May 2, 2025

Advancements and Potential Futures Capabilities, Models, and Agents

Dubberly Design Office

Al has attracted billions in investments and generated lots of hype. "In the short term, we overestimate the effects of a new technology but underestimate them in the long term." – Roy Amara



PART ONE

Al's broader accomplishments

Dubberly Design Office 🔸 Al Advancements 🔸 May 2, 2025



Even just recently, AI has made extraordinary accomplishments:

- Predicting protein folding (with DeepFold)
- Detecting tumors from radiograms in cases of lung cancer (with EIRL X-Ray)
- Beating the best humans at Chess and Go (Deep Blue and AlphaGo)
- Creating realistic video (with Sora)
- Decoding language from brain activity (from Meta)
- Identifying disease pathways and treatments (with Genentech Lab in a Loop)

Predicting 3D structure of 200 million proteins (with AlphaFold)

AlphaFold, developed by DeepMind, revolutionized biology by predicting 3D protein structures from amino acid sequences with remarkable accuracy—solving a challenge that had stumped scientists for decades.

Its impact is profound: what once took years in the lab can now be done in minutes, accelerating research in drug discovery, disease understanding, and molecular biology.

In recognition of this breakthrough, Demis Hassabis and John Jumper were awarded the 2024 Nobel Prize in Chemistry. The Nobel committee honored their work as a "revolution in the life sciences," noting how AlphaFold has provided structure predictions for over 200 million proteins nearly all known to science.





Detecting tumors from radiograms in cases of lung cancer (with EIRL X-Ray)

EIRL, developed by LPIXEL Inc., is an AI system that supports radiologists by detecting lung nodules in chest X-rays with high accuracy. It improves early lung cancer detection by increasing diagnostic sensitivity—up to 13% for non-specialists—without raising false positives. It represents a major step forward in AI-assisted healthcare diagnostics, particularly in radiology.



Beating the best humans at Chess and Go (Deep Blue and AlphaGo)

In 2016, Google DeepMind's AlphaGo made history by defeating world champion Lee Sedol at the complex board game Go—a feat long thought to be decades away due to Go's immense strategic depth. AlphaGo's victory marked a milestone in artificial intelligence, showcasing the power of deep learning and reinforcement learning. It stunned experts with creative, non-human moves and demonstrated that AI could surpass top human intuition in highly abstract domains. This achievement paved the way for advanced AI systems in science, healthcare, and beyond.





Creating realistic video from text prompts (with Sora)

In 2024, OpenAl unveiled Sora, a cutting-edge Al model capable of generating high-quality, realistic video from text prompts. Sora represents a major leap in multimodal Al, combining spatial and temporal coherence to produce videos with dynamic motion, complex scenes, and consistent characters. Unlike earlier models, Sora can maintain visual consistency across extended sequences, opening new possibilities in filmmaking, simulation, and education.











Decoding language from brain activity (from Meta)

In 2025, Meta AI achieved a significant milestone in non-invasive brain-computer interfaces. The system interprets natural language production by analyzing neural patterns associated with typing.

Their AI model decodes brain activity captured via magnetoencephalography (MEG) and electroencephalography (EEG) to reconstruct typed sentences, reaching up to 80% character-level accuracy.





Using AI to decode language from the brain and advance our understanding of human communication https://ai.meta.com/blog/brain-ai-research-human-communication/



Identifying disease pathways and treatments (with Genentech Lab in a Loop)

In a 2023 TED talk, Aviv Regev — executive VP and head of research and development at Genentech — describes a "lab in a loop", a framework for using Al in disease research and drug discovery.

She describes how AI / ML models were trained to identify disease pathways from gene sequences, and how other models could be trained to identify drugs that target those pathways as well as develop personalized treatments for individual patients.



Aviv Regev, "Can Al help develop new medicines?" (2023) https://www.ted.com/talks/aviv_regev_can_ai_help_develop_new_medicines



PART TWO

Al's practical capabilities today

Al has also made accomplishments in more practical areas

- Generating PowerPoint presentations (with ChatGPT)
- Summarize and write documents (with GPT, etc.)
- Writing detailed reports, citing sources (with ChatGPT Deep Research)
- Natural-language search and information retrieval (with ChatGPT, Google)
- Generating and analyzing spreadsheets, generating graphs from data (with ChatGPT)
- Drafting wireframes from a text prompt (with Figma First Draft)
- Helping code prototypes (with Cursor, Co-pilot, Codeium, Windsurf, Vercel)
- Generating illustrations for user scenario storyboards (with GPT-40 or Dall-E)

Generating PowerPoint presentations (with ChatGPT)



Summarize long documents, articles, or emails (with GPT, etc.) Give it a PDF file or URL, or just paste text directly into the chat.



