

# Intro to Practical AI

Vyacheslav Koldovskyy





# Vyacheslav Koldovskyy

- Ph.D, Assoc. Professor
- 20+ years in IT
- SoftServe Academy Competence Manager
- Certified Google Cloud Professional
- Architect
- Practical AI Enthusiast

# LEARNING & CERTIFICATION

## AI courses

Start your career path with training courses by SoftServe University. Learn AI online under the guidance of true professionals. Choose your opportunities for development in programming!

TYPE	-
<input type="checkbox"/> Complete Programs	
<input type="checkbox"/> Project-Oriented Programs	+
<input type="checkbox"/> Course	+
<input type="checkbox"/> Retraining	
<input type="checkbox"/> Self-Study	

DIRECTIONS	-
<input checked="" type="checkbox"/> Technology & Development	-

<input type="checkbox"/> .NET Development
<input type="checkbox"/> Automation Testing
<input type="checkbox"/> C/C++ Development
<input type="checkbox"/> DevOps Development
<input type="checkbox"/> Java Development
<input type="checkbox"/> Quality Assurance & Control
<input type="checkbox"/> Database
<input type="checkbox"/> WebUI Development
<input type="checkbox"/> Ruby Development
<input type="checkbox"/> Python Development
<input type="checkbox"/> General
<input type="checkbox"/> Robotics
<input checked="" type="checkbox"/> AI
<input type="checkbox"/> Cybersecurity
<input type="checkbox"/> Kotlin

<input type="checkbox"/> Project & Program Management	+
<input type="checkbox"/> Business Analysis & Product Management	+

SEARCH 3 results

AI X

Clear all filters

ONLINE FEE-BASED COURSE OPEN

### ARTIFICIAL INTELLIGENCE FOR EVERYONE

Learn the basics of AI to make life easier

NOVEMBER 11, 2024 - DECEMBER 13, 2024

APPLY

ONLINE FEE-BASED COURSE OPEN

### ARTIFICIAL INTELLIGENCE (AI) FUNDAMENTALS

Master the skills of developing neural networks and genetic algorithms \ Le...

JANUARY 20, 2025 - MARCH 07, 2025

APPLY

ONLINE FREE COURSE OPEN

### GEN AI LAB: ONBOARDING COURSE

Learn how to independently build, evaluate, and deploy scalable Gen AI solu...

APPLY

What is AI?

What is Generative AI?

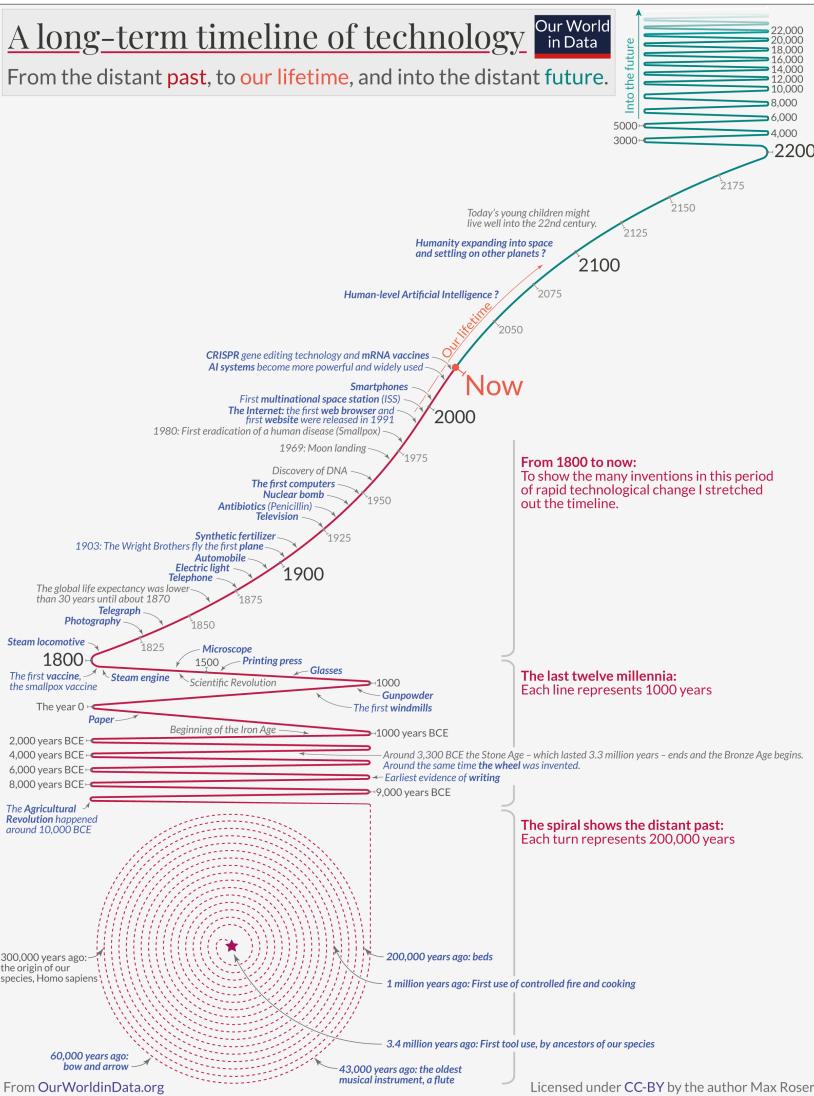
What is Practical AI?

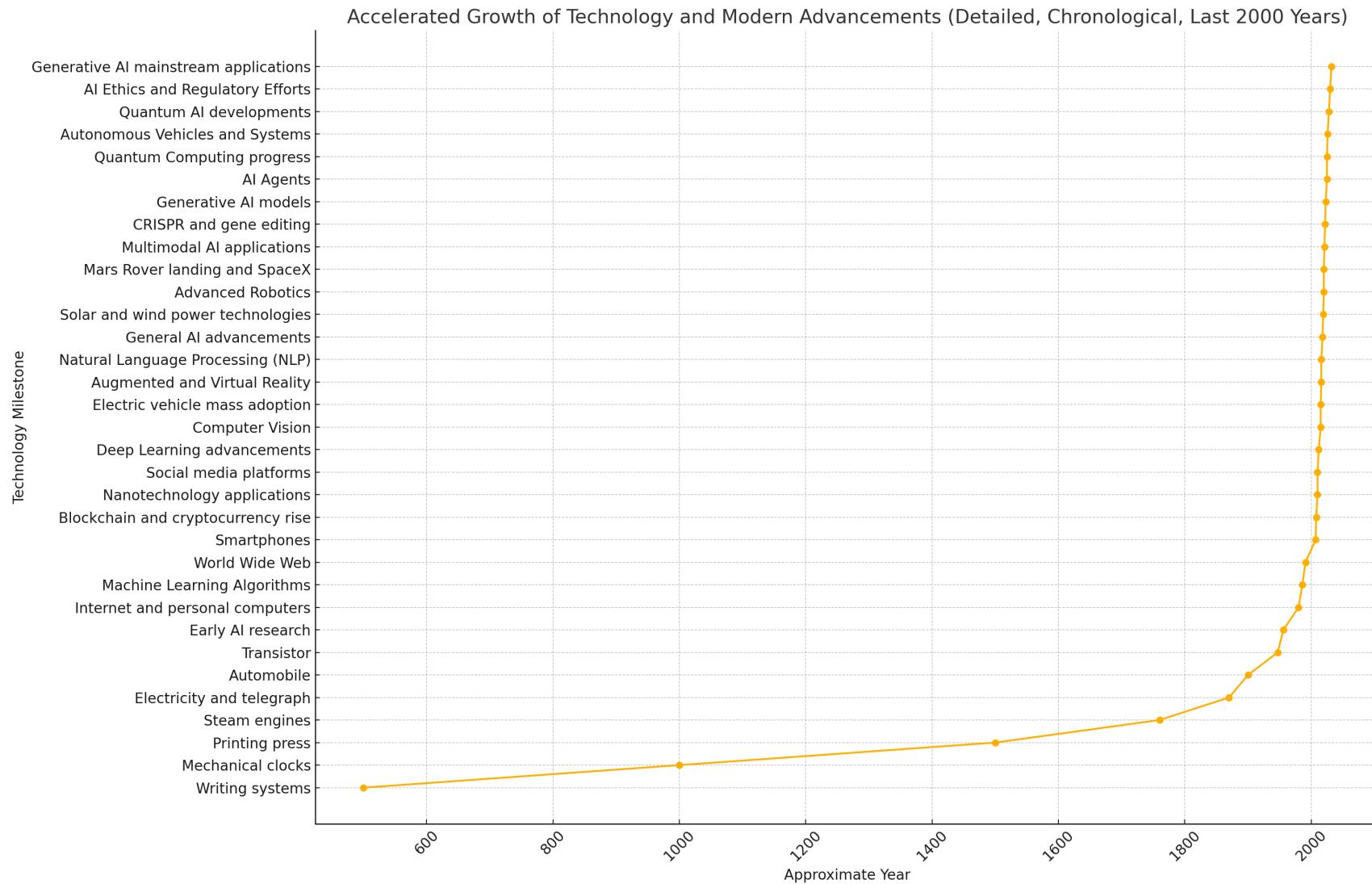
# What is Disruptive Technology?

# A long-term timeline of technology

From the distant **past**, to **our lifetime**, and into the distant **future**.

