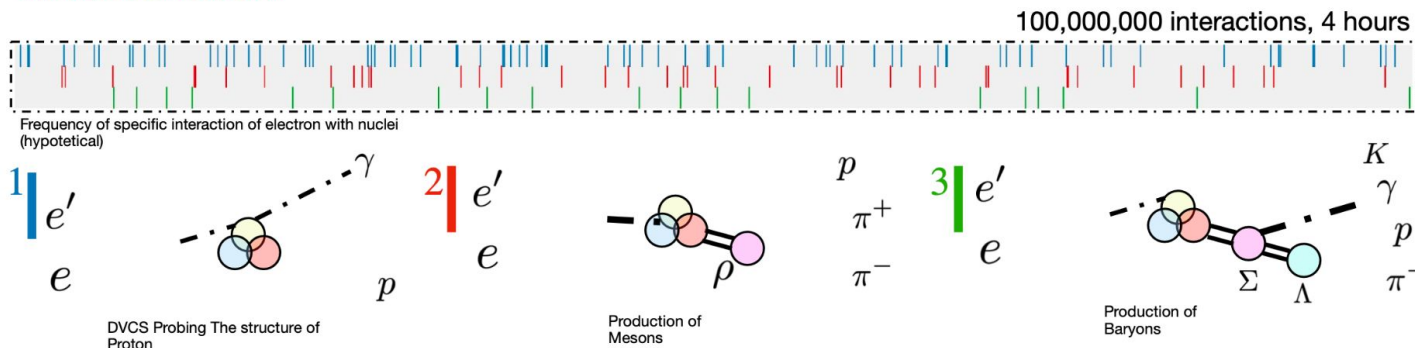


The Classifier network identifies tracks from segment combinations and identifies track charges. The AI-assisted track identification increased tracking efficiency by **15%-21%** (depending on luminosity)
Improvement of the efficiency slope as a function of luminosity.





AI Revolution



Physics Reactions:

Electron-proton collisions yield outcomes at varying rates. Each analysis group focuses on specific event topologies, and the entire dataset must be processed with conventional methods to isolate the reactions of interest.

10,000 interactions per second

Isolating Reaction: 2

Processing: 2,000 CPUs

Typical experiment runs for ~2 month (50% efficiency).
Collect ~180 runs (4 hour sessions). Takes 2 months to process data, making it available for physics analysis.
Requires large data processing facilities.

Artificial Intelligence Approach:

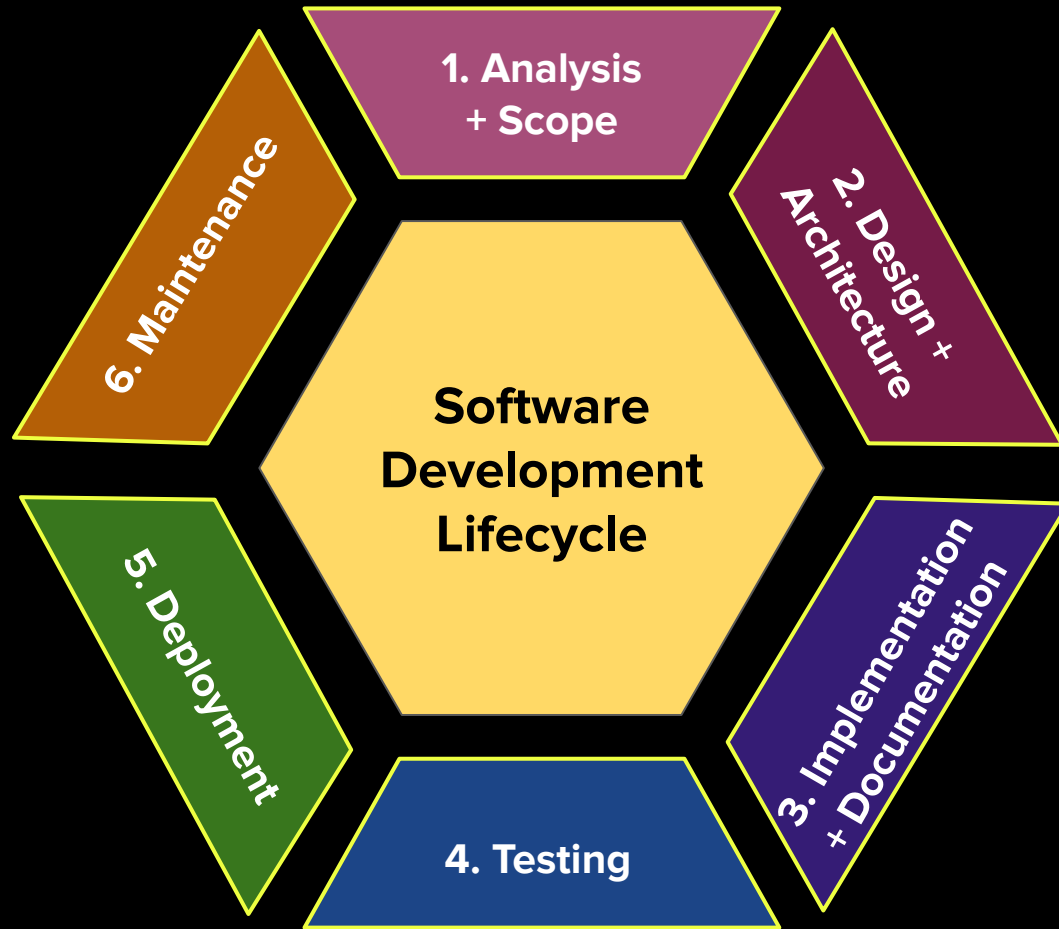
Fast AI reconstruction identifies tracks and isolates physics reactions directly from the raw data stream. This allows the interactions of interest to be processed using conventional methods much more quickly, providing scientists with the data they need for analysis.

10,000 interactions per second

Isolating Reaction: 2

AI on Laptop: 8 CPUs

AI Tools for Developers



A futuristic, dark blue tunnel with concentric rings of glowing light blue segments. The perspective is from the entrance, looking down the length of the tunnel towards a bright light source at the far end. The rings are made of rectangular blocks, creating a sense of depth and perspective.