🕕 🕜 ChatGPT 4.5 🗸

Summarize this article for me: https://openai.com/index/introducing-gpt-4-5/

OpenAl introduced GPT-4.5 on February 27, 2025, as a research preview of its most advanced language model to date. This release emphasizes enhanced unsupervised learning, leading to improved pattern recognition, broader knowledge, and more natural conversational abilities. GPT-4.5 is designed to assist with tasks like writing, programming, and problem-solving, offering a more intuitive user experience with reduced hallucinations. The Verge +12 OpenAl Help Center +2 OpenAl +3

Trained on Microsoft Azure AI supercomputers, GPT-4.5 showcases advancements in architecture and optimization, resulting in a model that is more reliable across various topics. While it lacks chain-of-thought reasoning found in models like OpenAI's o-series, GPT-4.5 compensates with its enhanced creativity and emotional intelligence.

OpenAl +1 OpenAl Help Center -1

Notably, GPT-4.5 has been recognized for its ability to pass an authentic Turing test, convincingly mimicking human conversation in a significant majority of interactions. This achievement underscores its advanced conversational capabilities and the strides made in Al-human interaction. Live Science

Currently, GPT-4.5 is accessible to ChatGPT Pro users and developers via the API, with plans to expand availability to other user tiers. Wikipedia





1 Share

Drafting documents like memos, emails, or speeches (with GPT, etc.) Provide outlines or notes and let ChatGPT or Claude generate a coherent document.

ChatGPT 40 ~		1 Share	:
	I'm an executive giving a talk on leadership at a conference coming up. Write me a speech using the outline I've drafted below:		
	 Opening – Establish Credibility and Connection What personal story or pivotal leadership moment can I share? What is my core belief about leadership? What challenge or opportunity does this audience face — and how is leadership the answer? 		
	 2. Defining Leadership - How do I define leadership (in contrast to management, for example)? - What principles or values guide my leadership? - What does effective leadership look like in today's environment 		
	(e.g., hybrid work, Al, generational shifts)?		
	 3. Lessons from Experience What leadership decisions am I most proud of, and why? What did I get wrong — and what did I learn? How have I grown as a leader over time? What moments tested me the most? 		
	 4. Frameworks and Tools - Do I use any frameworks to guide my team (e.g., servant leadership, radical candor, OKRs)? - How do I develop other leaders? - How do I approach change, uncertainty, or crisis as a leader? 		
	5. People and Culture - How do I build trust? - How do I shape organizational culture? - How do I make sure people feel seen and valued?		
	6. Leadership in Practice - What does leadership look like in day-to-day behavior? - How do I handle disagreement, failure, or complexity? - How do I align people around a shared vision or goal?		
	7. Looking Ahead - What leadership traits will matter most in the next 5–10 years? - How do we prepare future leaders? - What are my hopes for the next generation of leadership?		
Ask anything	***		
+ 🌐 Search 🗬	Deep research & Create image & 🖓		
	ChatGPT can make mistakes. Check important info.		



Writing detailed reports, citing sources (with ChatGPT Deep Research) For example, asking it to research competitors and collect a detailed competitive analysis.

C	ChatGPT 4o ~	▲ Share	1
	Map the EMR/CRM/RCM software market for me. Compare different offerings to Kipu Health, Focus primarily on behavioral healthcare (BHC), but include general products if they can also be applied in BHC contexts.		
	To create a thorough comparison of EMR/CRM/RCM software with a focus on behavioral healthcare and reference to Kipu Health, could you clarify the following:		
	1. Are you interested in U.Sbased providers only, or should this include international vendors?		
	2. Should the comparison focus on specific features (e.g., patient engagement, billing, outcome tracking)?		
	3. Are you looking for enterprise-grade platforms, SMB-focused options, or a mix?		
	4. Would you like pricing information included where available?		
	Once I have this, I can pull together a detailed market map and comparison.		
	 US-based only. All of the above, compare any major features you can think of and highlight any unique feature offerings. mix of both no, I don't need pricing info 		
	Got it. I'll map out the U.Sbased EMR, CRM, and RCM software market with a focus on behavioral healthcare. I'll include both behavioral health-specific platforms and general solutions that support BHC workflows, comparing them to Kipu Health. I'll evaluate a mix of enterprise and SMB-level offerings and		
	I'll let you know once the analysis is ready		

(Note: All platforms discussed are US-based. Citations are provided for feature claims. See the comparison table at the end for a summary.)

features:

ChatGPT 40

Comparison of U.S. Behavioral Health EMR/CRM/RCM Platforms

Introduction

Behavioral healthcare (BHC) organizations - from small therapy practices to large addiction treatment centers – have unique needs for electronic records and practice management. They require systems that handle clinical documentation, patient engagement, scheduling, billing and revenue cycle management (RCM), compliance (e.g. HIPAA, 42 CFR Part 2), interoperability with other health systems, robust analytics/reporting, and convenient mobile access for clinicians. In recent years, a variety of software platforms have emerged (or evolved from general EHRs) to serve these needs. This report maps the U.S. market of EMR/EHR, CRM, and RCM solutions focused on behavioral health, comparing major features across leading platforms. Both behavioral health-specific solutions and general-purpose EHR/RCM systems (adapted for behavioral health contexts) are included. Each platform's capabilities in areas like clinical workflows, patient engagement portals, scheduling, billing/RCM, compliance tools, interoperability, analytics, and mobility are compared - with standout features highlighted. A special focus is given to how these solutions stack up against Kipu Health, a prominent behavioral health and addiction treatment platform.

Behavioral Health-Specific Software Platforms

This category includes systems purpose-built for mental health and substance use treatment settings. They typically come pre-configured with behavioral health templates (e.g. therapy notes, treatment plans, outcome measures) and compliance features tailored to BHC regulations. Many offer all-in-one solutions integrating EMR/EHR functionality with practice management, billing/RCM, and sometimes CRM for client relationship management. Below we detail key platforms, including those serving large enterprise behavioral health providers and those for small to mid-size practices, and how each compares to Kipu in

• Kipu Health: Kipu is a cloud-based platform specifically designed for addiction treatment and behavioral health organizations. It is an end-to-end solution covering the entire patient lifecycle from admissions and CRM through clinical documentation to billing/RCM selectitub.com . Kipu's CRM module is a standout - it manages leads, referrals and marketing campaigns (tracking inquiry sources and ROI on marketing spend) which helps drive admissions growth selecthub.com . On the clinical side, Kipu supports comprehensive electronic medical record (EMR) functions: customizable intake assessments, treatment plans, progress notes, outcome measures tracking, and discharge planning selection.com . It includes features like e-prescribing, medication management, lab integrations (for toxicology results, etc.), and even built-in telehealth with screen-sharing selecthub.com . Kipu's billing/RCM tools handle coding (ICD-10, DSM-5), claim scrubbing, electronic claim submission/tracking, and payment processing selecthub.com . Compliance is also a focus – it offers a governance, risk & compliance (GRC) module to assist with regulatory compliance and auditing. Analytics are built-in via Kipu Insights, delivering dashboards on patient census, outcomes, financial metrics, and more velective.com . Mobile access: Kipu is web-based, accessible on tablets for clinicians; its CRM even allows on-the-go lead updates kipuhealth.com . Unique strengths: Kipu's specialization means it natively supports workflows like detox assessments, rehab admissions, and alumni aftercare tracking (its CRM helps manage alumni contacts and automate outreach for relapse prevention) kipulikakik.com . Compared to general EHRs, Kipu provides a more targeted feature set for behavioral health, including marketing/admissions tools that most medical EHRs lack. (One noted

limitation has been integration with external systams - some users desired smoother interoperability The sale of the state of the Prime Ask anything

\$ (III)

+
 Gearch
 Peep research
 ···

ChatGPT can make mistakes. Check important info.

ChatGPT 40 ~

1 Share

1 Share

cover behavioral health along with physical health services, but doesn't inherently include marketing CRM or some of the addiction-specific tools out-of-the-box.

 Other general systems: Meditech Expanse (used in some psychiatric hospitals), eClinicalWorks (used by some behavioral health clinicians especially in small clinics, offering progress note templates and telehealth), DrChrono, eMDs/Aprima, and Sevocity are examples of general ambulatory EHRs that mental health providers might use. They generally all provide core EMR features, patient scheduling, e-prescribing, portals, and billing. Usually, what distinguishes them from behavioral-health-specific platforms is the need for the user to configure behavioral health content and the lack of specialized tools like built-in outcome measures or rehab-specific workflows. For instance, none of these would include an alumni program tracker or referral CRM natively. However, they often integrate with third-party apps; e.g., a practice on eClinicalWorks might use a separate outcomes measurement app or a CRM tool and sync data.

In summary, general-purpose EHR/RCM platforms can absolutely be used in behavioral health and often are in settings that are part of larger healthcare organizations or where the provider's needs are more standard (document, prescribe, bill). But when directly compared to platforms like Kipu or others purpose-built for behavioral health, the general systems might require more customization and may lack certain niche features (while excelling in interoperability and broad support). Many behavioral health organizations face the choice of integration vs specialization - integrate into a big general EHR for data continuity, or use a specialized BH software for tailored features. Increasingly, interoperability improvements (HL7/FHIR APIs) are allowing specialized BH EHRs to connect with general systems, aiming to give providers the "best of both."