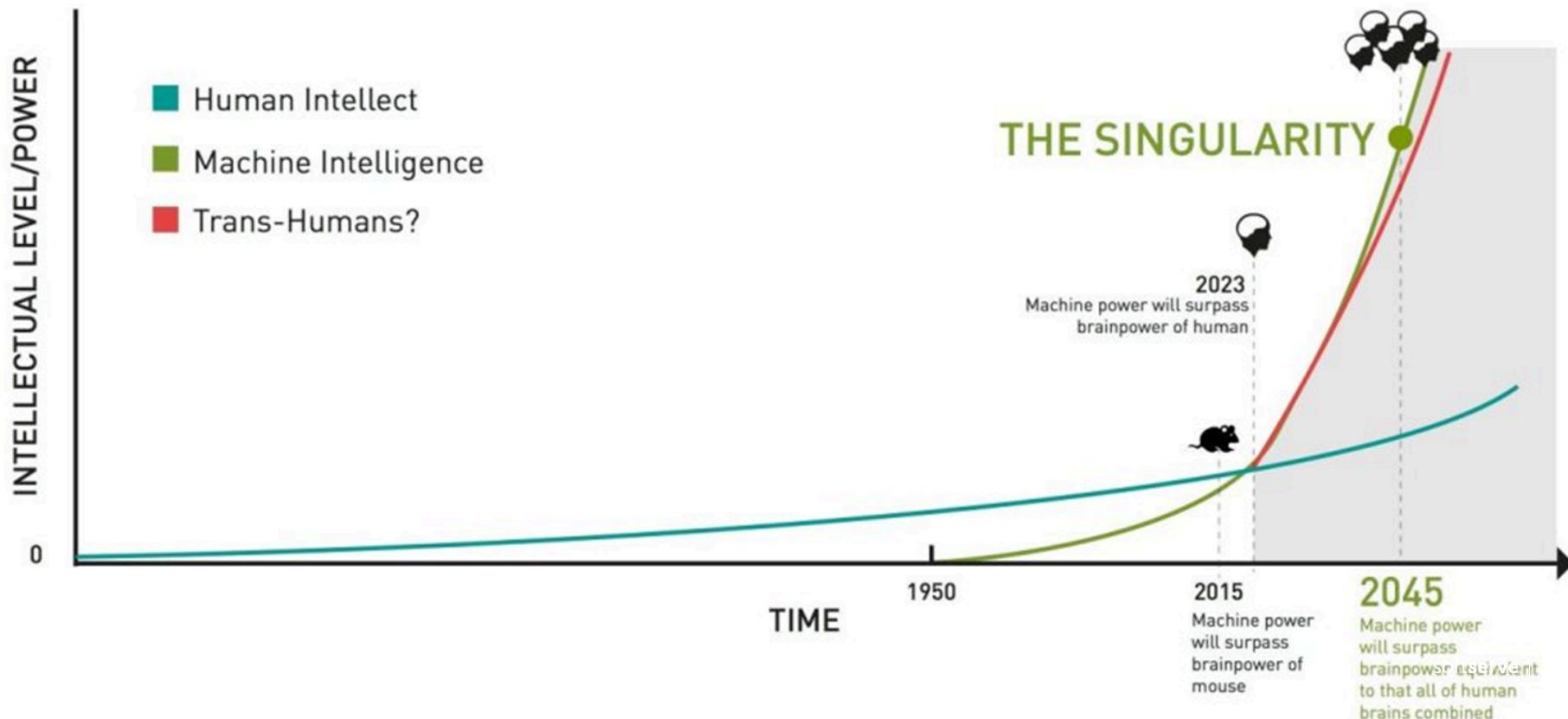
Our World  
in Data





# SINGULARITY TIMELINE

Rise in human intellect could be driven by integrating with machines in the future



What is AI?  
AI vs Non-AI?

# AI vs Non-AI Applications

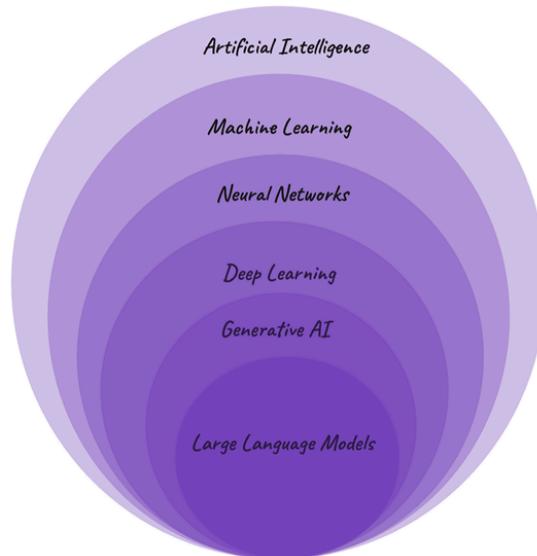
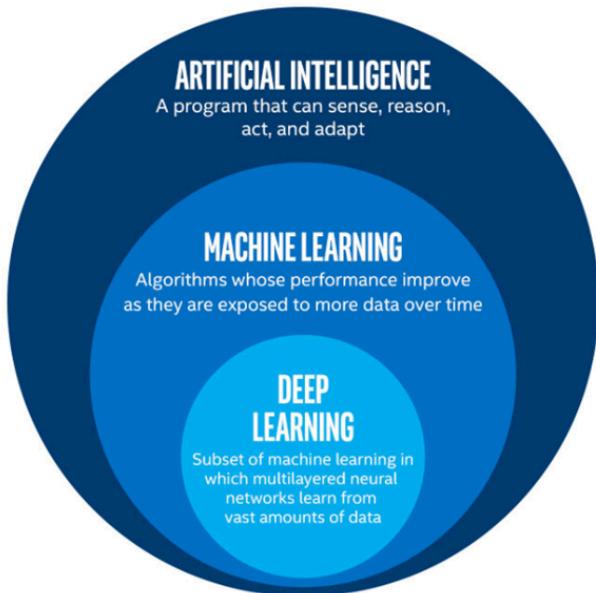
## Non-AI

- Fixed, rule-based logic
- Predictable outputs
- Manual updates required
- Limited to programmed scenarios
- Static behavior patterns

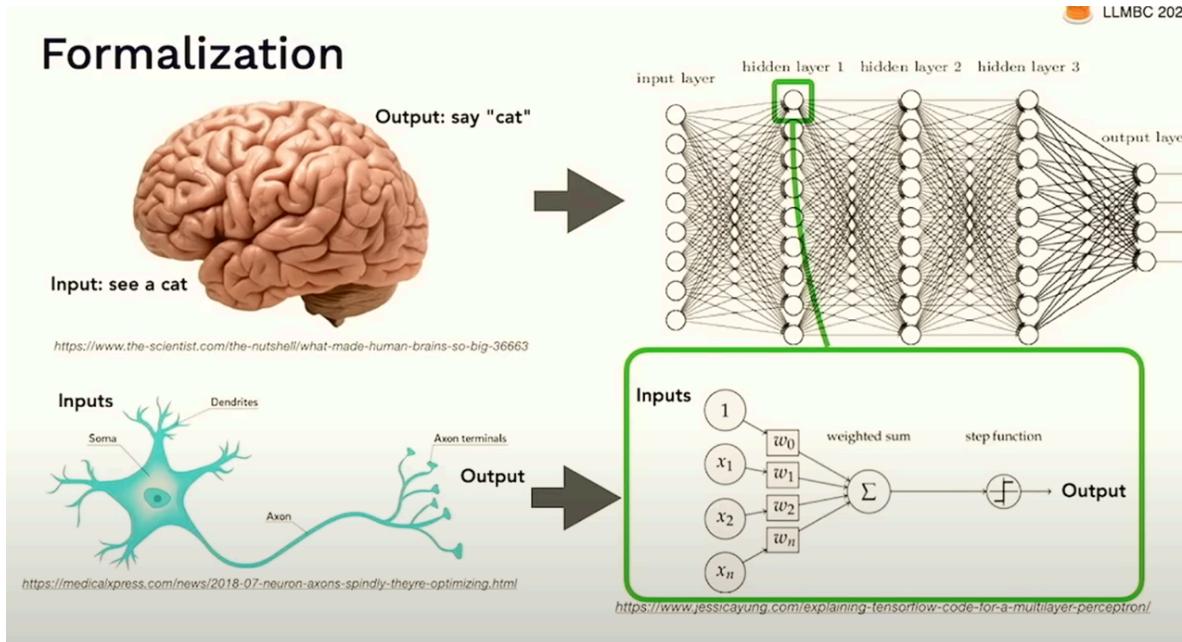
## AI

- Learning from data
- Adaptive responses
- Self-improving capabilities
- Handles unexpected scenarios
- Dynamic behavior patterns

# AI Landscape



# AI Concepts



# Neural Network

The neural network is a set of matrices

Called "parameters" or "weights"

LLMBC 2023

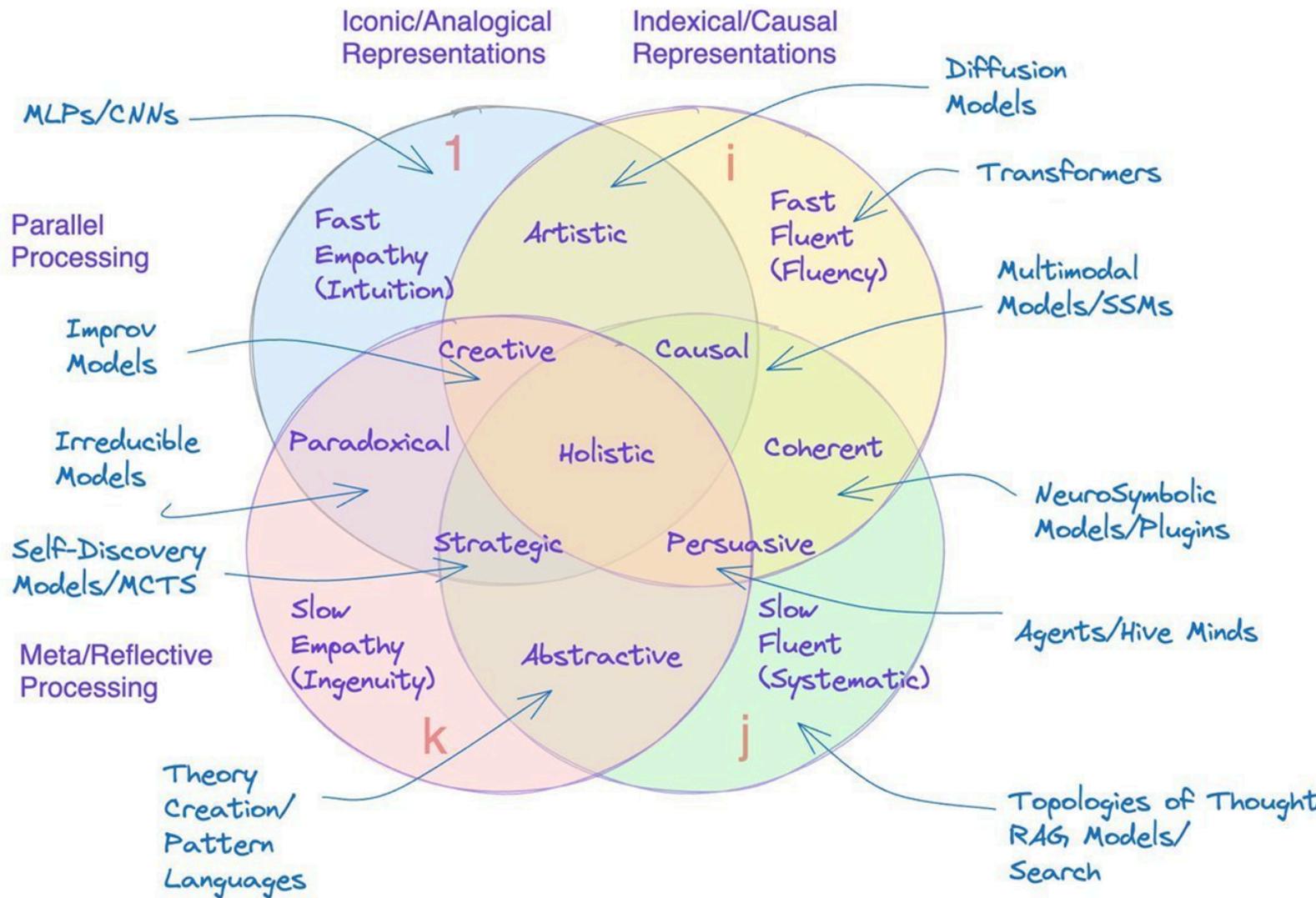
The diagram illustrates a neural network structure with five layers: input layer, hidden layer 1, hidden layer 2, hidden layer 3, and output layer. The input layer has 8 nodes, hidden layer 1 has 5 nodes, hidden layer 2 has 4 nodes, hidden layer 3 has 3 nodes, and the output layer has 2 nodes. Every node in one layer is connected to every node in the subsequent layer. Three vertical blue rectangles highlight the connections between the input layer and hidden layer 1, hidden layer 1 and hidden layer 2, and hidden layer 2 and hidden layer 3 respectively.