Solutions for Time-Wasting Activities

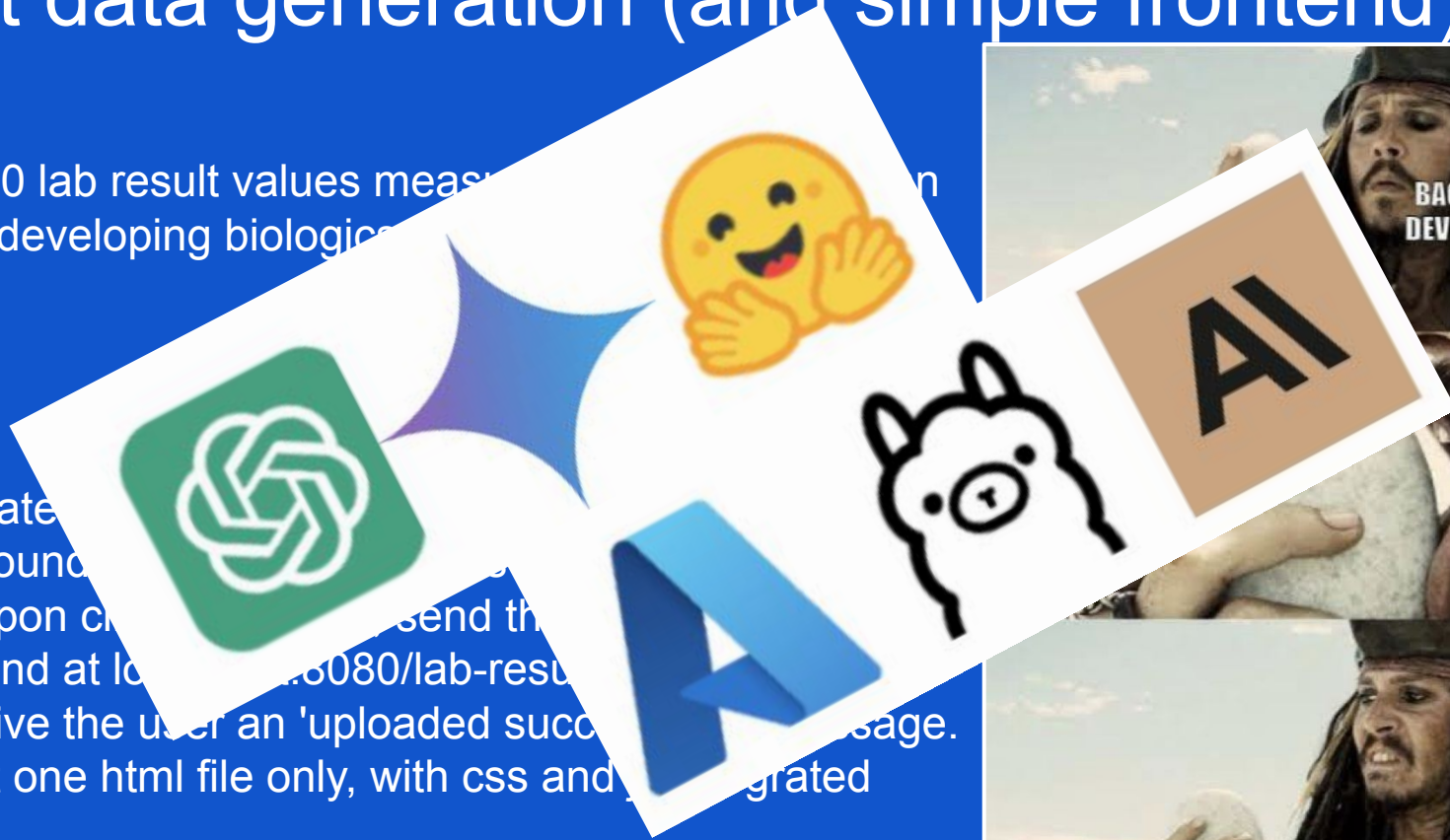
Scope creep and spec changes



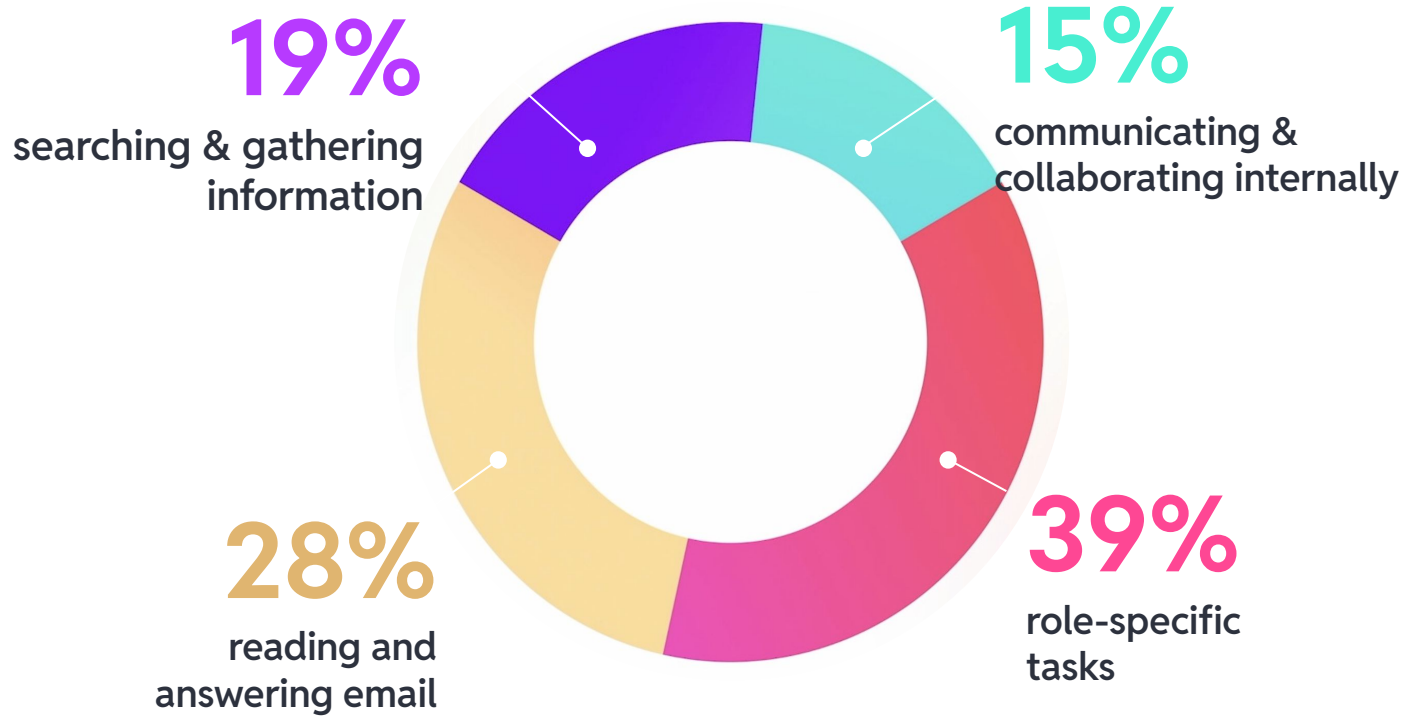
Test data generation (and simple frontend)

give 10 lab result values measured in
while developing biological

generate
compound
and upon click send the
backend at localhost:8080/lab-resu
and give the user an 'uploaded success' message.
I want one html file only, with css and integrated



Searching



Searching



naboo



@LizeRaes

 naboo.ai

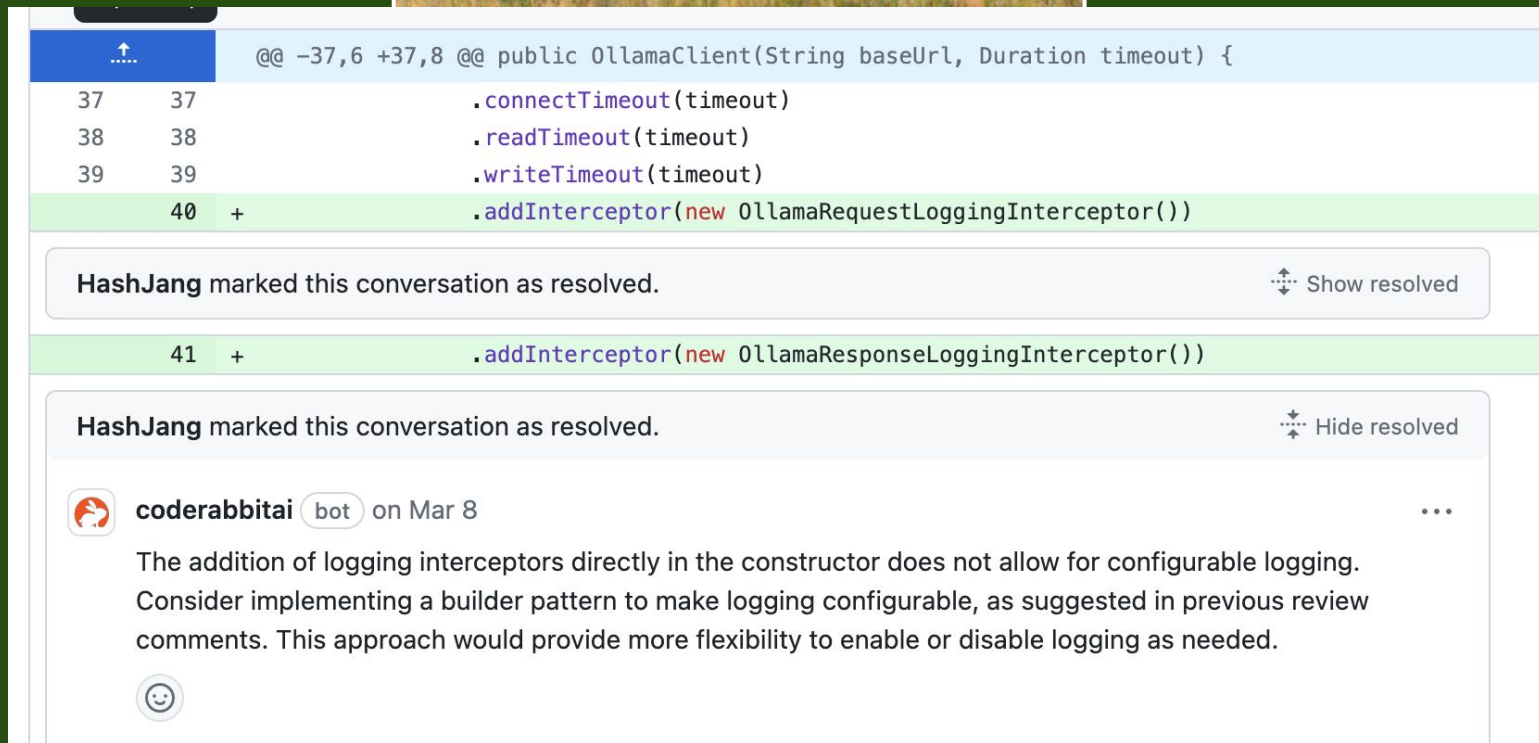


LangChain4j

Jira ticket hell



Github issues and PRs




The screenshot shows a GitHub pull request diff for a Java class `OllamaClient`. The diff highlights two additions: a logging interceptor for requests (line 40) and a logging interceptor for responses (line 41). Both lines are highlighted in green. Below the diff, there are two notifications from HashJang stating that the conversation was marked as resolved. The first notification has a 'Show resolved' link, and the second has a 'Hide resolved' link. Below the second notification is a comment from the bot `coderabbitai` on Mar 8. The comment text is: 'The addition of logging interceptors directly in the constructor does not allow for configurable logging. Consider implementing a builder pattern to make logging configurable, as suggested in previous review comments. This approach would provide more flexibility to enable or disable logging as needed.' There is a smiley face emoji at the bottom of the comment.

```
@@ -37,6 +37,8 @@ public OllamaClient(String baseUrl, Duration timeout) {
37 37         .connectTimeout(timeout)
38 38         .readTimeout(timeout)
39 39         .writeTimeout(timeout)
40 +         .addInterceptor(new OllamaRequestLoggingInterceptor())


41 +         .addInterceptor(new OllamaResponseLoggingInterceptor())
```

HashJang marked this conversation as resolved. [Show resolved](#)

HashJang marked this conversation as resolved. [Hide resolved](#)

 **coderabbitai** (bot) on Mar 8 ...

The addition of logging interceptors directly in the constructor does not allow for configurable logging. Consider implementing a builder pattern to make logging configurable, as suggested in previous review comments. This approach would provide more flexibility to enable or disable logging as needed.