Feature Comparison Across Major Platforms

All the above platforms cover the fundamental needs of electronic health records and practice management, but they differ in depth or approach to specific features. The table below provides a highlevel comparison of how key features are supported across several notable behavioral health platforms, including both specialized systems (like Kipu and its direct competitors) and a few general EHRs. This highlights strengths and unique offerings of each:

Platform	Primary Focus / Users	Clinical Documentation & Treatment Planning
Kipu Health	Behavioral health & addiction centers (SMB to enterprise)	Full BH EMR: assessments, progress notes, personalized treatment plans, outcome tracking selections . Includes detox/withdrawal management tools.
Alleva	Behavioral health & SUD treatment (SMB to mid-size)	Comprehensive BH clinical notes and treatment planning. Customizable forms and automated compliance prompts.
Lightning Step	Behavioral health & addiction centers (Enterprise-capable)	Full EMR with BH templates; supports configurable workflows and documentation across levels of care.
BestNotes	Behavioral health & addiction (SMB clinics, residential, outpatient)	Full BH documentation suite with treatment plans, group notes. Templates auto-populate to reduce double entry somweredvice.com . Outcome questionnaires can be sent to patients and auto-recorded
Qualifacts CareLogic	Behavioral health agencies, community mental health	 Extensive BH clinical functionality: assessments, progress notes, treatment plans, eMAR, and support
+	윢 Deep research ···	ę 🐠
	ChatGPT can make m	istakes. Check important info.

Natural-language search and information retrieval (with ChatGPT, Google)







Sources APTON

0000000

Yesterday

\$



NVIDIA's market capitalization has surged, recently surpassin

MORNINGSIDE HEIGHTS HARLEM 8 W 112th St 107th St EAST HARLE W92. EIn **Absolute Bagels** Open now until 7pm · Bagels · 9.1 (1107 A staple since the early 1990s, Absolute E flavored bagels and schmears, including Directions Website Call Sources





Generating, analyzing, and automating spreadsheets (with ChatGPT)