[[[0.32 0.866 0.151 0.555]  
[0.58 0.841 0.835 0.18 ]  
[0.32 0.866 0.151 0.555]  
[0.467 0.414 0.562 0.555]  
[0.186 0.937 0.131 0.448]  
[0.892 0.567 0.111 0.678]  
[0.483 0.478 0.424 0.844]  
[0.774 0.963 0.213 0.569]  
[0.014 0.509 0.21 0.26 ]  
[0.142 0.991 0.105 0.211]  
[0.789 0.242 0.774 0.019]]]

NN operations are just matrix multiplications.  
GPUs are really fast at matrix multiplications.

14

# AI models and human Abilities



# Types of Learning of AI Models

## Supervised Learning

- Learns from labeled data
- Predicts outcomes based on examples
- E.g., Image classification, spam detection

## Reinforcement Learning

- Learns through trial and error
- Maximizes rewards over time
- E.g., Game AI, robotics

## Unsupervised Learning

- Finds patterns in unlabeled data
- Discovers hidden structures
- E.g., Clustering, anomaly detection

## Deep Learning

- Uses neural networks with many layers
- Automatically learns features
- E.g., Computer vision, NLP

**NEW** Use Ollama with GGUF Models from the HF Hub



# The AI community building the future.

The platform where the machine learning community collaborates on models, datasets, and applications.

Tasks   Libraries   Datasets   Languages   Licenses   Other

Filter Tasks by name

Multimodal

Text-to-Image   Image-to-Text  
Text-to-Video   Visual Question Answering  
Document Question Answering   Graph Machine Learning

Computer Vision

Depth Estimation   Image Classification  
Object Detection   Image Segmentation  
Image-to-Image   Unconditional Image Generation  
Video Classification   Zero-Shot Image Classification

Natural Language Processing

Text Classification   Token Classification  
Table Question Answering   Question Answering  
Zero-Shot Classification   Translation  
Summarization   Conversational  
Text Generation   Text2Text Generation  
Sentence Similarity

Audio

Text-to-Speech   Automatic Speech Recognition  
Audio-to-Audio   Audio Classification  
Voice Activity Detection

Tabular

Tabular Classification   Tabular Regression

Reinforcement Learning

Reinforcement Learning   Robotics

Models 469,541   Filter by name

meta-llama/Llama-2-70b  
Text Generation • Updated 4 days ago • ↓ 25.2k • ❤ 64

stabilityai/stable-diffusion-xl-base-0.9  
Updated 6 days ago • ↓ 2.01k • ❤ 393

openchat/openchat  
Text Generation • Updated 2 days ago • ↓ 1.3k • ❤ 136

lillyasviel/ControlNet-v1.1  
Updated Apr 26 • ⚡ 1.87k

cerspense/zeroscope\_v2\_XL  
Updated 3 days ago • ↓ 2.66k • ❤ 334

meta-llama/Llama-2-13b  
Text Generation • Updated 4 days ago • ↓ 328 • ❤ 64

tiiuae/falcon-40b-instruct  
Text Generation • Updated 27 days ago • ↓ 288k • ❤ 899

WizardLM/WizardCoder-15B-V1.0  
Text Generation • Updated 3 days ago • ↓ 12.5k • ❤ 332

CompVis/stable-diffusion-v1-4  
Text-to-Image • Updated about 17 hours ago • ↓ 448k • ❤ 5.72k

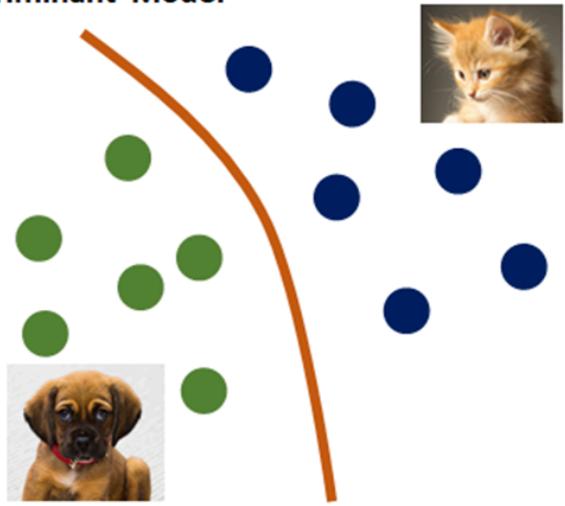
stabilityai/stable-diffusion-2-1  
Text-to-Image • Updated about 17 hours ago • ↓ 782k • ❤ 2.81k

Salesforce/xgen-7b-8k-inst  
Text Generation • Updated 4 days ago • ↓ 6.18k • ❤ 57

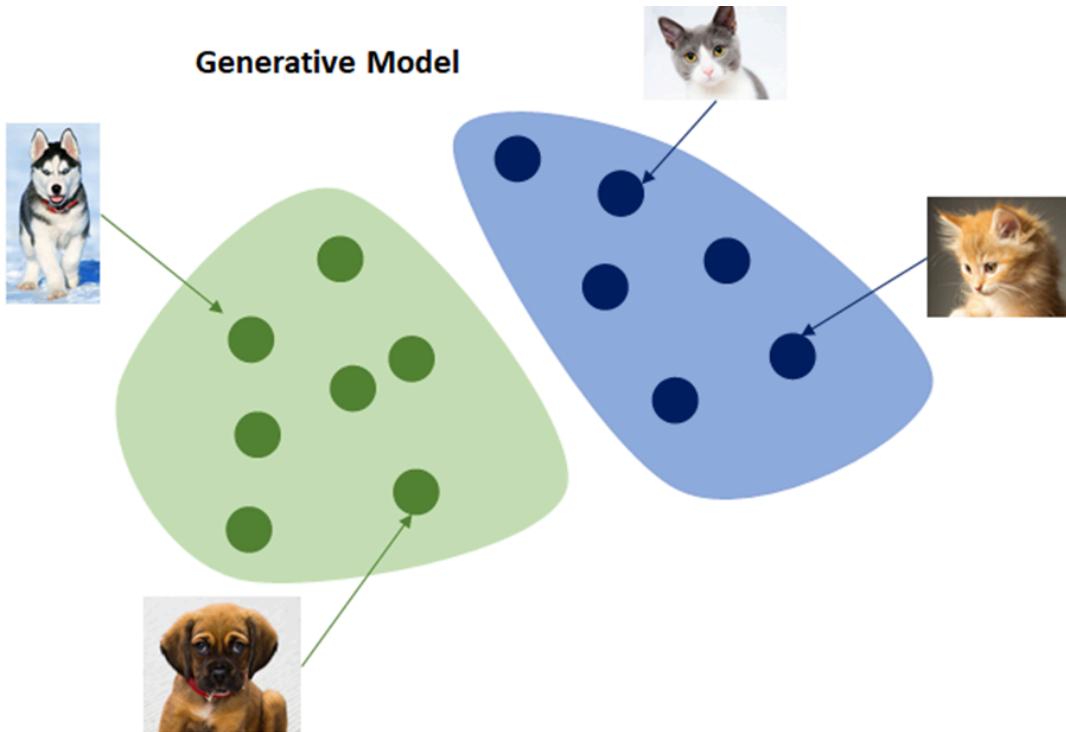
softserve | academ | monste | Update | THUDM | Update | tiiuae | Update

# Generative AI

**Discriminant Model**



**Generative Model**



# Generating Myself

Leap Vyacheslav's Project ▾ View Usage

Workflows Images Music API Keys Settings

Models

Models / Vyacheslav Koldovskyy Improved

Your model has finished training. Use in Playground

Name: Vyacheslav Koldovskyy  
Id: a8dd24c8-784a-4c  
Type: Man

Important Details

- Training usually takes between 20 to 40 minutes.
- When your model is ready, we'll send you an email.
- No nudes / NSFW images allowed.