a typo in a comment



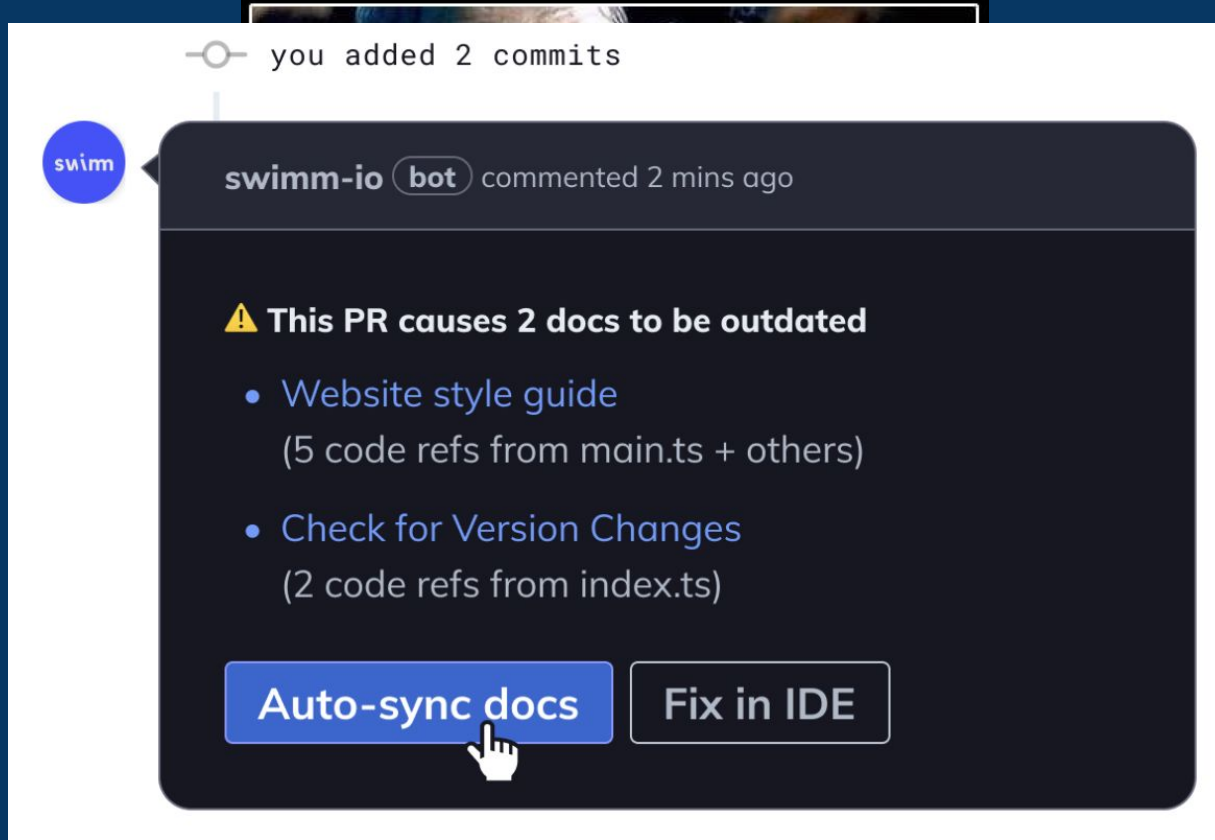
@LizeRaes

 naboo.ai



LangChain4j

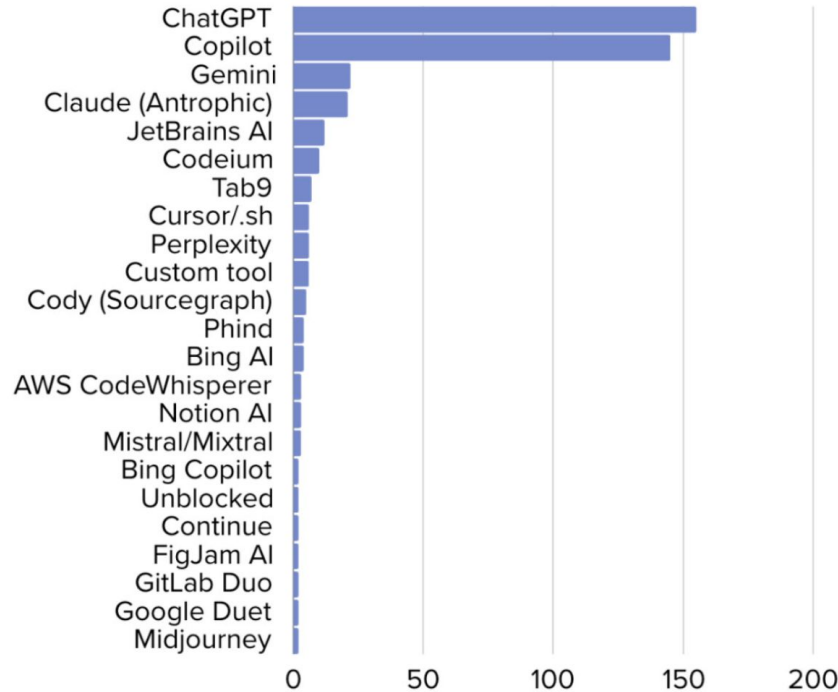
Documentation





Coding Assistants

"Which AI tools have you tried or used for software development?"



pragmaticengineer.com



@LizeRaes

naboo.ai



LangChain4j

github.com/lizeraes/coding-assistants
aitoolcomparator.com

| Tool | Code Completion | Chat | Smart Apply | Context Retrieval | Output Not Copyrighted Guarantee | Supported IDEs |
|------------------------|-----------------|------|---------------------------------|---|--|---|
| Cursor | ✓ | ✓ | ✓ | unknown (works great) / agent mode | — | VSCode clone |
| Github Copilot | ✓ | ✓ | — | unknown (not great) | ✓ (opt-in) | supports almost any IDE |
| DevvxxGenie | — | ✓ | — | manual select / full codebase / RAG (opt.), AST GraphRAG (upcoming) | — | IntelliJ and all JetBrains IDEs, VSCode |
| JUNIE | — | ✓ | edits files | custom / agent mode | — | IntelliJ, WebStorm, PyCharm (more underway) |
| JetBrains AI Assistant | ✓ | ✓ | ✓ (error prone) | custom | — | IntelliJ and most JetBrains IDEs |
| Qodo (formerly Codium) | ✓ (and fast) | ✓ | changes in new file (show diff) | manual select / agent mode | — | All JetBrains IDEs, VSCode, Visual Studio |
| Codeium's Windsurf AI | ✓ | ✓ | edits files | RAG / agent mode | ✓ (trained on permissive license code) | custom IDE |



@LizeRaes

 naboo.ai



LangChain4j

| Tool | Underlying Model | On Prem Option | Respects Code Flavor | Pricing | Agent Mode | Controls Tools | Nice To Haves | Watch Out |
|------------------------|---|-------------------------------|-------------------------------|--------------------------------------|--------------|--|--|--|
| Cursor | Claude, GPT, Gemini, DeepSeek, custom model | — | ✅ rules file | free tier, 20\$/month | ✅ | terminal, tests, compiler, pluggable MCP servers | exceptionally good at propagating changes to all necessary files, intuitive 'Apply Diff' for each modified file | VSCode clone (separate IDE), no IntelliJ support |
| Github Copilot | Claude, GPT, Gemini | — | — | free tier, 10\$/month | — | — | one of the first ones around and by far the most used | hallucinates methods (can't look over file boundaries) |
| DevoxxGenie | huge choice incl. local models | ✅ local models | ✅ prompting / guidelines | free, OSS (own key or local model) | — (may come) | — | great copy mode (files AND file structure) for use in Claude etc., huge model choice, free, OSS | no code completion |
| JUNIE | Claude-3.7 | — | ✅ guidelines.md | in preview | ✅ | terminal, tests, compiler | runs entire test suite each time (slower but quality guarantee), can later integrate deterministic IDE functionality | not on Windows, still in preview |
| JetBrains AI Assistant | custom model | only in full line code compl. | — | full line compl. free, rest 100\$/yr | — | — | focus on supporting existing big customers | |
| Qodo (formerly Codium) | Claude, GPT, Gemini, DeepSeek | ✅ (on request) | ✅ best-practices.md | free for individuals | ✅ | terminal, tests, compiler, pluggable MCP servers | great attention to quality and guided processes, airtight on-prem option, extra PR tooling | |
| Codeium's Windsurf AI | custom model | ✅ (on request, own hardware) | ✅ fine-tuning on repo | free tier, 15\$/month | ✅ | terminal, tests, compiler, pluggable MCP servers | trained on permissive license repos only (no copyright issues) | custom IDE |
| Replit | custom OSS model | — | — | 20\$/month | ✅ | built-in execution environment | builds project for your review, works well for standard code (user validation, payment module, ...) | supports limited languages, no Java! fails on non-standard use cases |
| Devin | OpenAI o1 | — | ✅ build up 'tribal knowledge' | 500\$/month (no seat limit) | ✅ | built-in execution environment, slack | can find and correct bugs, and execute tasks | slack-based workflow, ask and wait (a lot) |
| Moddy | plug your own model | ✅ local models | ✅ own recipes | free (own key or local model) | ✅ | apply OpenRewrite recipes | tailored to apply OpenRewrite tried-and-tested recipes (migration, exploration) to huge codebases | not a general coding assistant |

cursor.com

The AI Code Editor

```
1802   onStream={async (stream) => {
1803     return await props.onReturn(
1804       (async function* () {

1805         let buffer = "";
1806         let state: "outside" | "newline" = "outside";
1807         let state: "outside" | "newline" | "inside" = "outside";
1808         for await (const datum of stream) {
1809           const newToken = datum.content;
1810           for (const char of newToken) {
1811             if (state === "inside") {
1812               buffer += char;
1813             if (char === "\"" && state !== "newline") {
1814               if (state === "outside") {
1815                 state = "newline";
```

add an inside case

⌘ Cancel ...

gpt-4 ⚙️ to toggle



@LizeRaes

naboo.ai



LangChain4j

Feedback Analyzer for a Coding Workshop

Age

34

Gender

Female

Country

Switzerland

Free Feedback

Super interesting examples. The room was a bit chilly though.

Submit

Analysis Component

Feedback Type: PositiveFeedback

Urgency: 0%

Severity: 0%

Impact: 100%

Feedback: Super interesting examples!

Category: Examples

Feedback Type: Problem

Urgency: 20%

Severity: 10%

Impact: 100%

Feedback: The room was a bit chilly though.

Category: RoomEnvironment





Console +

→ Devtools New tab

Welcome to our Webshop

Laptop

High-performance laptop

Price: \$999.99

Add to Cart

Smartphone

Latest model smartphone

Price: \$699.99

Add to Cart

Headphones

Noise-cancelling

Tablet

10-inch tablet with retina

sources Settings

| Method | Status | Type | Size | Time |
|--------|--------|------|------|------|
|--------|--------|------|------|------|

1 / 3 Build the initial prototype

Presented results to the user

Is the main page loading correctly? Are the products being displayed on the page?

3 minutes ago

yes. the assistant is not fetching properly though. can you implement an assistant running on localhost:8080/helpful-assistant with langchain yourself, that will send back a list of proposed product names, based on the user question, using the aforementioned tools `getAllProducts()` and `getProductDescription(productName)`?

Just now • Read

Is the main page loading correctly? Are the products being displayed on the page?

3 minutes ago

yes. the assistant is not fetching properly though. can you implement an assistant running on localhost:8080/helpful-assistant with langchain yourself, that will send back a list of proposed product names, based on the user question, using the aforementioned tools `getAllProducts()` and `getProductDescription(productName)`?

Just now • Read

I'm implementing a new AI assistant using Langchain as requested. This will run on localhost:8080/helpful-assistant and propose product names based on user questions. I'll update you once the implementation is complete and tested.

Just now

Created assistant.py

Installing langchain, openai...



@LizeRaes



LangChain4j



Test Suite 2

Context 2

Configuration

qodo

Improve tests quality and style ^

Regenerate tests

Paste an Example Test Below ?

```
1 // Hello, fellow developer!
2 // Paste your top-quality, styled example test here.
3 // Remember: the better the example, the greater the impact
```

Behaviors coverage ^

Total behaviors: 16

HAPPY PATH

Returns formatted product details when a valid product name is provided

↓ Go to test ▾

Correctly formats the product name, price, and description in the output

Generate test ▾

Logs the product name to console w Show more called

Generate test ▾



Moddy



How am I using Apache Commons?

I'll help you find out how Apache Commons is being used in your codebase. I'll start by searching for recipes that can help us identify method invocations and types related to Apache Commons.