1	🕑 ChatG	PT 40 >										£ G											
Weath	er and Emission	s Data									₹ ⁴ ×	generate a big spreadsheet of a hunch of fake made-up data on											
#	Date	City	Temperature_C	Humidity_%	CO2_ppm	NO2_ppb	PM2.5_ug/m3	Wind_Speed_m/s	Precipitation_mm	Solar_Radiation_W/m2		weather statistics, emissions, and some other related things											
1	2022-01-01	Springfield	-1.49749156543524	58.87019516168519	393.31332537100604	20.6550172738447	14.679580823674106	6.806043628322675	0.0	383.0980759272759													
2	2022-01-02	Springfield	36.610083774395335	30.08176918627654	382.2410088735507	31.094711264695768	11.905520902664724	0.5278057024277322	0.0	206.78270793833786		Weather and Emissions											
3	2022-01-03	Springfield	16.49180762563287	31.099250314707447	364.3547181474206	36.016252366779504	8.76917899255109	6.35752407974559	1.5864048292685076	350.5015093010943		Here's a sample of a large spreadsheet with fake weather and emissions data across multiple cities an											
4	2022-01-04	Springfield	7.419066382417027	33.85111078239092	400.7673917301708	20.998718233456643	12.791359307366129	1.7905950056269333	0.6230165641623442	288.007050628745		dates. It includes temperature, humidity, CO ₂ and NO ₂ levels, PM2.5 concentrations, wind speed,											
5	2022-01-05	Springfield	15.659467858359916	87.17462399709052	402.827257118696	21.250528978641505	10.02570636044261	14.22625570940953	0.0	805.4411689020184		precipitation, and solar radiation. Let me know if you want the full dataset exported or visualized!											
6	2022-01-06	Springfield	4.585534611678529	38.87745974543379	393.93300928614633	18.564719719131237	11.899449432660608	5.3232831245345	0.0	903.1064613510782													
7	2022-01-07	Springfield	23.015331099310615	71.24126646569607	384.2245783754098	24.774325304480662	12.466535310897294	11.133357758832148	0.0	122.4041657384017													
8	2022-01-08	Springfield	15.396330188500304	56.13430594041414	414.7000635490909	24.868213017104722	12.00309776353378	6.6678516328290165	2.1214222763061286	734.7304366982811													
9	2022-01-09	Springfield	27.045384114696915	61.42641607837355	394.096104741351	17.662020753384795	14.517492979864414	5.547005521746617	3.3720812286718695	888.823342542246													
10	2022-01-10	Springfield	21.868813281193585	79.93580017027129	380.3161474655677	32.04771746964922	17.458732091053996	13.481097049045442	0.0	626.1522377528229													
11	2022-01-11	Springfield	20.877440328349078	46.48311650587021	406.0240553897764	22.529946185355158	13.451117618589008	9.867077329432389	0.0	446.74557954524164													
12	2022-01-12	Springfield	19.224114255727336	80.47103275869661	392.44412096533574	22.15133494555824	12.825087177043807	0.42151071901563697	0.0	777.4929185689479													
13	2022-01-13	Springfield	13.734895937723431	58.92059814489786	391.54802270882294	29.210242266204787	13.958024803440713	12.537856030843463	0.23320817291721868	313.1765445241007													
14	2022-01-14	Springfield	17.479880676295274	51.60889822972977	428.08444804079653	32.55597595947978	21.19919698405915	10.263872596914506	0.0	953.5426433507203													
15	2022-01-15	Springfield	30.581752408999087	31.767913872720936	402.99981421174334	33.610258059334456	7.383717038347001	9.6238305870267	2.9343062190192537	922.9737938567422													
16	2022-01-16	Springfield	14.121594591998303	82.85283252431105	415.47451973969504	20.72525973453911	13.618526074090802	13.560989582631507	0.0	699.4257246915521													
17	2022-01-17	Springfield	13.303104622496672	45.88244898992351	393.5393215768791	29.580465171503448	8.708424712388007	6.27851828584609	0.6849285401999123	421.4753907371478													
18	2022-01-18	Springfield	44.44712447070329	45.16079167152236	402.03854095495154	33.89572416873499	11.38242986405988	13.272807184328514	0.0	480.2344298182899													
19	2022-01-19	Springfield	19.58969422650621	58.83177916656178	394.2065039904482	20.15225999285111	17.03813630400013	3.9547070723337443	0.0	396.9832097644141													
20	2022-01-20	Springfield	6.552839020236215	72.12376002495805	407.6766424878641	20.91023353193291	18.999266904296732	12.232251828790936	2.2085044749663956	447.83640953205764													
21	2022-01-21	Springfield	15.684104946369573	74.82970007357221	403.48282017848993	31.37623040523139	10.213344830634007	13.887776251843867	0.013027668180812219	845.9708956283104													
22	2022-01-22	Springfield	11.510306932239976	66.30742607886364	394.36831993484213	26.061891797166343	2.9378689410293717	4.955799878081865	0.0	408.3334867703767													
23	2022-01-23	Springfield	1.567356620924432	67.40071001628864	395.0089389823429	25.695038834147933	8.894286367054487	2.2263383032841193	1.4163316932408323	276.04865955384696													
24	2022-01-24	Springfield	4.484564387731233	88.63806681492292	398.14255328017293	18.781330514546724	14.291949729268488	5.049704487667689	0.0	896.4679931541564													
25	2022-01-25	Springfield	9.723959975509317	56.191277341060854	384.3726761821596	26.111585143642362	10.028878032689775	10.929459923265123	1.702377415348721	723.2992517218117													
26	2022-01-26	Springfield	26.668928387300586	76.43292188764349	358.5758957138169	26.703867895247555	18.502765023960862	5.8884853687273795	0.0	916.7085294768438													
27	2022-01-27	Springfield	5.9349783796162505	61.33598047053236	411.58020329192317	27.766964922555154	8.906260533092873	8.312314996547114	0.0	795.4017616943336													
28	2022-01-28	Springfield	11.336246097885553	37.99782989175018	406.15615608581766	22.758334557652287	8.361410290140252	13.007410895989906	0.0	423.0983909245677													
29	2022-01-29	Springfield	18.81821465458534	31.482647491610518	378.6628451224619	19.374684408893334	10.952805982403383	5.698972922526108	0.0	438.0554787764137		↔ "14.121594591998303" cell											
30	2022-01-30	Springfield	20.834536026490913	51.26017997560131	391.11127269937015	21.909884207436853	11.27414727693941	9.929276396392165	0.0	713.5160027482209		Massage ChatCBT											
31	2022-01-31	Springfield	1.0128620456864539	37.03978168725175	410.62431608111496	28.723603108363164	18.316427929700257	4.538000236549615	0.0	220.78114321673655													
32	2022-02-01	Springfield	18.85273344523185	53.159442311157314	385.0760071233534	31.684131170358725	10.018476008977462	4.438845975722206	0.0	631.1820720644929													

Generating graphs from data (with ChatGPT)

Graphs can be exported for you to use them in other documents.