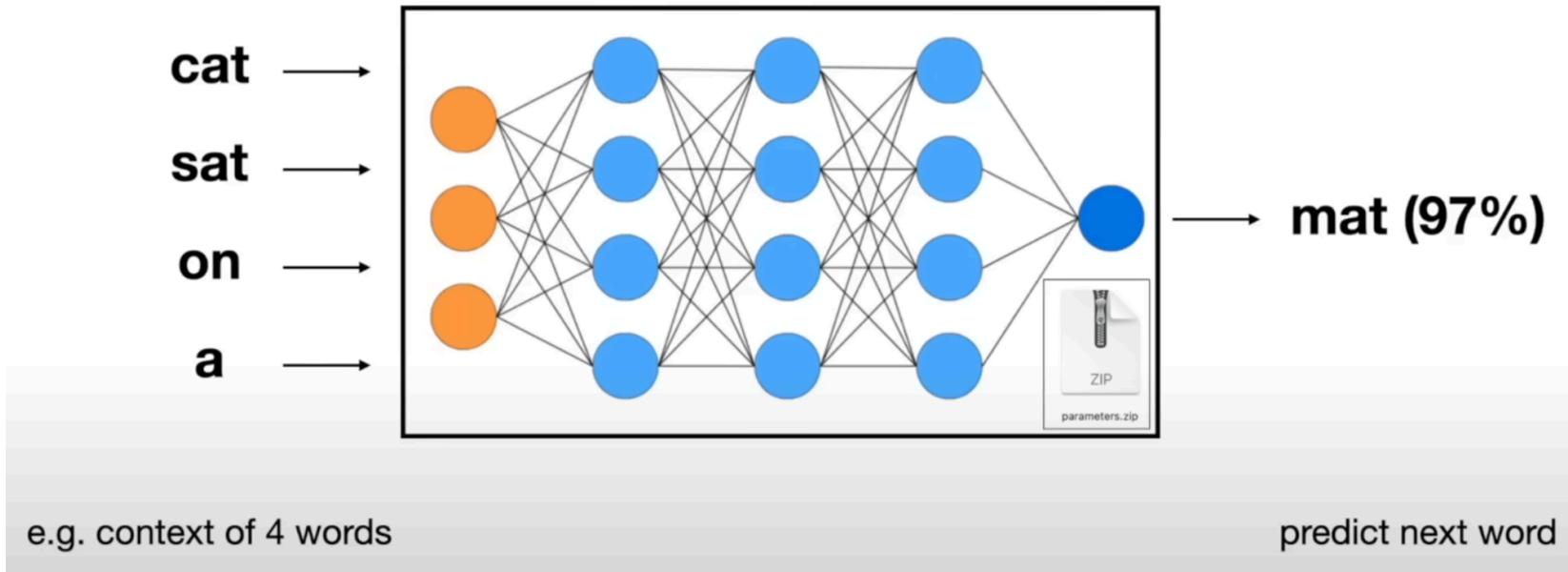


A screenshot of the Leap AI web interface. The top navigation bar shows 'Leap' and 'Vyacheslav's Project'. Below it are tabs for 'Workflows', 'Images' (which is selected), 'Music', 'API Keys', and 'Settings'. Under the 'Models' section, it says 'Your model has finished training.' and 'Use in Playground'. It shows the model's name 'Vyacheslav Koldovskyy', ID 'a8dd24c8-784a-4c', and type 'Man'. An 'Important Details' section lists: 'Training usually takes between 20 to 40 minutes.', 'When your model is ready, we'll send you an email.', and 'No nudes / NSFW images allowed.' Below this are three rows of generated portrait images of the same man with different expressions and backgrounds.



# LLM - Large Language Models

# Why Language Model?



# Why Large?

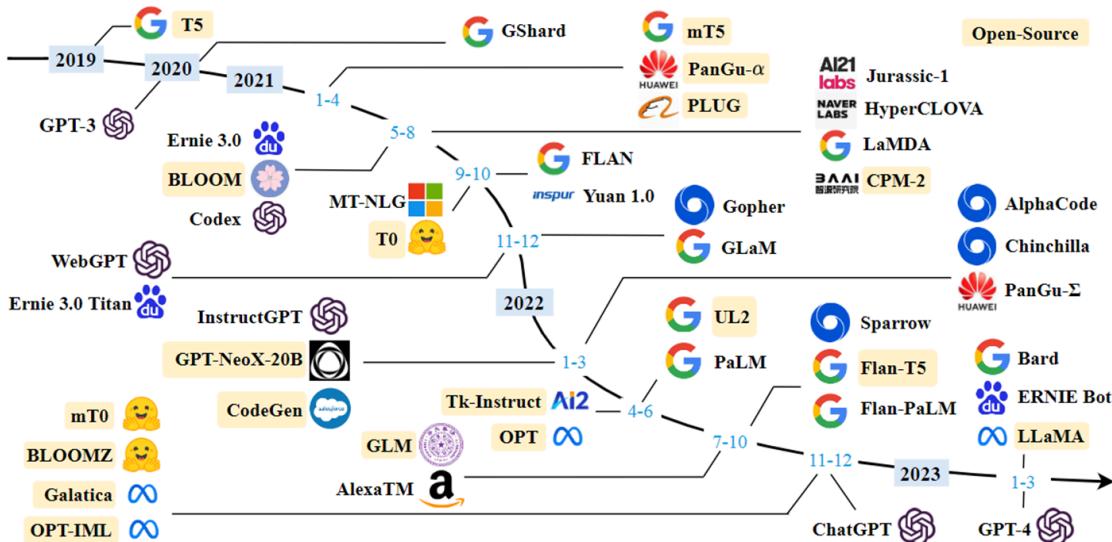
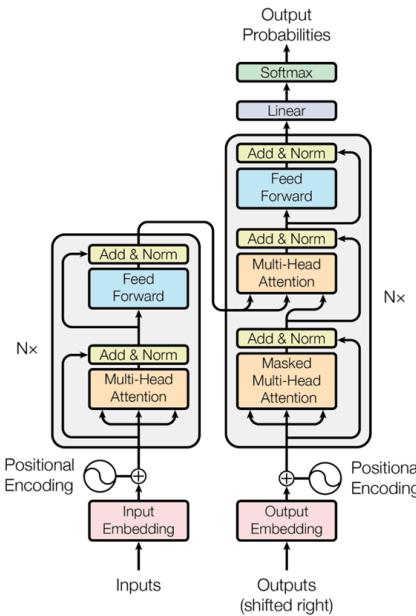


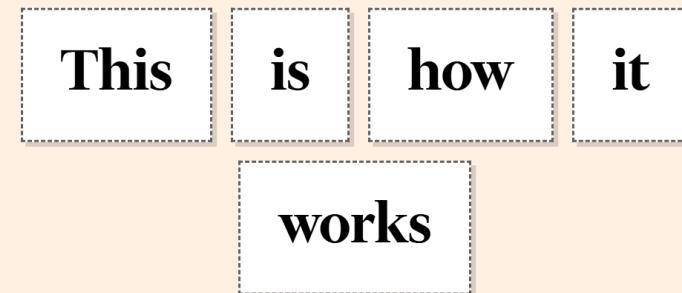
Fig. 1. A timeline of existing large language models (having a size larger than 10B) in recent years. We mark the open-source LLMs in yellow color.

# Transformer



Artificial Intelligence

## Generative AI exists because of the transformer



By Visual Storytelling Team and Madhumita Murgia in London SEPTEMBER 12 2023

Which LLM is the best?

# Chatbot Arena LLM Leaderboard: Community-driven Evaluation for Best LLM and AI chatbots

[Blog](#) | [GitHub](#) | [Paper](#) | [Dataset](#) | [Twitter](#) | [Discord](#) | [Kaggle Competition](#)

Chatbot Arena ([Imarena.ai](#)) is an open-source platform for evaluating AI through human preference, developed by researchers at UC Berkeley [SkyLab](#) and [LMSYS](#). With over 1,000,000 user votes, the platform ranks best LLM and AI chatbots using the Bradley-Terry model to generate live leaderboards. For technical details, check out our paper.

Chatbot Arena thrives on community engagement — cast your vote to help improve AI evaluation!

New Launch! Copilot Arena: [VS Code Extension](#) to compare Top LLMs

Arena NEW: Overview Arena (Vision) Arena-Hard-Auto Full Leaderboard

Total #models: 157. Total #votes: 2,118,851. Last updated: 2024-10-22.

Code to recreate leaderboard tables and plots in this [notebook](#). You can contribute your vote at [Imarena.ai](#)!

Category		Overall		Apply filter		Overall Questions					
				<input type="checkbox"/> Style Control	<input type="checkbox"/> Show Deprecated	#models: 157 (100%)	#votes: 2,118,851 (100%)				
Rank* (UB)	▲ Rank (StyleCtrl)	▲ Model	▲	Arena Score	▲ 95% CI	▲ Votes	▲ Organization	▲ License	▲ Knowledge Cutoff	▲	▲
1	1	<a href="#">ChatGPT-4o-latest (2024-09-03)</a>		1340	+4/-5	31927	OpenAI	Proprietary	2023/10		
1	1	<a href="#">o1-preview</a>		1337	+4/-5	19924	OpenAI	Proprietary	2023/10		
3	5	<a href="#">o1-mini</a>		1309	+5/-4	21425	OpenAI	Proprietary	2023/10		
3	3	<a href="#">Gemini-1.5-Pro-002</a>		1303	+5/-5	13957	Google	Proprietary	Unknown		
4	3	<a href="#">Gemini-1.5-Pro-Exp-0827</a>		1299	+4/-3	32393	Google	Proprietary	2023/11		
6	8	<a href="#">Grok-2-08-13</a>		1290	+4/-4	39193	xAI	Proprietary	2024/3		
6	11	<a href="#">Yi-Lightning</a>		1286	+4/-4	18864	01 AI	Proprietary	Unknown		
6	3	<a href="#">GPT-4o-2024-05-13</a>		1285	+3/-2	101733	OpenAI	Proprietary	2023/10		
9	14	<a href="#">GLM-4-Plus</a>		1275	+5/-4	18695	Zhipu AI	Proprietary	Unknown		
9	17	<a href="#">GPT-4o-mini-2024-07-18</a>		1274	+4/-3	41440	OpenAI	Proprietary	2023/10		
9	20	<a href="#">Gemini-1.5-Flash-002</a>		1271	+5/-5	11155	Google	Proprietary	Unknown		
9	13	<a href="#">Gemini-1.5-Flash-Exp-0827</a>		1269	+4/-3	25509	Google	Proprietary	2023/11		
9	26	<a href="#">Llama-3.1-Nemotron-70b-Instruct</a>		1267	+7/-7	4416	Nvidia	Llama 3.1	2023/12		
10	5	<a href="#">Claude_3.5_Sonnet</a>		1268	+3/-2	79710	Anthropic	Proprietary	2024/4		
11	24	<a href="#">Grok-2-Mini-08-13</a>		1267	+3/-4	32949	xAI	Proprietary	2024/3		
11	7	<a href="#">Meta-Llama-3.1-409b-Instruct-fp8</a>		1267	+4/-3	41936	Meta	Llama 3.1 Community	2023/12		
11	7	<a href="#">Gemini_Advanced_App_(2024-05-14)</a>		1266	+4/-2	52236	Google	Proprietary	Online		

[Arena \(battle\)](#)[Arena \(side-by-side\)](#)[Direct Chat](#)[Leaderboard](#)[About Us](#)

## Chatbot Arena LLM Leaderboard: Community-driven Evaluation for Best LLM and AI chatbots

[Blog](#) | [GitHub](#) | [Paper](#) | [Dataset](#) | [Twitter](#) | [Discord](#) | [Kaggle Competition](#)

Chatbot Arena ([imarena.ai](#)) is an open-source platform for evaluating AI through human preference, developed by researchers at UC Berkeley [SkyLab](#) and [LMSYS](#). With over 1,000,000 user votes, the platform ranks best LLM and AI chatbots using the Bradley-Terry model to generate live leaderboards. For technical details, check out our [paper](#).