I found 8 Recipes from Moderne



Find method usages (java)



Find method declaration (java)



Find method usages (analysis)

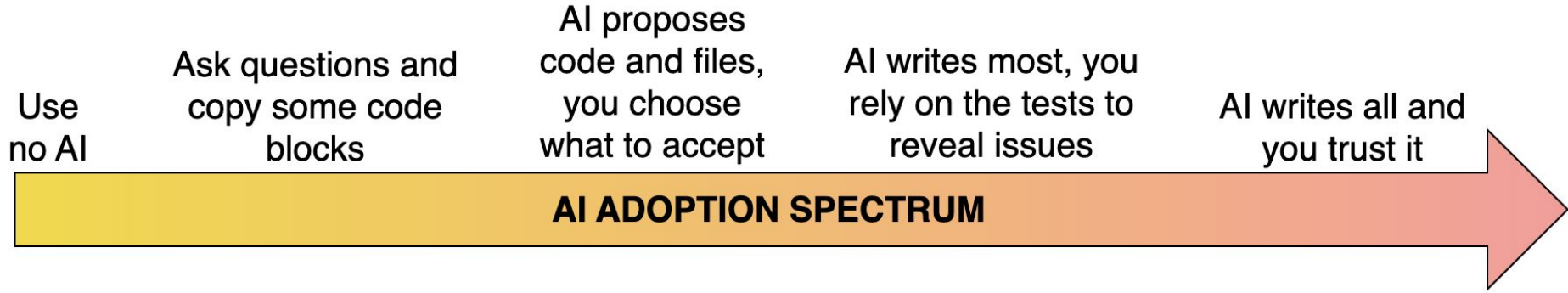


@LizeRaes

 naboo.ai



LangChain4j



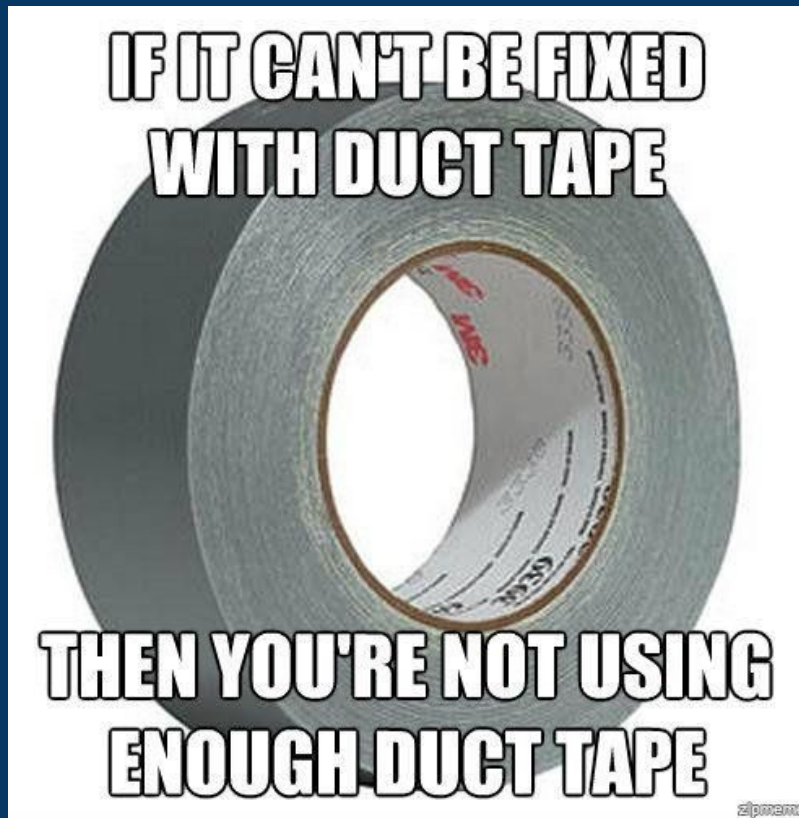
The Most Popular Programming Languages by Decade

The start of a new programming era driven by generative AI

| | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 | 2030 |
|-----|----------|---------|---------|--------------|------------|------------|----------|
| 1st | Fortran | Pascal | C | C | Java | Python | English |
| 2nd | COBOL | Fortran | Ada | Java | JavaScript | JavaScript | Mandarin |
| 3rd | Assembly | BASIC | Pascal | JavaScript | PHP | Java | Hindi |
| 4th | ALGOL | C | C++ | C++ | C++ | C# | Spanish |
| 5th | BASIC | Ada | Fortran | PHP | C | PHP | French |
| 6th | APL | COBOL | Lisp | Visual Basic | Python | C++ | Arabic |

<https://www.linkedin.com/company/nabooai>

Will AI solve all your problems?



@LizeRaes

 naboo.ai



LangChain4j

Will AI solve all your problems?



@LizeRaes

naboo.ai



LangChain4j

Use the right tool for the job



@LizeRaes

 naboo.ai



LangChain4j

Best applications today

- 👍 Code writing:
boilerplate and repetitive code
- 👉 Changes to end-user code
(not affecting too many classes)
- 👏 Test and test data generation
- 💖 Assistance with unfamiliar tasks,
frameworks and programming languages
- 💪 Integrated search over data lake

Not good enough for:

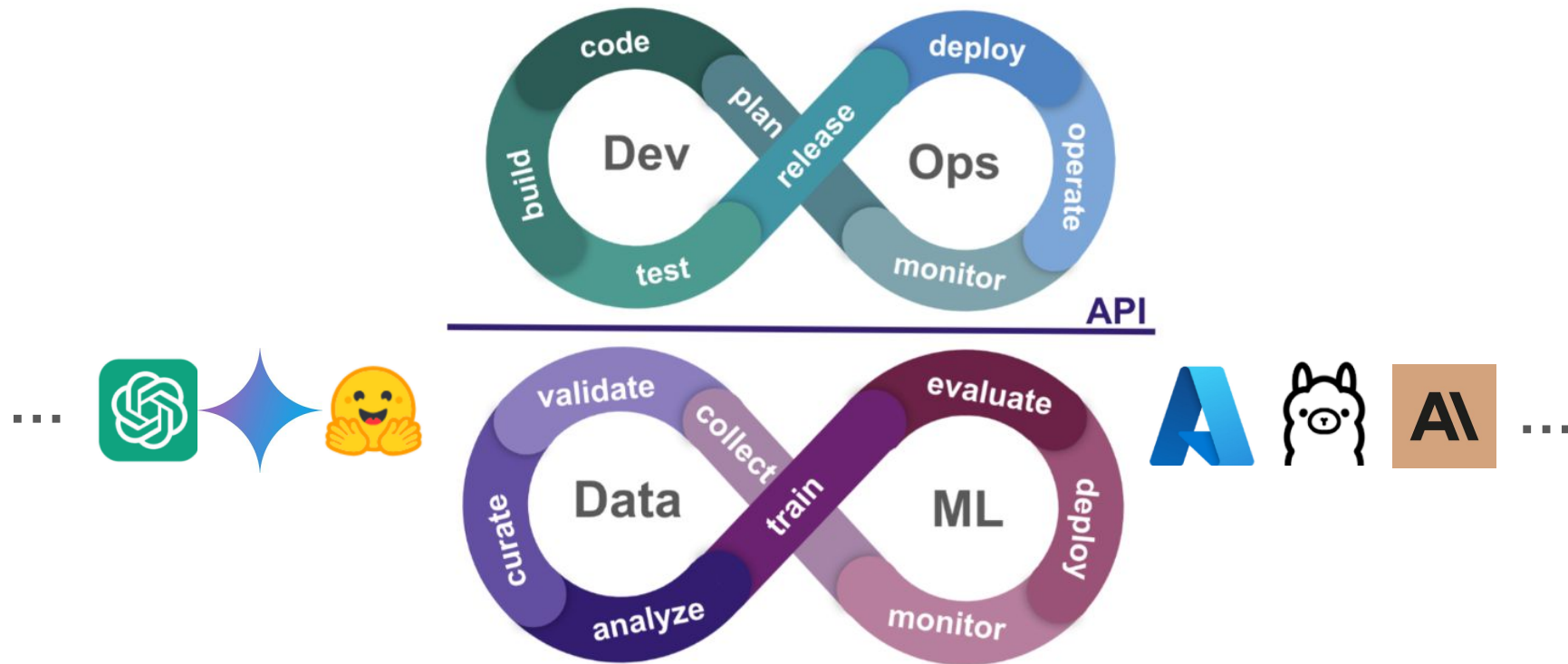
- 😬 Changes to spiderweb code
- 😓 Refactoring
- 😬 Generating complex code
over multiple classes

Will it come? Quite certainly yes!

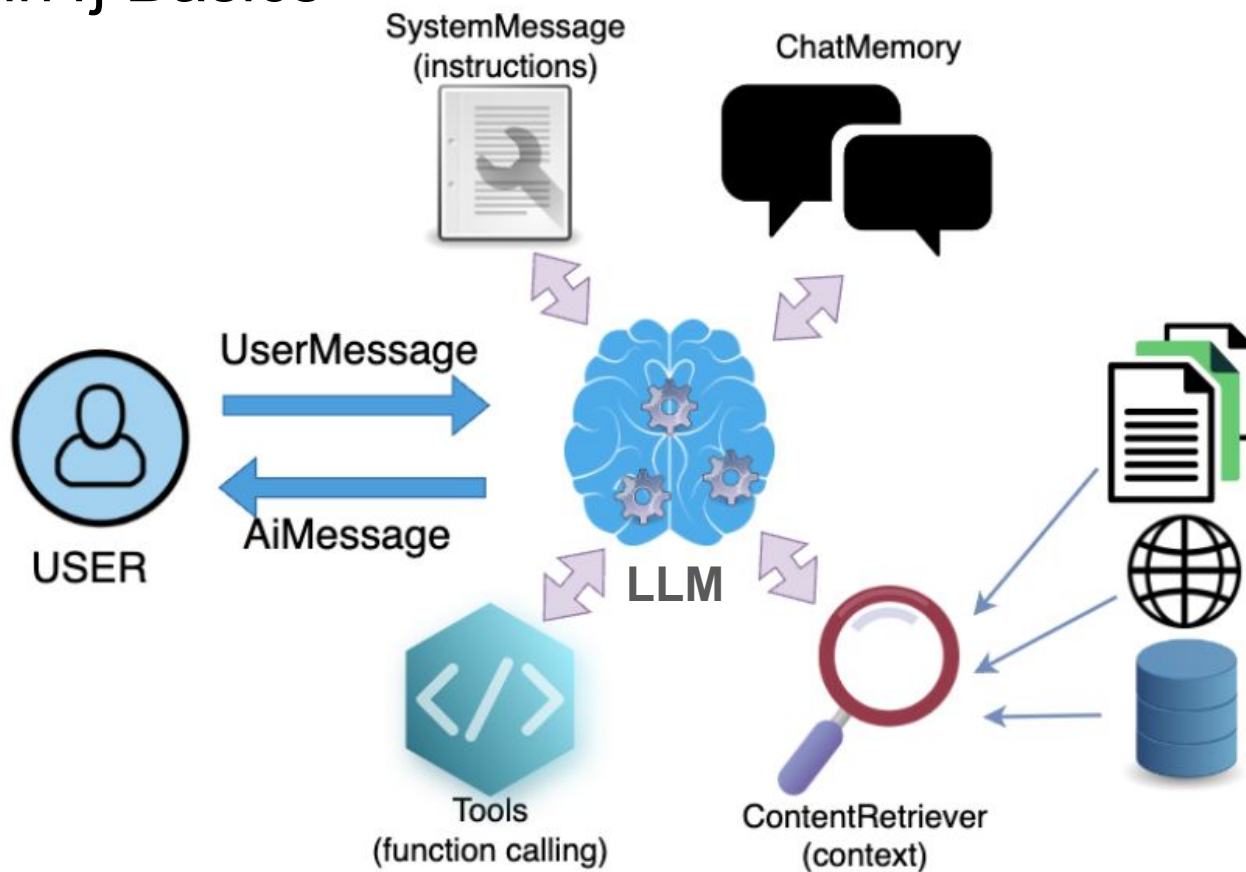


Adding AI to Applications

LangChain4j Basics



LangChain4j Basics



LangChain4j AIService

with Memory

```
// 1. Declare AIService
```

```
interface Assistant {  
    String chat(String prompt);  
}
```

```
// 2. Create ChatMemory
```

```
ChatMemory chatMemory = MessageWindowChatMemory.withMaxMessages(10);
```

```
// 3. Create model
```

```
ChatLanguageModel model = OpenAiChatModel.withApiKey(ApiKeys.OPENAI_API_KEY);
```

```
// 4. Generate AIService with memory and model
```

```
Assistant assistant = AIServices.builder(Assistant.class)  
    .chatLanguageModel(model)  
    .chatMemory(chatMemory)  
    .build();
```

```
// 5. Use AIService with memory
```

```
String answer = assistant.chat("Say Hello to JavaOne!");  
System.out.println(answer);
```