Drafting wireframes from a text prompt (with Figma First Draft)



N	
-	

			Al beta
oonto			
lents			
me			
ienta			
ents			
Ionto			
ma			
nents			
	e options abo	ove are i	ust the
a library, TI			COLD 21 20 C.D.
a library. Tl oon. <u>Learn r</u>	lore		
a library. Tl oon. <u>Learn r</u>	lore		
a library. Tl oon. <u>Learn r</u>	<u>iore</u>		
a library. Tl on. <u>Learn r</u>		<u>~</u> 2	

Helping code prototypes (with Cursor, Co-pilot, Codeium, Windsurf, Vercel)

DMPOSER
rrent File nake it easier to switch certificates in th
dify the code to make certificate switchi Il be to enhance the TlsAccept trait a I in the TlsSettings . Here are the ke
existing code
Trait] t TlsAccept { Called during TLS handshake to se Returns true if certificates were of fn certificate_callback(&self,
false
ic fn on_cert_error(&self, _error:
// Default implementation does no
default no—op implementation tha (Default)] s Current File × p(೫Y),↑to select
s p

Cursor Al



Vercel v0



Generating illustrations for user scenario storyboards (with GPT-40 or Dall-E)



Tasks

Tasks

Right on Time: As the Current Time reaches the Alarm Time, the alarm triggers perfectly, waking Emily up calm and ready for her flight.



Tasks

Ending the Day with Ease: After a long day, Emily disables the alarm for the next morning, grateful for an app that keeps her on track without added stress.







Tasks

Automatic Time Updates: The app seamlessly adjusts the Current Time to the local time zone, eliminating Emily's worry about setting her alarm incorrectly.







Table of Al products based on inputs and outputse.g., text-to-text, text-to-image, image-to-image

		Output								
1. 21	-14	Text	Image	Music	Video	Speech	Code	Object	Proteins	Drugs
Input	Text	Text-to-text GPT-3.5 / 4 (OpenAl) LLaMA (Meta) LaMDA (Google)	Text-to-image Dall-E (OpenAl) Midjourney Stable Diffusion Runway ML	Text-to-music	Text-to-video Runway ML	Text-to-speech TTS (OpenAl)	Text-to-code GPT 3.5 / 4 (OpenAl)	Text-to-object Excalidraw ~ GPT → Code → Object	Text-to-proteins	Text-to-drugs
	Image	Image-to-text (image recognition) GPT-4 Turbo w/ vision (OpenAI)	Image-to-image Dall-E (OpenAl) Runway ML NVIDIA Canvas	Images-to-music	Image-to-video Runway ML <i>Deepfakes</i>	Images-to-speech	Images-to-code	<mark>Images-to-object</mark> <u>MakeltReal</u> (tldraw)	Images-to-proteins	Images-to-drugs
	Music	Music-to-text Music tagging Al	Music-to-image	Music-to-music	Music-to-video	Music-to-speech	Music-to-code	Music-to-object	Music-to-proteins	Music-to-drugs
	Video	Video-to-text	Video-to-image	Video-to-music	Video-to-video Deepfakes	Video-to-speech	Video-to-code	Video-to-object	Video-to-proteins	Video-to-drugs
	Speech	Speech-to-text (voice recognition) Whisper (OpenAI)	Speech-to-image	Speech-to-music	Speech-to-video	Speech-to-speech ChatGPT Voice Chat Eleven Labs Voice Agent Deepfakes	Speech-to-code	Speech-to-object	Speech-to-proteins	Speech-to-drugs
	Code	Code-to-text	Code-to-image	Code-to-music	Code-to-video	Code-to-speech	Code-to-code Copilot (OpenAl, GitHub)	Code-to-object	Code-to-proteins	Code-to-drugs
	Object	Object-to-text	Object-to-image	Object-to-music	Object-to-video	Object-to-speech	Object-to-code	Object-to-object <u>MakeltReal</u> (tidraw)	Object-to-proteins	Object-to-drugs
	Proteins	Proteins-to-text	Proteins-to-image	Proteins-to-music	Proteins-to-video	Proteins-to-speech	Proteins-to-code	Proteins-to-object	Protein-to-protein	Protein-to-drugs
	Drugs	Drugs-to-text	Drugs-to-image	Drugs-to-music	Drugs-to-video	Drugs-to-speech	Drugs-to-code	Drugs-to-object	Drugs-to-proteins	Drug-to-drug Pharma.Al



PART THREE

Deploying agents in organizations

A 'sparring partner' agent could help train new sales representatives.

As a role playing exercise, the trainee pretends to sell to the agent, who tests the trainee by asking questions about the product as though it is a prospective buyer.







Agent Prospective Buyer

The agent uses product information and a scoring rubric to score the trainee's performance.



knowledge fed to





Product Information Rubric

Similar 'sparring partner' agents could be useful in other areas an agent could stress-test a new product or feature by ensuring the product team has addressed any necessary documentation or considerations before getting started.



knowledge fed to





Product Development Requirements Rubric

 \blacksquare

Alternatively, the agent might evaluate the usability of a product by playing the role of a user.