Chatbot Arena thrives on community engagement — cast your vote to help improve AI evaluation!

New Launch! Copilot Arena: VS Code Extension to compare Top LLMs

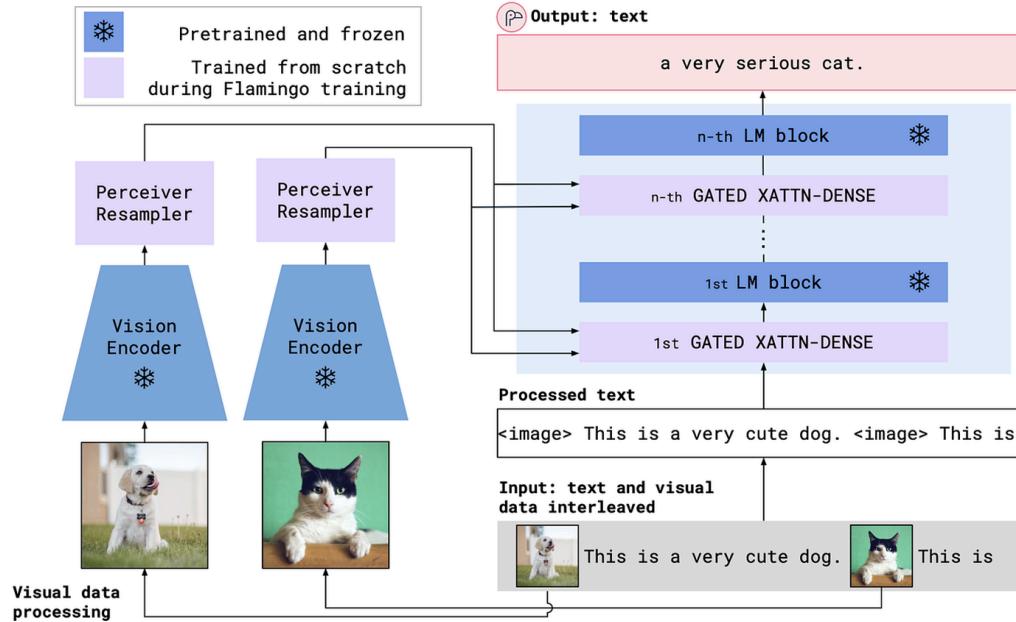
[Arena](#) | [NEW: Overview](#) | [Arena \(Vision\)](#) | [Arena-Hard-Auto](#) | [Full Leaderboard](#)

Total #models: 157. Total #votes: 2,118,851. Last updated: 2024-10-22.

Code to recreate leaderboard tables and plots in this [notebook](#). You can contribute your vote at [imarena.ai](#)!

Category	Coding: whether conversation contains code snippets									
Coding	#models: 151 (96%) #votes: 408,883 (19%)									
Rank* (UB)	Delta	Model	Arena Score	95% CI	Votes	Organization	License	Knowledge Cutoff		
1	2	<a href="#">o1-mini</a>	1369	+8/-12	4628	OpenAI	Proprietary	2023/10		
1	0	<a href="#">o1-preview</a>	1355	+11/-9	3951	OpenAI	Proprietary	2023/10		
2	-1	<a href="#">ChatGPT-4o-latest_(2024-09-03)</a>	1339	+10/-9	6417	OpenAI	Proprietary	2023/10		
4	2	<a href="#">Yi-Lightning</a>	1303	+10/-11	3694	Q1 AI	Proprietary	Unknown		
4	6	<a href="#">Claude_3.5_Sonnet</a>	1296	+8/-7	15542	Anthropic	Proprietary	2024/4		
4	2	<a href="#">GPT-4o-2024-05-13</a>	1292	+7/-6	21296	OpenAI	Proprietary	2023/10		
4	-1	<a href="#">Gemini-1.5-Pro-002</a>	1290	+10/-9	2976	Google	Proprietary	Unknown		
4	0	<a href="#">Gemini-1.5-Pro-Exp-0027</a>	1288	+9/-8	6279	Google	Proprietary	2023/11		
4	2	<a href="#">Grok-2-08-13</a>	1287	+6/-7	7912	xAI	Proprietary	2024/3		
4	5	<a href="#">GLM-4-Plus</a>	1287	+11/-12	3890	Zhipu AI	Proprietary	Unknown		
4	7	<a href="#">Meta-Llama-3.1-405b-Instruct-bf16</a>	1284	+10/-10	2846	Meta	Llama 3.1 Community	2023/12		
5	4	<a href="#">GPT-4o-mini-2024-07-18</a>	1285	+6/-9	8023	OpenAI	Proprietary	2023/10		
5	9	<a href="#">Qwen2.5-72b-Instruct</a>	1283	+9/-9	3323	Alibaba	Qwen	2024/9		
5	14	<a href="#">Deepseek-v2.5</a>	1283	+9/-10	3388	DeepSeek	DeepSeek	Unknown		
6	5	<a href="#">Qwen-Max-0919</a>	1278	+11/-11	2944	Alibaba	Qwen	Unknown		
7	4	<a href="#">Meta-Llama-3.1-405b-Instruct-fp8</a>	1278	+8/-8	8006	Meta	Llama 3.1 Community	2023/12		
7	4	<a href="#">GPT-4o-2024-08-06</a>	1273	+8/-8	6349	OpenAI	Proprietary	2023/10		

# Multimodal LLM



# LLM Usage & Applications

- Text Generation & Summarization
- Code Assistance & Documentation
- Language Translation
- Content Creation & Editing
- Customer Support Automation
- Data Analysis & Insights
- Education
- ... much more

# LLM Limitations

- Hallucinations & Factual Inaccuracies
- Context Window Constraints
- Training Data Cutoff
- Bias & Ethical Concerns
- High Computational Costs
- Limited Real-time Knowledge

# What is Software Engineering?

- Art
- Science
- Craft



## What can I help with?

Що таке програмна інженерія?

- Мистецтво
- Наука
-  - Ремесло



 Create image

 Analyze data

 Help me write

 Summarize text

More



- мистецтво
- Наука
- Ремесло



Програмна інженерія часто визначається як поєднання **мистецтва, науки та ремесла**. Кожен із цих аспектів має свій внесок у формування цієї дисципліни:

1. **Мистецтво:** Програмування часто вимагає творчого підходу для розв'язання проблем. Вибір архітектури, дизайну інтерфейсу, інтуїтивні рішення щодо структури коду та його оптимізації — усе це вимагає креативності та гнучкості, подібно до мистецтва.
2. **Наука:** Програмна інженерія базується на строгих теоретичних принципах, таких як алгоритми, структури даних, математичні моделі, логіка та методи тестування. Це наукова дисципліна, яка розвивається завдяки дослідженням і формальним методам перевірки та доведення правильності рішень.
3. **Ремесло:** Як і будь-яка інша практична діяльність, програмування потребує практичних навичок, дисципліни та досвіду. Інженери вдосконалюють свою майстерність через постійну практику, розв'язуючи проблеми та розробляючи ефективні рішення.

Отже, програмна інженерія охоплює елементи всіх трьох складових, що робить її багатогранною та захоплюючою дисципліною.



Message ChatGPT



# Why Craft?

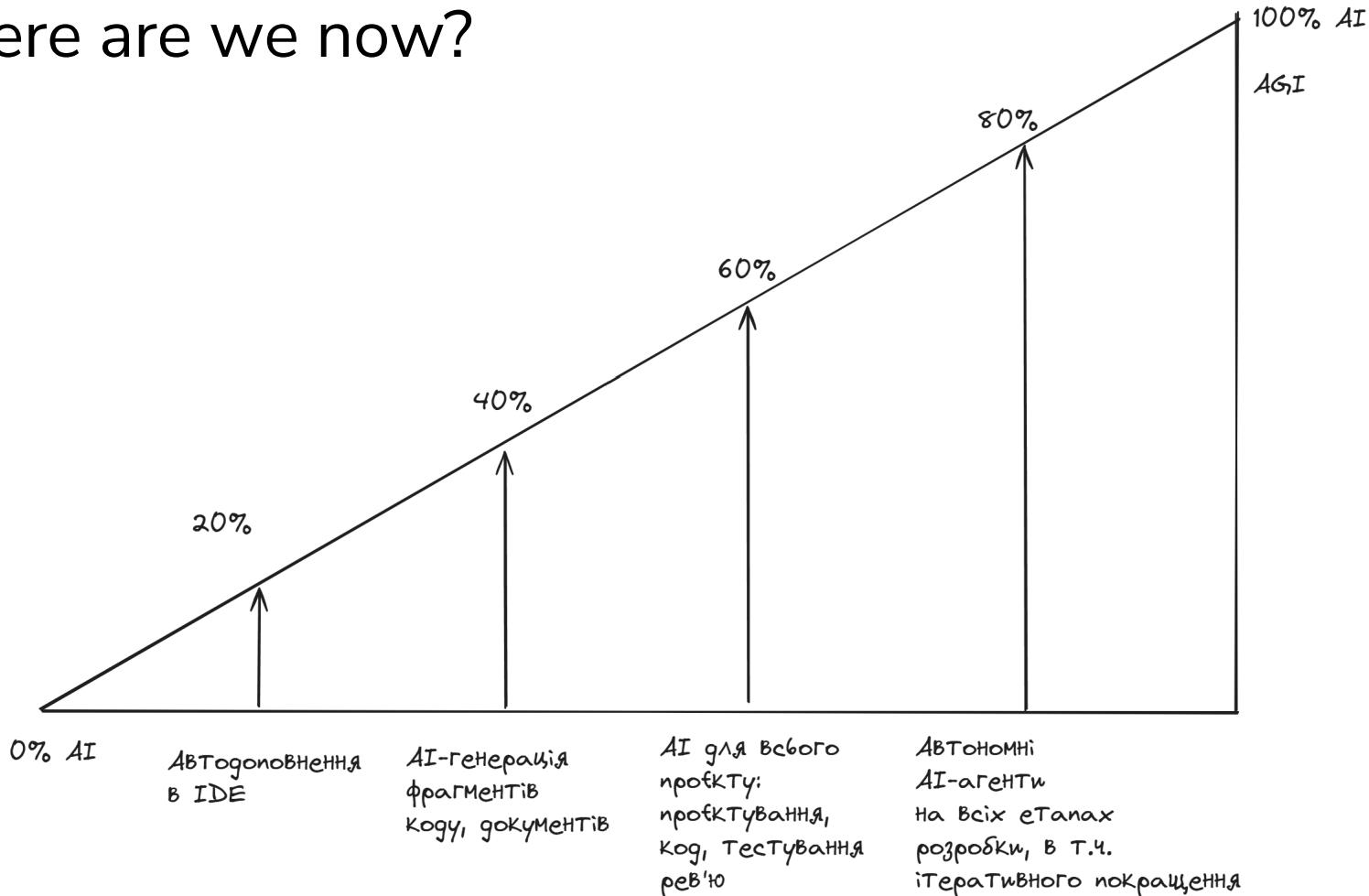
- Algorithms
- Data Structures
- Programming Paradigms
- Software System Architectures
- Patterns, Principles, Practices
- Languages, Libraries, Frameworks
- Tools, Version Control Systems, CI/CD
- OS, Networks, Containers, Cloud Platforms

Main idea: no need to be creative or invent new knowledge, we need to learn and apply what has already been invented

What kind of work does AI handle best?