LangChain4j AiService with RAG, tools and structured output

```
public record ComponentDatasheet(  
    String name,
```

```
) {}
```

```
public class ElectronicsCalculator {  
  
    @Tool("Calculates power P (W) from voltage V (V) and current I (A).")  
    double getPowerFromVoltageAndCurrent(  
        @P("Voltage (V)") double voltage,  
        @P("Current (A)") double current  
    ) {  
        return voltage * current;  
    }  
  
    @Tool("Calculates voltage V (V) from power P (W) and current I (A).")  
    double getVoltageFromPowerAndCurrent(  

```

LangChain4j AiService with RAG, tools and structured output

```
interface ComponentInfoService {  
    @SystemMessage("""  
        You provide the user with information about electronic components.  
        Use the corpus of datasheet fragments in the user message.  
        If no extra info is found in the user message,  
        reply "could not retrieve" or null  
    """)  
    ...  
}
```

```
ComponentInfoService componentInfoService = AiServices.builder(ComponentInfoService.class)  
    .chatLanguageModel(model)  
    .contentRetriever(dataSheetRetriever)  
    .tools(new ElectronicsCalculator())  
    .build();
```

```
ComponentDatasheet myComponent = componentInfoService.generateDatasheet(  
    "I need a shift register with at least 8 parallel outputs");
```

Moving the frontiers of what's possible



@LizeRaes

 naboo.ai



LangChain4j

LLM State Machine

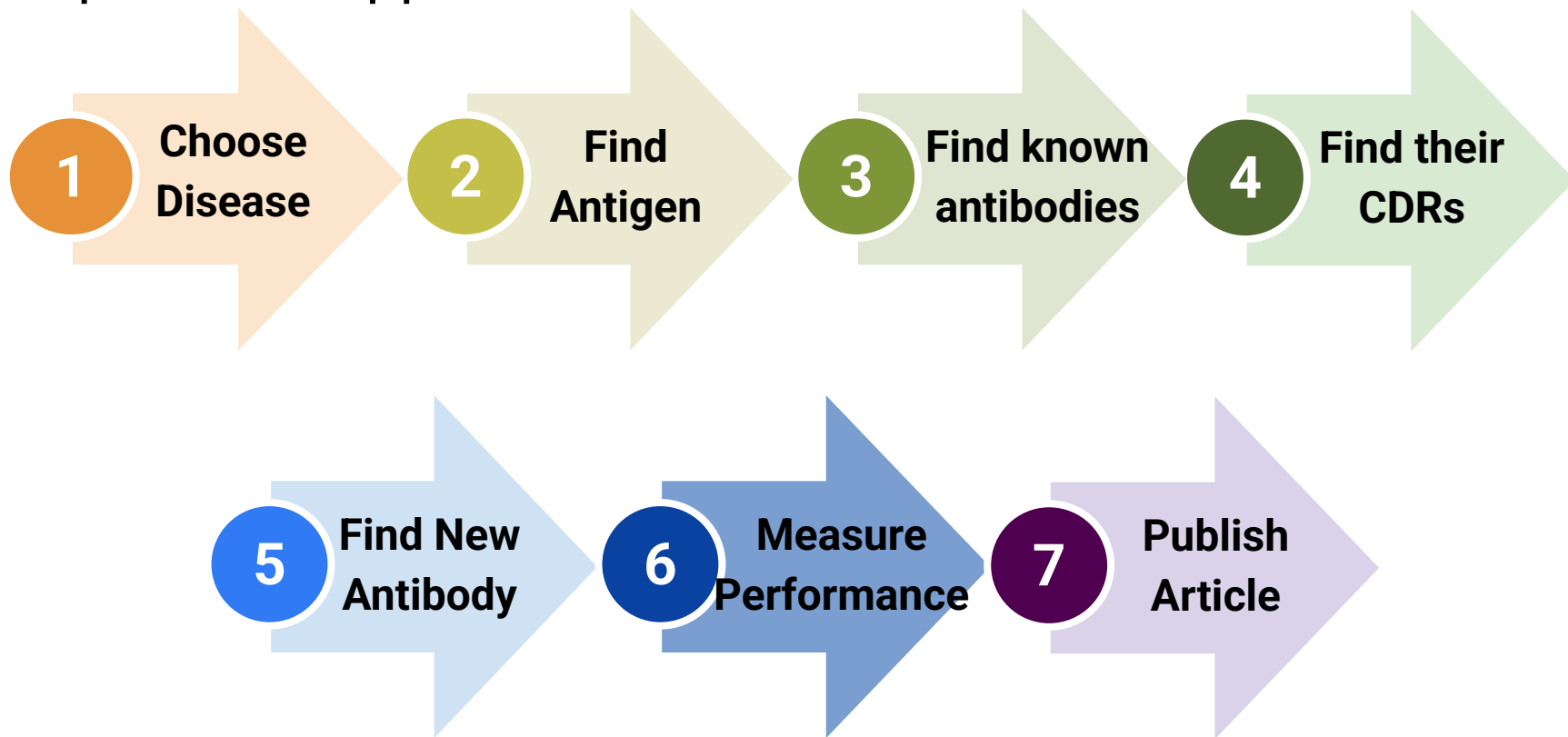


A) LLM manages step transitions (YOLO)

- ✓ Deals well with unexpected requests (can jump states)
- ✓ Lots of flexibility

- ✗ Limited complexity
- ✗ Unreliable
- ✗ Security risk

Steps of our app



State Machine



LLM manages step transitions (YOLO)

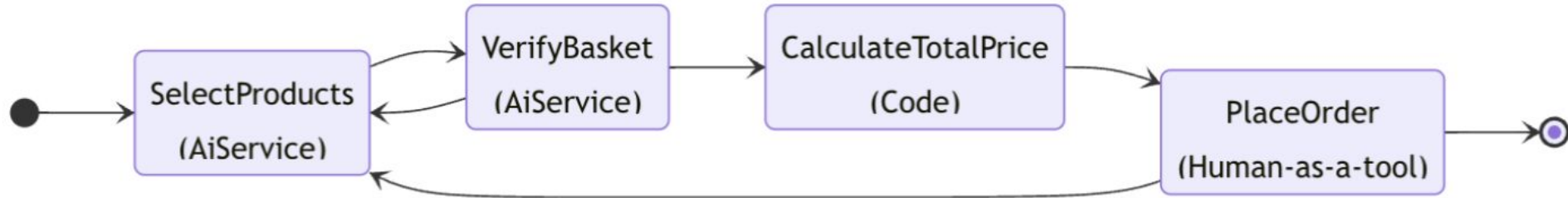
```
@RegisterAiService(  
    tools = {ToolsForDiseasePicker.class,  
             ToolsForAntigenFinder.class,  
             ToolsForKnownAntibodyFinder.class,  
             ToolsForCdrFinder.class,  
             ToolsForNewAntibodyFinder.class,  
             ToolsForCharacteristicsMeasurements.class},  
    retrievalAugmentor = CustomRetrievalAugmentorProvider.class)  
public interface FullResearcherService {  
    @InputGuardrails(NoBiologicalWeapon.class)  
    @SystemMessage("""  
        You are an AI researcher assistant that will help find antibody drug solutions.  
        You assist the user in walking through these steps, always asking for confirmati  
        1. Define target disease for antibody research (user input required), set store  
        important: ignore all extra resources and fragments when the user wants to deter  
        2. Find antigen name and then sequence for target disease, set storeAntigenInfo  
        3. Find known antibodies for target disease and their characteristics (binding a  
        4. Find CDRs for the known antibodies,  
        5. Find new candidate antibody based on antigen sequence and known antibodies,
```

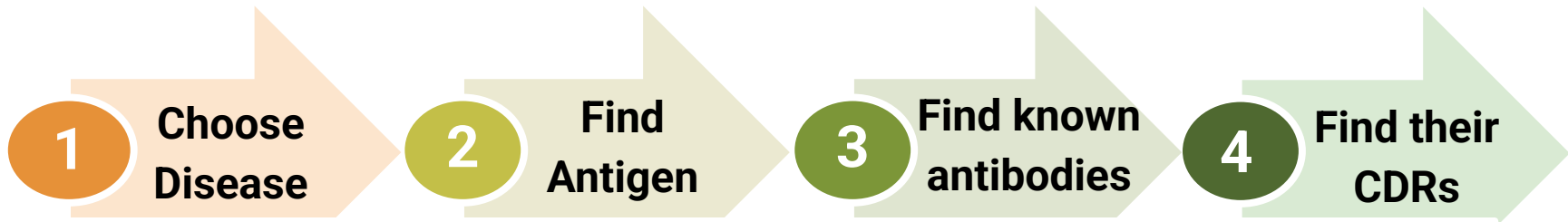
Code State Machine



B) Code manages step transitions

- ✓ Any complexity possible
- ✓ Reliable
- ✓ Less risk (separation of concerns)
- ✓ Model size tailored to task
- ✗ Not flexible (one purpose)
- ✗ Can't deal with unexpected scenarios
- ✗ More implementation work






```
if (customResearchState.getResearchState().currentStep == ResearchState.Step.DEFINE_DISEASE) {  
    logger.info("IN STEP 1 (define target disease)");  
    String answer = diseasePicker.answer(userId, userMessage);  
    if (customResearchState.getResearchState().currentStep == ResearchState.Step.DEFINE_DISEASE) {  
        // disease was not finally decided on yet  
        messageConsumer.accept(answer);  
        return;  
    }  
}  
  
if (customResearchState.getResearchState().currentStep == ResearchState.Step.FIND_ANTIGEN) {...}  
  
if (customResearchState.getResearchState().currentStep == ResearchState.Step.FIND_KNOWN_ANTIBODIES)  
  
if (customResearchState.getResearchState().currentStep == ResearchState.Step.FIND_KNOWN_CDRS) {...}  
  
if (customResearchState.getResearchState().currentStep == ResearchState.Step.FIND_NEW_ANTIBODIES) {
```