Č Š

knowledge fed to

2



Agent User

Usability considerations Rubric

In research with real users, the agent could assist demos by recording data.



or Designer

Another agent might augment customer support, increasing support availability by assisting customers with questions. The agent would also be able to connect the customer with a human customer support rep.



knowledge fed to



Product **Documentation** / Knowledge Base

When a customer support representative is working with a customer, an agent could provide them with live assistance and surface relevant documentation.



PART FOUR

Models, infrastructure, and players.

Several major players have emerged with competitive Large Language Models (LLMs).



Standard LLMs

GPT Series

e.g.,

GPT-4

GPT-4o

GPT-4.5



Standard LLMs Gemini Series e.g., Gemini 2.5 Pro Gemini 2.5 Flash



Standard LLMs LLaMA Series e.g., LLaMA 2 LLaMA 3

Reasoning models

o Series

e.g.,

01

о3







Standard LLMs Claude e.g., Claude 3.5 Haiku Claude 3.5 Sonnet Claude 3.7 Sonnet

Reasoning models	Standard LLMs
e.g.,	Grok
DeepSeek-V3	e.g.,
DeepSeek R1	Grok 2
-	Grok 3

1. OpenAl

With significant investment from Microsoft, OpenAI spearheaded the modern LLM era through the launch of ChatGPT, powered first by GPT-3.5 and later by GPT-4. Their models set a new standard for conversational AI and sparked widespread interest.

2. Google DeepMind

Though Google had long been a pioneer in NLP research, its entrance into the consumer-facing LLM space came later. The release of PaLM, followed by the Gemini series, marked DeepMind's serious commitment to competing at the frontier of multimodal and reasoning-rich Al.

3. Meta

Meta joined the LLM race with the introduction of LLaMA. Unlike other major players, Meta made its models downloadable and fine-tunable, enabling broad adoption across academia and startups, and helping drive the open-source LLM movement forward.

4. Anthropic

Founded by former OpenAI researchers, Anthropic entered the LLM race with a strong emphasis on AI safety and alignment. Their Claude models, named after Claude Shannon, quickly became known for their balanced reasoning abilities and enterprise-readiness.

5. DeepSeek

DeepSeek emerged from China as a leading force in the global LLM landscape, releasing high-performance, multilingual models designed for reasoning and code generation. Its open-weight philosophy and competitive benchmarks attracted worldwide attention, reinforcing its status as a serious international contender.

6. xAI

Launched by Elon Musk, xAI entered the LLM space with its Grok series. Tightly integrated with X (formerly Twitter), Grok stood out for its direct consumer deployment, bold personality, and alignment with Musk's vision of "maximally truth-seeking" artificial intelligence.

In addition to APIs from model providers, several key infrastructure components have enhanced the ability for developers to connect LLMs to their applications:

Amazon Bedrock

Amazon Bedrock is a managed service from AWS that gives developers easy access to foundation models from providers like Anthropic, Meta, and Amazon. It handles all the underlying infrastructure, allowing teams to integrate AI into their applications via simple APIs, with options for finetuning and secure, scalable deployment.

Agent Builder Tools (e.g., ZeroWidth, Glean)

Agent builder platforms like Glean and ZeroWidth provide the tools to create LLM-based assistants that can plan, reason, and use external tools or APIs. They manage memory, context flow, and orchestration behind the scenes, making it easier to build agents that go beyond chat and actually perform tasks across systems.

Anthropic's Model Context Protocol (MCP)

MCP is an open standard developed by Anthropic for structuring and passing context between AI models, tools, and environments. It enables continuity and interoperability across sessions or platforms, helping Al agents retain memory and coordinate tasks in a more consistent, reliable way.

These infrastructure components enable developers integrate AI models into their applications.

Al Model

e.g., GPT LLaMA Gemini

Infrastructure

e.g.,

APIs Amazon Bedrock Agent-builder tools Model Context Protocol

Application

e.g., Cursor Replit Copilot Notion Al Figma





The evolution of ChatGPT

Dubberly Design Office 🔸 Al Advancements 🔸 May 2, 2025

When ChatGPT first launched in 2022, it was essentially just a chat UI for an LLM (a transformer model) which only handled text inputs for simple conversation.

	2023												2024												2025				
Nov Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
🗣 Chat UI																													
ঞ্জি GPT-3.5																													

New, more advanced GPT models introduced the capability to process image and voice inputs, and generate images and voice output.



Web search, plugin support, and code execution further extended ChatGPT's capabilities to access external information, analyze data, do calculations, manipulate files, and perform other tasks.