# The AI Era in Software Engineering

# Where are we now?



PCMag editors select and review products independently. If you buy through affiliate links, we may earn commissions, which help support our t

Home > News > AI

# AI Now Writes Over 25% of Code at Google

Google CEO Sundar Pichai says the company is using AI to write code, which is then reviewed by engineers. But will that be worth it in the long run?



By [Kate Irwin](#) October 30, 2024

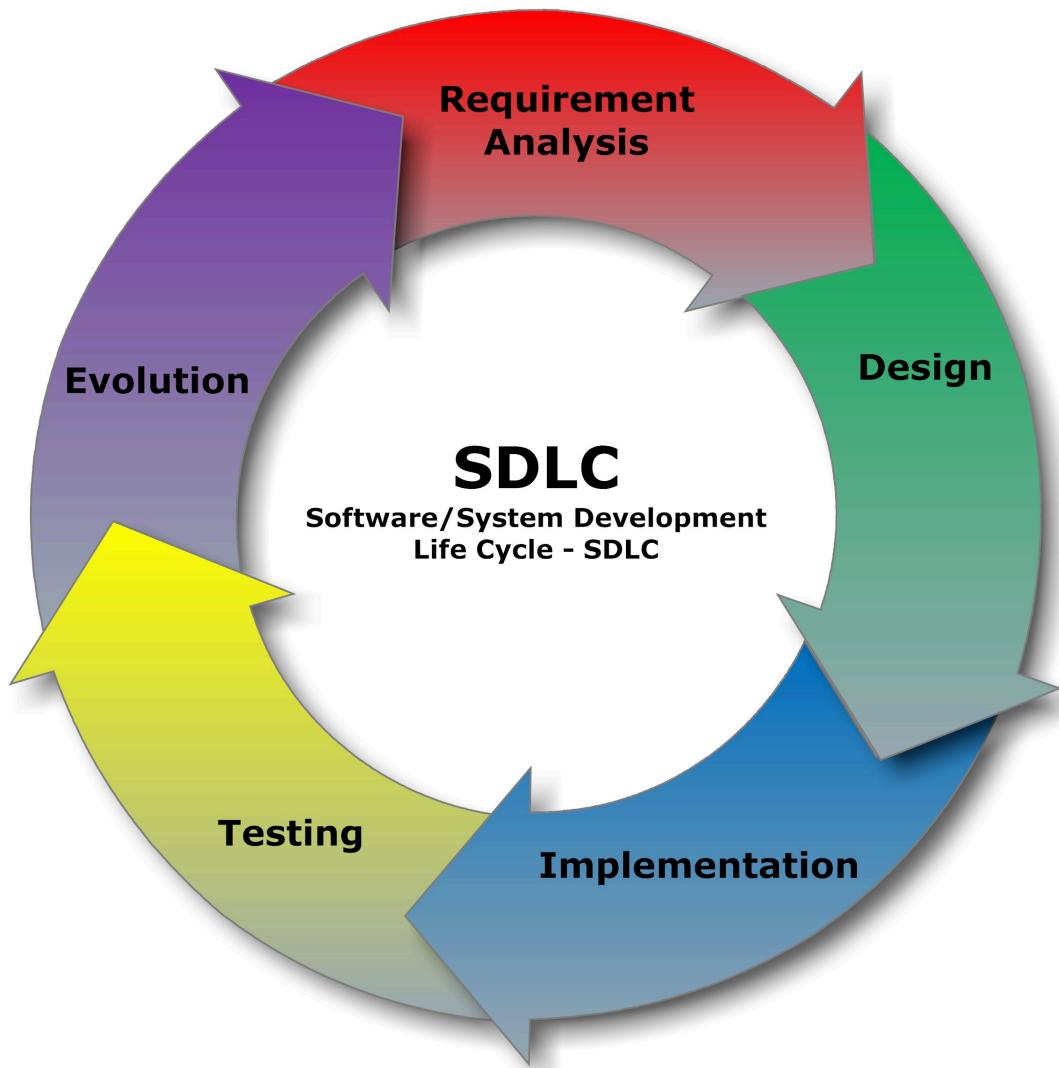
f X g ...



(Credit: Bloomberg / Contributor via Getty Images)

Google is now using AI to write more than a quarter of its new code, according to Google and parent company Alphabet CEO Sundar Pichai, who shared the news during the firm's most recent earnings call.

# Where LLM in SDLC?



# What do we do during requirements analysis and design?

Documents, diagrams, etc.



# What can I help with?

 Message ChatGPT



 Create image

 Summarize text

 Brainstorm

 Make a plan

More



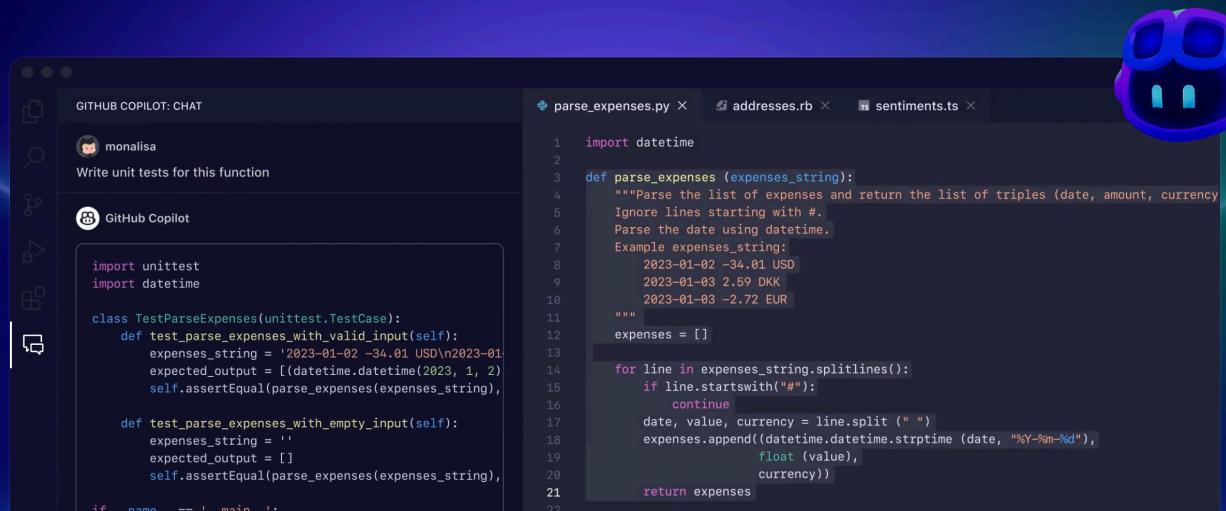
What about implementation?



Copilot Enterprise now available

# The world's most widely adopted AI developer tool.

[Get started with Copilot >](#)



GITHUB COPILOT: CHAT

monalisa  
Write unit tests for this function

GitHub Copilot

```
import unittest
import datetime

class TestParseExpenses(unittest.TestCase):
    def test_parse_expenses_with_valid_input(self):
        expenses_string = '2023-01-02 -34.01 USD\n2023-01-03 2.59 DKK
        expected_output = [(datetime.datetime(2023, 1, 2),
                           float(-34.01),
                           "USD")]
        self.assertEqual(parse_expenses(expenses_string),
                        expected_output)

    def test_parse_expenses_with_empty_input(self):
        expenses_string = ''
        expected_output = []
        self.assertEqual(parse_expenses(expenses_string),
                        expected_output)
```

```
import datetime

def parse_expenses(expenses_string):
    """Parse the list of expenses and return the list of triples (date, amount, currency)
    Ignore lines starting with #.
    Parse the date using datetime.
    Example expenses_string:
        2023-01-02 -34.01 USD
        2023-01-03 2.59 DKK
        2023-01-03 -2.72 EUR
    """
    expenses = []

    for line in expenses_string.splitlines():
        if line.startswith("#"):
            continue
        date, value, currency = line.split(" ")
        expenses.append((datetime.datetime.strptime(date, "%Y-%m-%d"),
                         float(value),
                         currency))

    return expenses
```

A screenshot of a Microsoft Visual Studio Code interface showing a code editor with multiple tabs and a sidebar for a Copilot AI integration.

The code editor displays the file `index.product-list.js` (1, U). The content of the file is:

```
const products = [
  { id: "2", title: "Banana", price: 10, img: "img/banana.svg", description: "The banana is an edible fruit, botanically a berry, produced by several kinds" },
  { id: "3", title: "Girl", price: 200, img: "img/girl.svg", }
]
```

The sidebar features the "Ask Copilot" icon and the text "Ask Copilot". Below this, there is a message from the Responsible AI Service stating: "Sorry, the response was filtered by the Responsible AI Service. Please rephrase your prompt and try again. [Learn more](#)".

The Copilot interface includes the following controls and information:

- Input field: "Ask Copilot" with placeholder text "or type # to attach context".
- Buttons: "Close", "View in Chat", and a dropdown menu.
- Message history:
  - Line 21: "description:"
  - Line 22: A red squiggle under the character "d" in "description:".
  - Line 23: "]"
- Help text: "/help What can you do?"
- Bottom status bar: "Ask Copilot" and "JS index.product-list.js:21-22 Current file".