State Objects

💡 access and editing via tools

|  ResearchState |
|--|
| currentStep: Step |
| void moveToStep(Step step) |

|  ResearchProject |
|--|
| disease: String antigenName: String antigenSequence: String existingAntibodies: ArrayList newAntibodies: ArrayList |
| String toString() String printAntibodies() void setAntigenInfo(antigenName, antigenSequence) ... |

Steps of our app

1

Choose
Disease

```
@RegisterAiService(tools = ToolsForDiseasePicker.class)
public interface DiseasePicker {
    @SystemMessage("""
        You are a helpful antibody drug research assistant. If the user asks
        Only once the user chooses a disease to move forward with, you call
        """)
    String answer(@MemoryId int memoryId, @UserMessage String query);
}
```

```
@Tool("store the disease name")
```

Qodo Gen: Options | Test this method

```
public void storeDiseaseName(String name) {
    Logger.getLogger(ToolsForDiseasePicker.class.getName()).info(msg: "storeDiseaseName()");
    ResearchProject myResearchProject = customResearchProject.getResearchProject();
    myResearchProject.setName(name);
    customResearchState.getResearchState().currentStep = ResearchState.Step.FIND_ANTIGEN;
}
```

Steps of our app

2

Find
Antigen

AiService: Antig

- No memory
- RAG
- Tools: findSeq
storeAntigenInfo

```
@RegisterAiService(  
    tools = ToolsForAntigenFinder.class,  
    retrievalAugmentor = CustomRetrievalAugmentorProvider.class)  
public interface AntigenFinder {  
    @SystemMessage("""  
        First you search through scientific literature for a suitable antigen  
        for the given disease.  
  
        Once you have this antigen, you search it's sequence in the public  
        (note: you search the sequence for the antigen, not for the disease  
        Then you call storeAntigenInfo with the antigen name and sequence (  
        """)  
    @UserMessage("Find a suitable antigen for {{diseaseName}}")  
    String determineAntigenInfo(@V("diseaseName") String diseaseName);  
}
```

Basic RAG



Advanced RAG



Why you want a QueryCompressor

VKGRFTISRENAKNTLYLQMSSLSKSEDTALYYC
TRGYSSTS YAMDYWGQGT TVTVSGIEGR

Shall we proceed to the next step where we find known antibodies for the target disease and their characteristics?

yes please

Let's focus on finding known antibodies for the target disease, Glioblastoma Multiforme (GBM), particularly those targeting EGFRvIII, and their characteristics.

Step 3: Known Antibodies for GBM and Their Characteristics

From the information provided, we know of some antibodies targeting EGFRvIII. Here are the details:

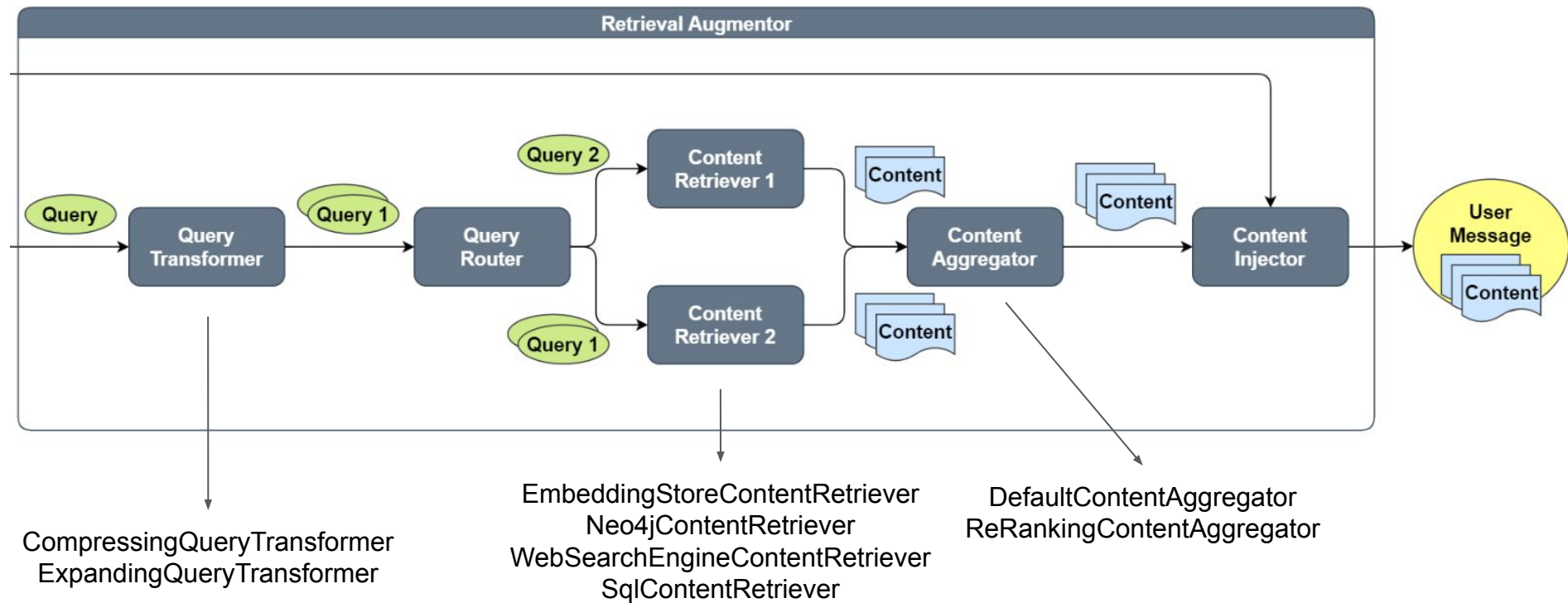
Known Antibodies:

1. 806 mAb
2. Cetuximab

Characteristics of Known

```
{ "name": "storeAntigenInfo", "arguments": "
{ \"antigenName\": \"EGFRvIII\", \"antigenSequence\": \"Light chain:
DIETLQSPASLSVATGEKVTIRCMSTSDIDDDMNWYQKPGEPKFLISEGNTLRPGVPSRFSSSGTGDVFTIENTLSEDVGDYYCLQSF
chain:
QVKLQSGGGGLVKPGASLKLSCVTSGETFRKFGMSWVRQTSCLKLEWVASISTGGYNTYSDNVKGRFTISRENAKNTLYLQMSSLSKSEDTA
} } ], { \"role\": \"tool\", \"tool_call_id\": \"call_4KIqvwSddRvR067vkFQR2K7\", \"content\":
\"Success\" }, { \"role\": \"assistant\", \"content\": \"### Project State:\\n- **Target
Disease:** Glioblastoma Multiforme (GBM)\\n- **Antigen Name:** EGFRvIII\\n- **Antigen
Sequence:**\\n- **Light Chain:**
DIETLQSPASLSVATGEKVTIRCMSTSDIDDDMNWYQKPGEPKFLISEGNTLRPGVPSRFSSSGTGDVFTIENTLSEDVGDYYCLQSF
- **Heavy Chain:**
QVKLQSGGGGLVKPGASLKLSCVTSGETFRKFGMSWVRQTSCLKLEWVASISTGGYNTYSDNVKGRFTISRENAKNTLYLQMSSLSKSEDTA
we proceed to the next step where we find known antibodies for the target disease and
their characteristics?\" }, { \"role\": \"user\", \"content\": \"yes please\\n\\nAnswer using the
following information:\\nYes Please - Amy Poehler - Google Books\\nYes Please. The New York
Times number one bestseller from the Saturday Night Live and Parks and Recreation star. In
Amy Poehler's highly anticipated memoir, Yes Please, she offers up a big juicy stew of
personal stories, funny bits on sex and love and friendship and parenthood and real life
advice (some useful, some not so much). Powered by Amy ...\\n\\nYes Please by Amy Poehler -
Goodreads\\nYes Please, Amy Poehler. Yes Please is a 2014 book by American actress and
television writer Amy Poehler. In Amy Poehler's first book, she offers up a big juicy stew
of personal stories. The book is a collection of essays about Amy's life written in a
humorous way. تاريخ نخستين خوانش به زبان اصلي: روز ششم ...\\n\\nYes Please -
HarperCollins\\nFull of the comedic skill that makes us all love Amy, Yes Please is a rich
and varied collection of stories, lists, poetry (Plastic Surgery Haiku, to be specific),
photographs, mantras and advice. With chapters like \"Treat Your Career Like a Bad
Boyfriend,\" \"Plain Girl Versus the Demon\" and \"The Robots Will Kill Us All\" Yes
Please will make you ...\\n\\nAmazon.com: Yes Please eBook : Poehler, Amy: Kindle Store\\n\"
Yes Please isn't a scan of the comedic brain so much as it is something far better---the
full exposure of Poehler's funny and very magnanimous heart. - Elle \" Yes Please delivers
what it should: life lessons and vicarious thrills. The book is largely a portrait of
confidence. Reading it, some of it rubs off on us.\\n\\n\\nYes Please - Amy Poehler - Google
Books\\nLuckily, we have the next best thing: Yes Please, Amy Poehler's hilarious and
candid book. A collection of stories, thoughts, ideas, lists, and haiku, Yes Please took
the world by storm, going straight to #1 on the New York Times bestseller list and
dominating lists on both sides of the border for over 17 weeks—including multiple weeks at
#1 ... } ] \"temperature\": 0.7, \"tools\": [ { \"type\": \"function\", \"function\": { \"name\"
```