Custom instructions, memory, and custom GPTs enable a more personalized experience, while canvas and projects improve collaboration, iteration, and organization.

OpenAI also introduced the o-series 'reasoning models' — these models are trained to 'think step-by-step' and are optimized for more structured problem-solving.

The 'Deep Research' capability leverages reasoning to search, interpret, and analyze massive amounts of text, images, and PDFs on the internet, pivoting as needed in reaction to information it encounters.

Models, agents and Multi-Agent Systems

The fundamental mechanism of AI is recognizing a pattern.

That requires training data; the more (high-quality) data the better.

The Al builds a model of the relationships in the data e.g., the probability that one word will follow another.

Then the Al uses the model to predict what comes after a prompt.

Results can be measured and stored as training data, enabling these models to continue to learn over time.

LLMs like GPT-40 are trained with massive datasets of text from across the internet.

Their vast datasets provide them with general knowledge across many different domains.

Knowledge Depth

Breadth

On the other hand, AI models can be trained to be highly specialized on a narrower scope of knowledge.

Breadth

General and specialized models (agents) could be combined in a Multi-Agent System (MAS). A MAS could be configured hierarchically or with a more decentralized arrangement.

In a hierarchical Multi-Agent System (MAS), agents could be connected with general agents 'in charge' — managing the system e.g., by delegating tasks to more specialized agents.

In a hybrid approach, agents could still communicate across the hierarchy.

Specialized

In the future we may have agents managing different domains of our daily lives.

Dubberly Design Office · Al Advancements · May 2, 2025

Home

Agent

For example, one agent might be in charge of managing agents for various household systems like energy usage, the home security system, HVAC, and plumbing.

These agents might, in turn, be in charge of more specialized agents, down to individual daemons in charge of monitoring specific processes or performing specific tasks.

A person with diabetes may use a closed-loop continuous glucose monitor (CGM) which passes data to an app that controls an insulin pump.

Intelligent agents could connect to this system and provide predictions and assistance.

APPENDIX ONE

Collaborative interaction with applications

Today, LLM agents have limited visibility into other applications. Their ability to assist the user by performing tasks directly is limited.

Application

Soon, agents may be embedded more deeply into operating systems, enabling a more collaborative conversation with the ability for them to perform actions.

OpenAl indicates that things are heading in that direction — ChatGPT's desktop app enables more direct sharing of application windows, while their upcoming Operator agent is fully capable of using a remotely hosted browser window.

Today, ChatGPT desktop app enables users to share certain application windows with ChatGPT directly.

OpenAl's 'Operator' agent, still in 'research preview' and not available to all users, is able to browse the web visually and take action.

APPENDIX TWO

Notes, caveats, and more to read

The Model Context Protocol (MCP) is an open standard introduced by Anthropic in November 2024 it's designed to streamline the integration of AI models with external data sources and tools.

Key features:

- **Universal Connectivity:** MCP enables AI models to access diverse data sources through a standardized protocol, eliminating the need for custom integrations for each data source.
- **Client-Server Architecture:** Developers can expose their data via MCP servers or build AI applications (MCP clients) that connect to these servers, allowing for secure, two-way data exchanges.
- **Pre-built Integrations:** Anthropic provides pre-built MCP servers for popular systems like Google Drive, Slack, GitHub, Git, Postgres, and Puppeteer, facilitating rapid deployment.
- **Open-Source Ecosystem:** The protocol is open-source, with SDKs available in multiple programming languages, encouraging community contributions and widespread adoption.

OpenAI and Google DeepMind have both announced decisions to adopt the MCP.

Caveats

A 2022 survey by TechRepublic reported that 97% of CEOs planned a significant investment in Al.

It also showed that 80% of AI products that are developed fail before they go to market and of the remaining 20%, 40% of those never succeed.

Most AI Products Fail [HBR, 2023]

85% of AI Implementations Will Fail by 2022 [Gartner, 2021]

AI Has a PoorTrack Record [Industry Week, 2022]

Additional Readings

The Quest for A.I. 'Scientific Superintelligence' (March 10, 2025) An ambitious start-up embodies new optimism that artificial intelligence can turbocharge scientific discovery. <u>https://www.nytimes.com/2025/03/10/technology/ai-science-lab-lila.html</u>

Preparing for the Intelligence Explosion, by Forethought (March 10, 2025) <u>https://www.forethought.org/research/preparing-for-the-intelligence-explosion</u>

Aviv Regev, "Can Al help develop new medicines?" (2023) https://www.ted.com/talks/aviv_regev_can_ai_help_develop_new_medicines

May 2, 2025

Al Advancements and Potential Futures

Dubberly Design Office