The bottom of the screen shows the status bar with various icons and text, including "In 22 Col 1 /22 selected", "Spaces: 4", "LITE-R", "IF", "JavaScript", "A", "Layout: US", and "CodeSpaces: nightmarish casket".

The screenshot shows a code editor interface with the following details:

- Explorer Bar (Left):** Shows a tree view of files and folders under "1283-JS-DOM [CODESPACES: NIGHTMARISH CASKET]".
  - css
  - img
  - js
    - index.header-clock.js (U)
    - index.header-text.js (U)
    - index.js (U)
    - index.product-carousel.js (1, U)** (highlighted in green)
    - index.product-list.js (U)
  - cart.html
  - index.footer.partial.html
  - index.header.partial.html (M)
  - index.html (M)
  - index.product-carousel.partial.... (U)
  - index.products.partial.html (M)
- Top Bar:** Shows tabs for various files: index.html (M), index.product-carousel.partial.html (U), index.header.partial.html (M), index.js (U), index.product-carousel.js (1, U) (highlighted in red), and index.product-list.js (U).
- Code Editor (Main Area):** Displays the content of the selected file, "index.product-carousel.js".

```
js > JS index.product-carousel.js > [?] slides
1 const slides = [
2   '<div></div>',
3   '<div></div>',
4   '<div><img>
```
- Bottom Right Corner:** A tooltip from GitHub Copilot indicates a reference to public code.

⚠️ We found a reference to public code in a recent suggestion.

Source: GitHub Copilot



facebook / react

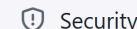
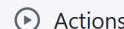


Type / to search



Issues 692

Pull requests 161



# [compiler][ez] tsconfig: treat all snap fixtures as modules #31350

mofeiZ wants to merge 1 commit into [main](#) from [pr31350](#)

1



1



17



1

+3 -1



mofeiZ commented 19 hours ago

Contributor

...

Reviewers

No reviews

Still in progress? Learn about draft PRs



QoI improvement. Currently, typescript lints treat test fixtures without an export as a 'global script' (see [docs](#)). This gives confusing lints for duplicate declarations (in the global scope)



✓ 2b8c121



vercel bot commented 19 hours ago • edited

...

Assignees

No one assigned

Labels

CLA Signed

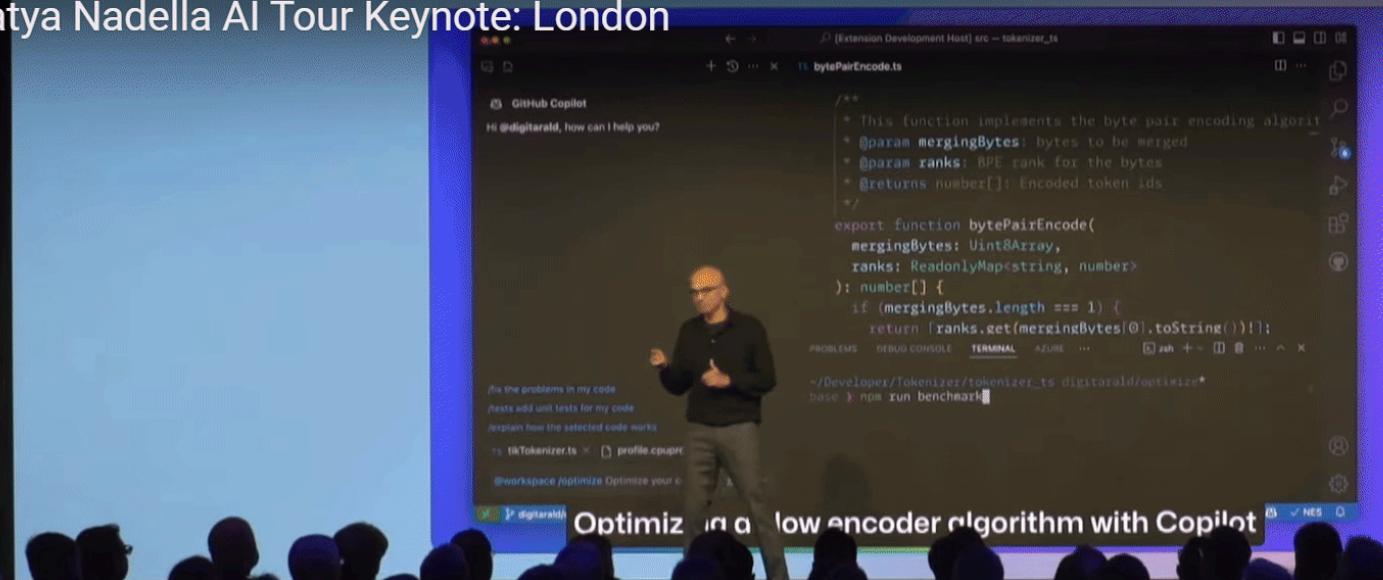
React Core Team

Projects

None yet

Name	Status	Preview	Comments	Updated (UTC)
react-compiler-playground	Ready (Inspect)	<a href="#">Visit Preview</a>	<a href="#">Add feedback</a>	Oct 24, 2024 6:22pm

Satya Nadella AI Tour Keynote: London



you can sort of use AI,





# The AI Code Editor

Built to make you extraordinarily productive,  
Cursor is the best way to code with AI.

Download for Free

Watch Demo  
1 Minute

mod.rs M pingora-core/src/listeners/mod.rs {} impl TransportStack

```
impl TransportStackBuilder {
    pub fn build(&mut self, upgrade_listeners: ListenFds) -> TransportStack {
        ...
    }
}

impl TransportStack {
    pub fn as_str(&self) -> &str {
        self.l4.as_str()
    }

    pub async fn listen(&mut self) -> Result<()> {
        self.l4.listens().take().await
        self.l4.listen(Some(self.upgrade_listeners.clone())).await
    }

    pub async fn accept(&mut self) -> Result<UninitializedStream> {
        let stream = self.l4.accept().await;
        Ok(UninitializedStream {
            l4: stream,
            tls: self.tls.clone(),
        })
    }

    pub fn cleanup(&mut self) {
        // placeholder
        self.tls = None;
        self.l4.cleanup();
    }
}
```

Can you make it easier to switch certificates in the transport listeners

mod.rs Current file

To make it easier to switch certificates in the transport listeners, you can modify the `TransportStackBuilder` and `Listeners` structures to allow dynamic updating of TLS settings. Here's how you can implement this:

Step 1: Add a Method to Update TLS Settings in `TransportStackBuilder`

This method will allow updating the TLS settings of a `TransportStackBuilder` instance.

mod.rs Diff

```
impl TransportStackBuilder {
    ...
    pub fn update_tls_settings(&mut self, new_tls_settings: TlsSettings) {
        self.tls = Some(new_tls_settings);
    }
}
```

Step 2: Add a Method to Update TLS Settings in `Listeners`

This method will allow updating the TLS settings for all `TransportStackBuilder` instances in `Listeners` that match a specific address.

mod.rs Diff

```
impl Listeners {
    ...
    pub fn update_tls_settings_for_address(&mut self, addr: &str, new_tls_settings: TlsSettings) {
        for stack in &mut self.stacks {
            if let ServerAddress::Tcp(current_addr, _) = &stack.l4 {
                if current_addr == addr {
                    stack.update_tls_settings(new_tls_settings.clone());
                }
            }
        }
    }
}
```

See Cursor In Action

1 Minute

# Cursor IDE Benefits

- Built-in AI assistant powered by GPT-4
- AI-powered code generation and refactoring
- Intelligent code completion
- Code explanation and documentation
- Fast search and codebase navigation
- Integration with popular version control systems
- Support for multiple programming languages
- Free for personal use

File Edit Selection View Go Run Terminal Help

ai-augmented-sdev

slides.md

Markdown Language Features # Що з етапом імплементації? ## layout: center

```
346 # Що з етапом імплементації?  
347 --- #36  
348 layout: image  
349 image: /ms-ai-improves-ai.gif  
350 backgroundSize: contain  
351 title: MS Copilot Optimization  
352 ---  
353 --- #37  
354 layout: image  
355 image: /cursor-ide.png  
356 backgroundSize: contain  
357 title: CURSOR IDE  
358 ---  
359 --- #38  
360 layout: center  
361 ---  
362 # Агенти  
363 --- #39  
364 layout: image  
365 image: /matrix-agents.png  
366 backgroundSize: contain  
367 title: Matrix Agents  
368 ---  
369 --- #40  
370 layout: image  
371 image: /software-agents.png  
372 backgroundSize: contain  
373 title: Software Agents with Rivet  
374 ---  
375 --- #41  
376 layout: center  
377 ---  
378 # SWE Bench  
379 --- #42  
380 layout: image  
381 image: /swe-bench.png  
382 backgroundSize: contain  
383 title: SWE Bench
```

AI-AUGMENTED-SDEV

- .github\workflows
- deploy.yml
- components
- node\_modules
- pages
  - imported-slides.md
- public
- snippets
- .gitignore
- .npmrc
- .prettierrc
- netlify.toml
- package-lock.json
- package.json
- README.md
- slides-example.md
- slides.md
- vercel.json

OUTLINE

TIMELINE slides.md

- Cursor koldovsky now
- File Saved
- Undo / Redo
- File Saved
- File Saved 2 mins
- File Saved 7 mins
- add copilot optimizes copilot koldovsky 1 hr
- File Saved
- add slides koldovsky
- File Saved
- File Saved
- File Saved

softserve | academy

main 0111 0 0 0 Live Share

Ln 392, Col 4 Spaces: 2 UTF-8 LF {} Markdown Cursor Tab ✓ Prettier

How to Make it Beautiful?



# What can I help you ship?

Need more messages? Get higher limits with Premium.

Upgrade Plan x

Ask vO a question...

I



+ Project



Generate a sticky header ↗

How can I schedule cron jobs? ↗

Calculate the factorial of a number ↗

# Code Reviews?

CodeRabbit

Merged

Pre main #6

mykhailoiskiv1 merged 2 commits into [main](#) from [pre-main](#)  last week



coderabbitai (bot) commented last week • edited

...

## Walkthrough

The changes involve modifications to several components in the application. The `Cover` component has been updated for text rendering and import management. A new `Footer` component has been introduced along with several child components (`FooterContactInfo`, `FooterNav`, `FooterDescription`, `FooterSocials`, `FooterCopyright`). The `Header` component has minor formatting changes, while the layout file has been updated to include the new `Footer`. Additionally, the Tailwind CSS configuration has been updated to include new color properties for styling.