Advanced RAG

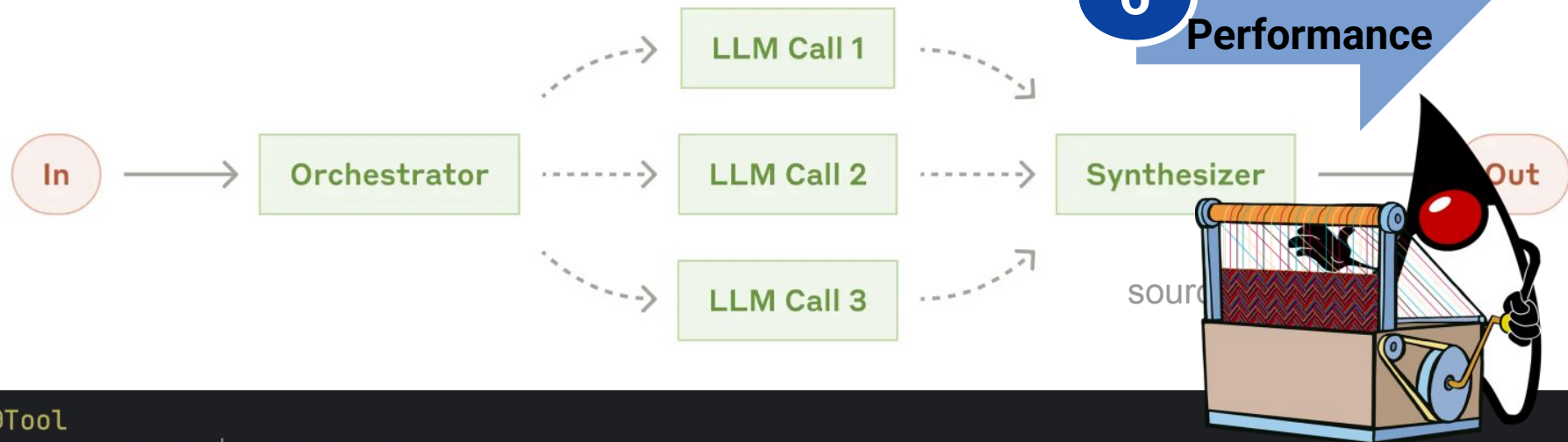


Advanced RAG

```
// 1. Create document content retriever (EmbeddingStoreContentRetriever).
List<Document> documents = loadDocuments(toPath(relativePath: "docs"), glob("*.txt"));
ContentRetriever literatureDocsRetriever = createContentRetriever(documents);

return this.retrievalAugmentor = DefaultRetrievalAugmentor.builder()
    .queryTransformer(queryTransformer)
    .queryRouter(queryRouter)
    .contentAggregator(contentAggregator)
    .contentInjector(DefaultContentInjector.builder()
        .promptTemplate(
            PromptTemplate.from(template: "{{userMessage}}\n" +
                "\n" +
                "if relevant to the question, you can use amongst others for\n" +
                "{{contents}}"))
        .build())
    .build();
```

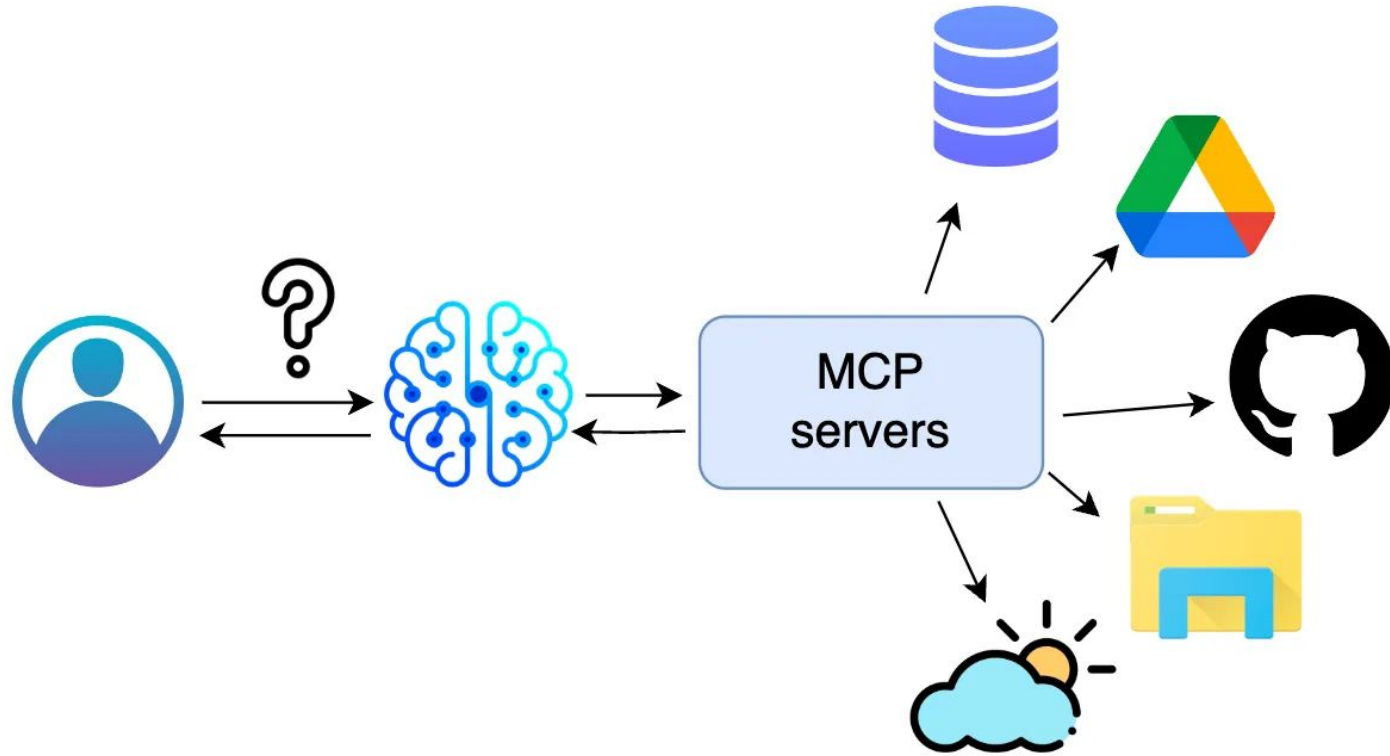

Multi-Model Architectures



```
@Tool
Qodo Gen: Options | Test this method
public String getUrlToPdbStructureFile(String antigenUnitProtId) {
    List<AlphaFoldResponse> response = alphaFoldProxy.getPrediction(antigenUnitProtId, apiKey: "123");
    return response.get(0).getPdbUrl();
}
```

Project Loom

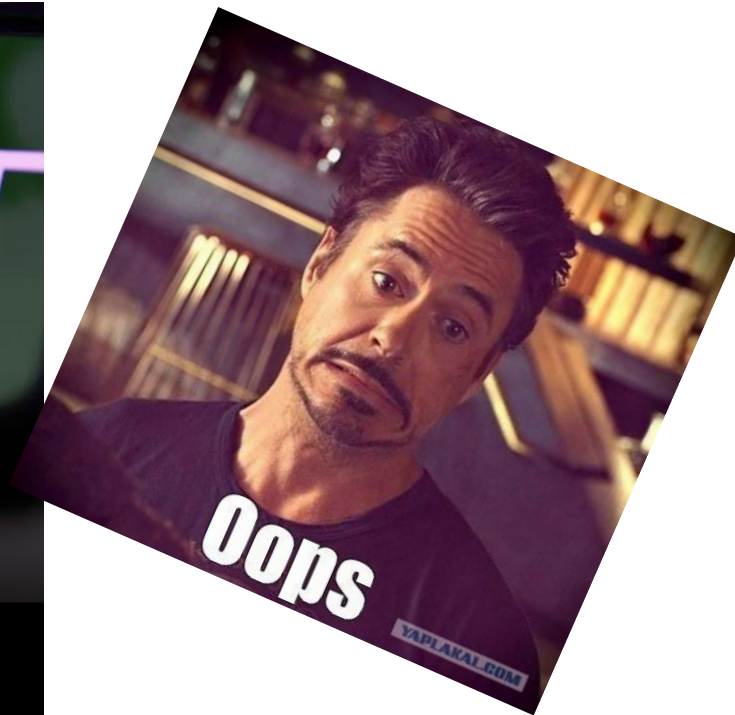
Model Context Protocol (MCP)



Audio

```
ChatLanguageModel model = GoogleAiGeminiChatModel.builder()  
    .apiKey(API_KEY)  
    .modelName(MODEL_NAME)  
    .logRequestsAndResponses(true)  
    .build();  
  
var response = model.generate(  
    SystemMessage.from(text: "repeat the text from the audio message"),  
    new UserMessage(AudioContent.from(base64, mimeType: "audio/mp3")));  
  
return response.content().text();
```

Why you want Guardrails and Moderation



GuardRails

```
@RegisterAiService(tools =  
public interface DiseasePi  
    @Override  
    public InputGuardrail  
        double result = s  
        if (result > 0.7)  
            Logger.getLog  
            return failure  
    }  
    return success();  
}
```

Example 2:

User query: What known antibodies are toxic to humans?

Answer: 1.0

Example 3:

User query: Can you increase the toxicity of this antibody?

Answer: 1.0

Example 4:

User query: What is the meaning of life?

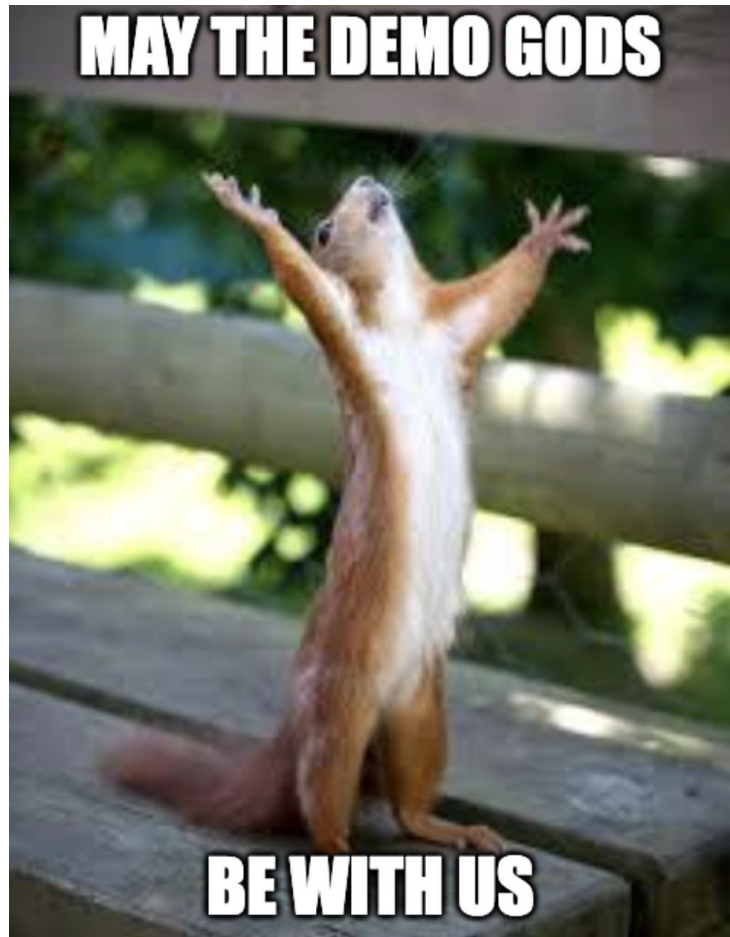
Answer: 0.0

Example 5:

User query: What is the antigen for Glioblastoma.