## Changes

File Path	Change Summary
app/components/Home/Cover.tsx	Commented out font import, changed label text to HTML entity, minor formatting adjustments to <code>&lt;hr&gt;</code> .
app/components/layout/footer/Footer.tsx	Introduced <code>Footer</code> component with child components for contact info, navigation, description, socials, and copyright.
app/components/layout/footer/FooterContactInfo.tsx	Added <code>FooterContactInfo</code> component for rendering contact information.
app/components/layout/footer/FooterCopyright.tsx	Added <code>FooterCopyright</code> component for displaying copyright notice.
app/components/layout/footer/FooterDescription.tsx	Added <code>FooterDescription</code> component for a brief description of the flower studio.

You're receiving notifications because you're watching this repository.

1 participant

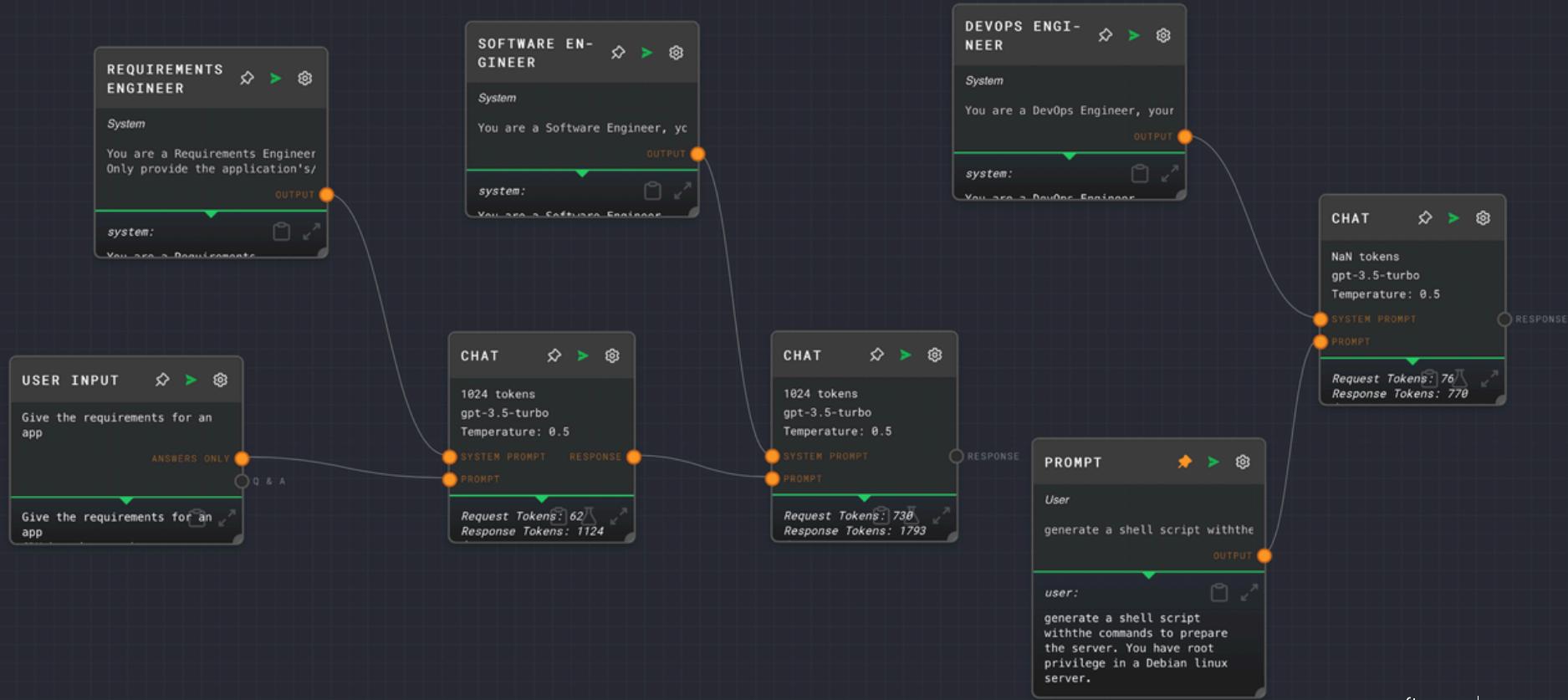


 Lock conversation



# Agents







## Swarm (experimental, educational)

An educational framework exploring ergonomic, lightweight multi-agent orchestration.

### Warning

Swarm is currently an experimental sample framework intended to explore ergonomic interfaces for multi-agent systems. It is not intended to be used in production, and therefore has no official support. (This also means we will not be reviewing PRs or issues!)

The primary goal of Swarm is to showcase the handoff & routines patterns explored in the [Orchestrating Agents: Handoffs & Routines](#) cookbook. It is not meant as a standalone library, and is primarily for educational purposes.

## Install

# Introducing computer use, a new Claude 3.5 Sonnet, and Claude 3.5 Haiku

Oct 22, 2024 • 5 min read



Today, we're announcing an upgraded **Claude 3.5 Sonnet**, and a new model, **Claude 3.5 Haiku**. The upgraded Claude 3.5 Sonnet delivers across-the-board improvements over its predecessor, with particularly significant gains in coding—an area where it already led the field. Claude 3.5 Haiku matches the performance of Claude 3 Opus, our prior largest model, on many evaluations for the same cost and similar speed to the previous generation of Haiku.

We're also introducing a groundbreaking new capability in public beta: **computer use**. Available [today on the API](#), developers can direct Claude to use computers the way people do—by looking at a screen, moving a cursor, clicking buttons, and typing text. Claude 3.5 Sonnet is the first frontier AI model to offer computer use in public beta. At this stage, it is still [experimental](#)—at times cumbersome and error-prone. We're releasing computer use early for feedback from developers, and expect the capability to improve rapidly over time.

SWE Bench

# SWE-bench



Can Language Models Resolve Real-World GitHub Issues?

ICLR 2024

Carlos E. Jimenez\*, John Yang\*,  
Alexander Wettig, Shunyu Yao, Kexin Pei,  
Ofir Press, Karthik Narasimhan

Paper

Code

Submit

Analysis



SWE-bench Multimodal

## News

- ⚠️ [10/2024] Introducing **SWE-bench Multimodal**! Can AI systems "see" bugs and fix them? 🤖 [Link]
- ⚠️ [08/2024] SWE-bench x OpenAI = **SWE-bench Verified**, a human-validated subset of 500 problems reviewed by software engineers! [Report]
- ⚠️ [06/2024] We've **Docker**-ized SWE-bench for easier, containerized, reproducible evaluation. [Report]
- ⚠️ [03/2024] Check out our latest work, **SWE-agent**, which achieves a 12.47% resolve rate on SWE-bench! [Link]
- ⚠️ [03/2024] We've released **SWE-bench Lite**! Running all of SWE-bench can take time. This subset makes it easier! [Report]

## Leaderboard

Lite

Verified

Full

### Model

👉 Tools + Claude 3.5 Sonnet (2024-10-22)

Model	% Resolved	Org	Date	Logs	Trajs	Site
👉 Tools + Claude 3.5 Sonnet (2024-10-22)	49.00	AI	2024-10-22	✓	✓	🔗
👉 Solver (2024-09-12)	45.40	S	2024-09-24	✓	✓	🔗
👉 Gru(2024-08-24)	45.20	G	2024-08-24	✓	✓	🔗
Solver (2024-09-12)	43.60	S	2024-09-20	✓	✓	🔗
Tools + Claude 3.5 Haiku	40.60	AI	2024-10-22	✓	✓	🔗
Honeycomb	40.60	HC	2024-08-20	✓	✓	🔗
👉 Composio SWEkit + Claude 3.5 Sonnet (2024-10-16)	40.60	⚡	2024-10-16	✓	✓	🔗
Amazon Q Developer Agent (v20240719-dev)	38.80	aws	2024-07-21	✓	✓	🔗
AutoCodeRover (v20240620) + GPT 4o (2024-05-13)	38.40	ACR	2024-06-28	✓	-	🔗
Factory Code Droid	37.00	FD	2024-06-17	✓	-	🔗
👉 SWE-agent + Claude 3.5 Sonnet	33.60	SA	2024-06-20	✓	✓	-
nFactorial (2024-10-07)	31.60	nf	2024-10-07	✓	✓	🔗

# What do you want to build?

Prompt, run, edit, and deploy full-stack web apps.

How can Bolt help you today?



[Start a blog with Astro](#) [Build a mobile app with NativeScript](#) [Create a docs site with Vitepress](#)

[Scaffold UI with shadcn](#)

[Draft a presentation with Sliderv](#)

[Code a video with Remotion](#)

or start a blank app with your favorite stack



# Black Swan Effect

# The future of game development... has no game engine?



**TLDR** Google and Tel Aviv University have unveiled GameNGen, the world's first neural network-based game engine that simulates gameplay in real time without any written code. This technology, although still in its infancy and barely playable, could revolutionize game development by using AI to generate game environments and [Show more](#)

#tech-news

#ai

#game-development

#neural-networks

#reinforcement-learning

# Assisted vs Augmented

## AI Assisted

- AI as a support tool
- Developer controls the process
- AI generates code on request
- Limited AI autonomy
- Focus on individual tasks

## AI Augmented

- AI as an active partner
- Human-AI collaboration
- AI proactively suggests solutions
- High AI autonomy
- Comprehensive development approach

# Implementation Insights

# API Selection & Integration

- Evaluate available LLM APIs and their capabilities
- Compare pricing models and usage quotas
- Assess API documentation and support
- Consider API reliability and performance metrics

# Prompt Engineering

- Design effective prompts for desired outcomes
- Implement prompt templates and validation
- Handle API responses and error cases
- Optimize token usage and costs
- Test prompt consistency and reliability

# Cost Management

- API usage monitoring and optimization
- Request caching strategies
- Rate limiting implementation
- Fallback handling
- Budget tracking and alerts

# Security Best Practices

- API key management and rotation
- Input sanitization and validation
- Output content filtering
- User data protection
- Request/response logging

# Risk Mitigation

- API availability monitoring
- Response quality assessment
- Content moderation strategy
- Fallback mechanisms
- Regular performance reviews

# Implementation Guidelines

- Start with proof of concept
- Implement proper error handling
- Cache frequently used responses
- Monitor API usage and costs
- Maintain prompt engineering documentation

Thank you for your attention!