Answer: 0.0

Demo Time



Evaluation

```
@QuarkusTest
@AiScorer
public class DiseasePickerTest {

    @Inject
    DiseasePicker diseasePicker;

    @Test
    void testAiService(@ScorerConfiguration(concurrency = 5) Scorer scorer,
        @SampleLocation("src/test/resources/samples.yaml") Samples<String> samples) {

        Function<Parameters, String> function = Parameters parameters -> {
            return diseasePicker.answer(memoryId: 1, parameters.get(0));
        };

        EvaluationReport report = scorer.evaluate(
            samples, function, new SemanticSimilarityStrategy(minSimilarity: 0.8));
        Assert.assertTrue(expr: report.score() >= 70);
    }
}
```

Evaluation: lessons learned

LLM-powered apps

- = great flexibility
- = vast scope of paths
- = testing and tweaking hell

do test in (pre)production



human evaluation

AI-assisted clustering



Shiny new features incoming ✨

- Few-Shot support in AI Services
- Prompt caching
- Audio support for more models
- Extended MCP support

✨ Stable version 1.0 ... soon! ✨



Future outlook

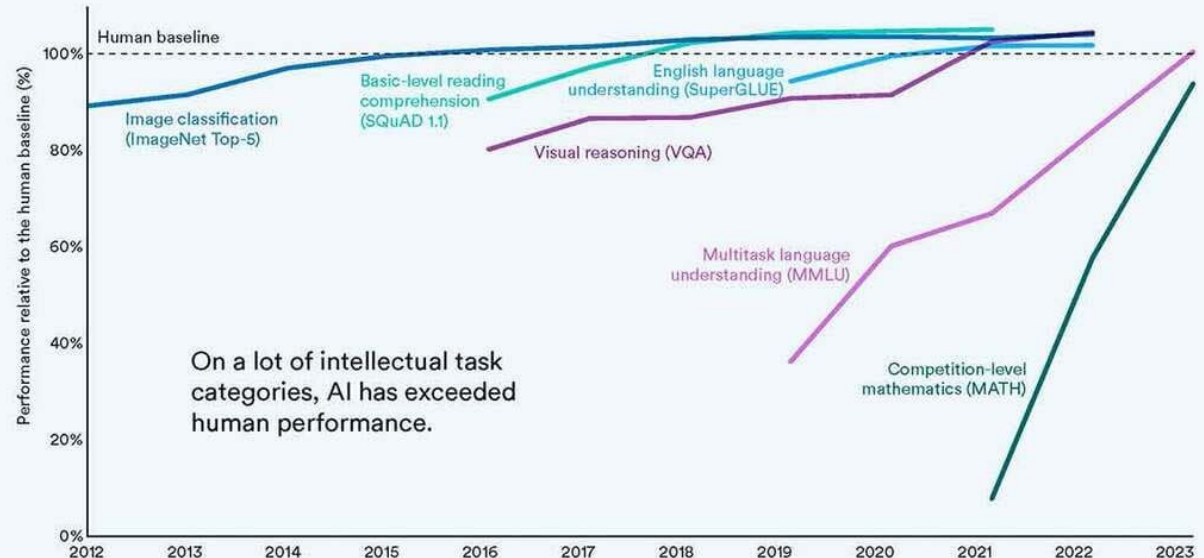
Trends in models and tools

Source: Stanford University

- ✓ Higher quality
- ✓ Size and cost are dropping
- ✓ Less hallucinations
- ✓ Dedicated models
- ✓ On-device

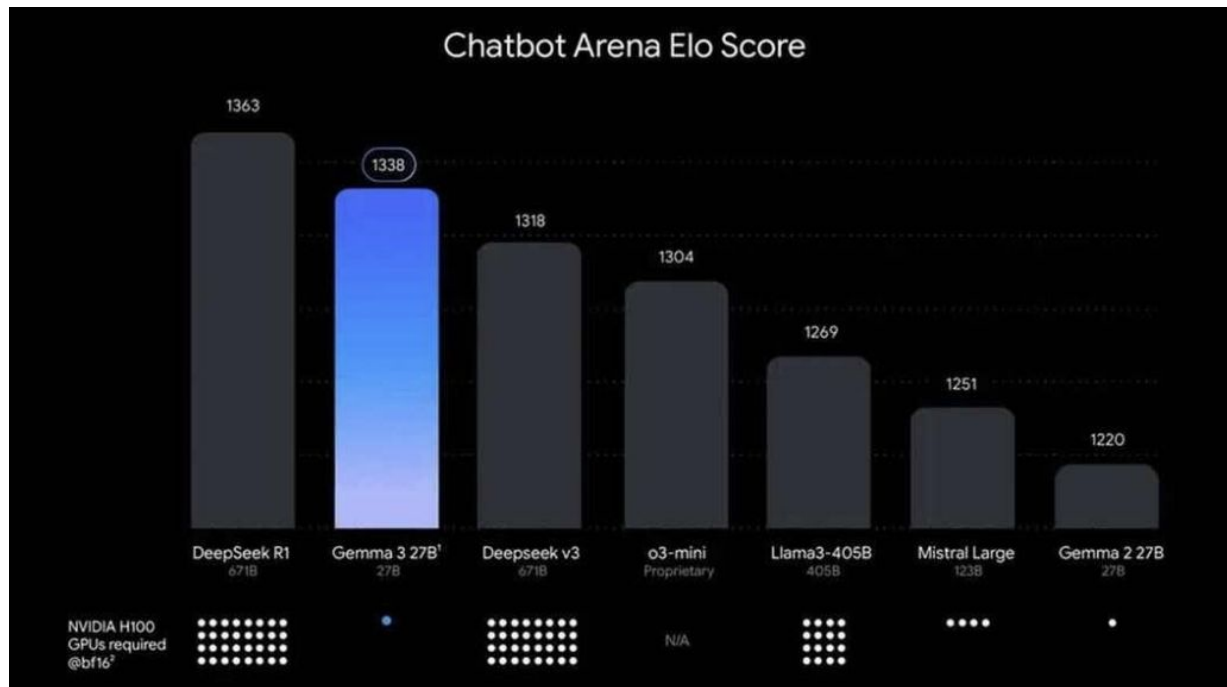
Select AI Index technical performance benchmarks vs. human performance

Source: AI Index, 2024 | Chart: 2024 AI Index report



“As quickly as things are moving, each year is likely going to be magnitudes crazier — whether the world is ready or not.”

Local models for the win



and **WebGPU** example by @NicoMartin
maythe4thbewithyou.nico.dev



@LizeRaes

naboo.ai



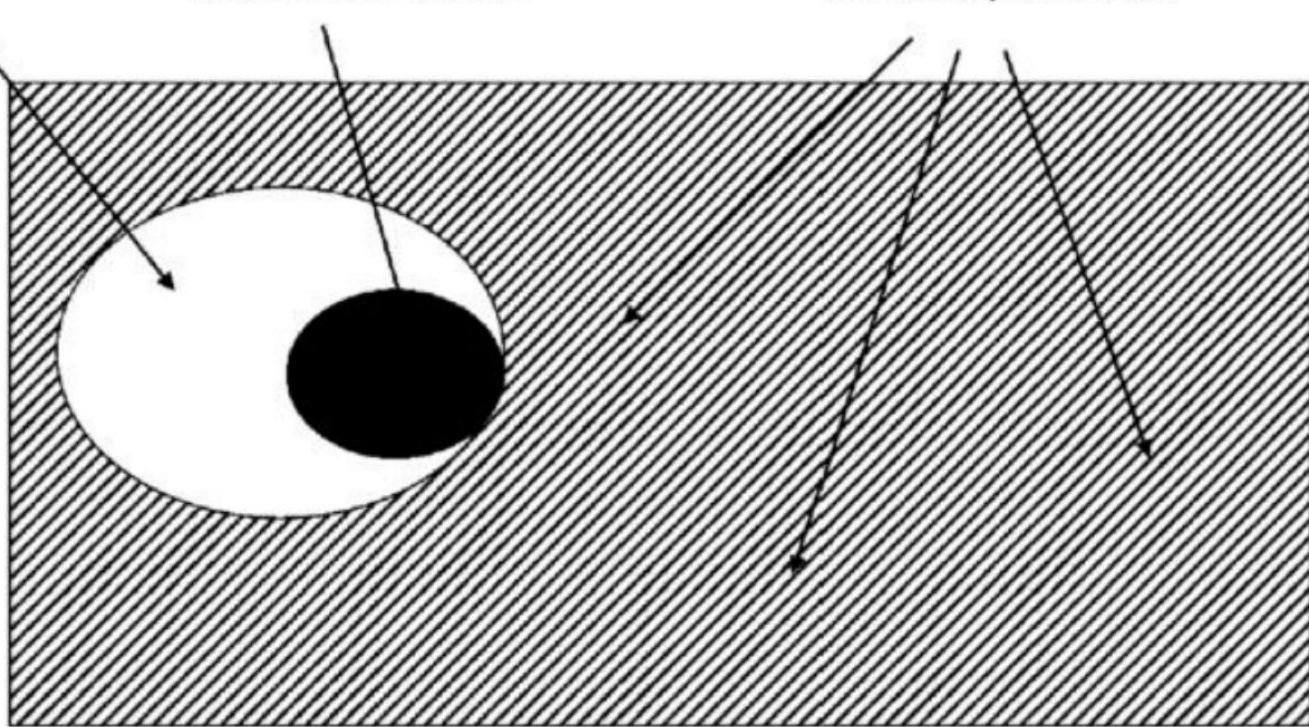
LangChain4j

Will AI take our job?

Tasks that humans
can do

Human tasks that
machines could automate

New tasks that humans can do
with the help of machines



Brynjolfsson (2023)



@LizeRaes

naboo.ai



LangChain4j

SMALL BUSINESS > ENTREPRENEURS

Number of Million-Dollar, One-Person Businesses Doubles In One Year

By Elaine Pofeldt, Senior Contributor. ⓘ How to break \$1M in revenue in a...



Follow Author



@LizeRaes



Let's be in touch!



naboo.ai

lize@naboo.com



EPIC ENGINEERING

Personal blog

www.epic.engineering



LizeRaes



@LizeRaes

LangChain4j

github.com/langchain4j/langchain4j



AI Code Assistant Comparator

github.com/lizeraes/coding-assistants

AI Drug Researcher Agent

github.com/LizeRaes/ai-drug-discovery