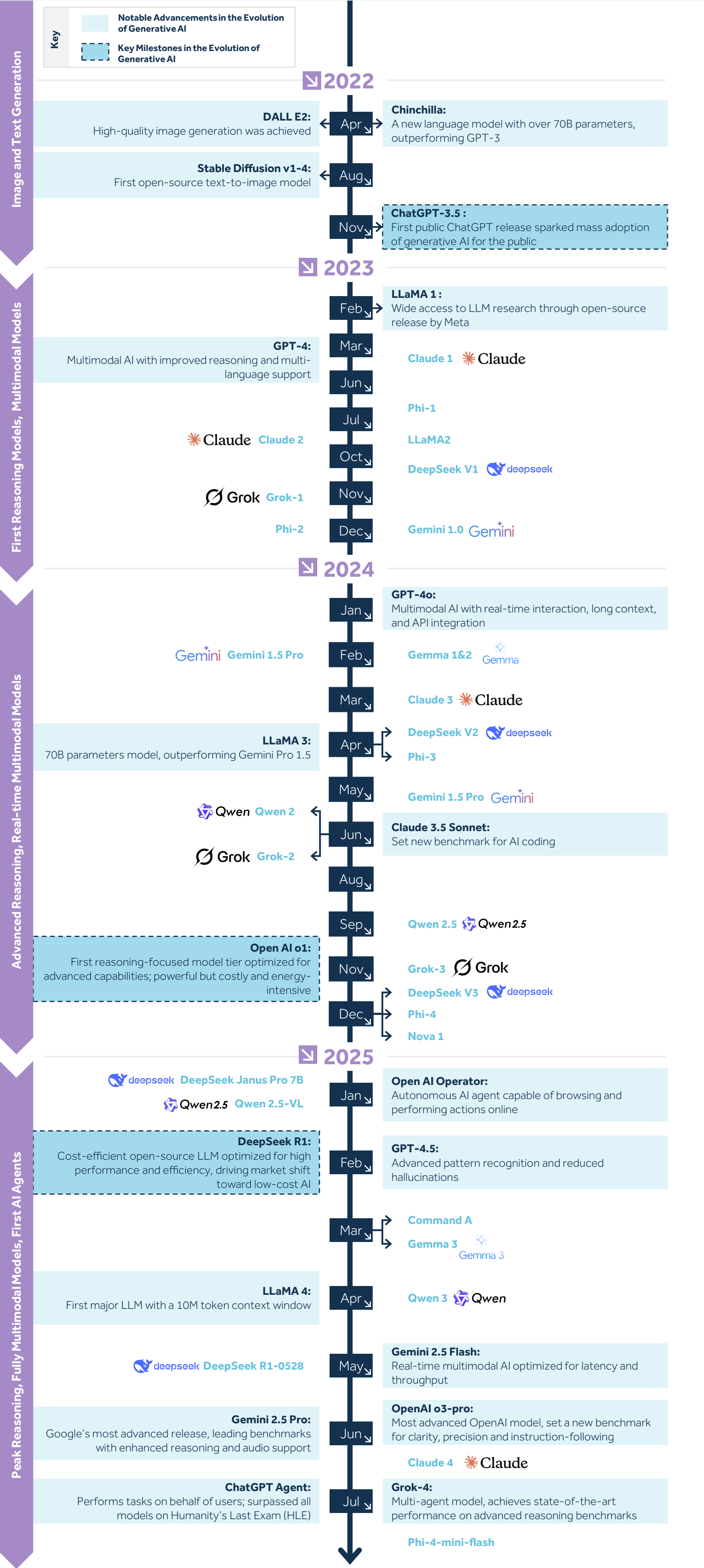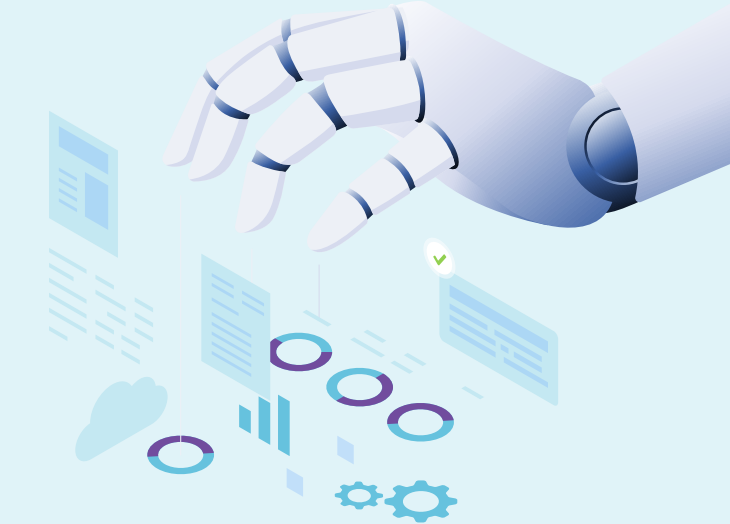# Re:so

Unlocking the Generative AI Era

Between 2022 and 2025, generative AI expanded through the mass adoption of commercial and open-source models by the public. This period is marked by the launch of advanced and specialized models that developed from text and image- generating LLMs to multimodal models with high reasoning capabilities

Key
- Notable Advancements in the Evolution of Generative AI
- Key Milestones in the Evolution of Generative AI

Image and Text Generation

2022

DALL E2:
High-quality image generation was achieved
— Apr —
Chinchilla:
A new language model with over 70B parameters, outperforming GPT-3

Stable Diffusion v1-4:
First open-source text-to-image model
— Aug —

— Nov —
ChatGPT-3.5 :
First public ChatGPT release sparked mass adoption of generative AI for the public

First Reasoning Models, Multimodal Models

2023

— Feb —
LLaMA 1 :
Wide access to LLM research through open-source release by Meta

GPT-4:
Multimodal AI with improved reasoning and multi-language support
— Mar —

— Jun —
Claude 1 ✴ Claude

— Jul —
Phi-1

✴ Claude **Claude 2**
— Oct —
LLaMA2

Ø Grok **Grok-1**
— Nov —
DeepSeek V1 🐋 deepseek

Phi-2
— Dec —
Gemini 1.0 Gemini

Advanced Reasoning, Real-time Multimodal Models

2024

— Jan —
GPT-4o:
Multimodal AI with real-time interaction, long context, and API integration

Gemini **Gemini 1.5 Pro**
— Feb —
Gemma 1&2 ✧ Gemma

— Mar —
Claude 3 ✴ Claude

LLaMA 3:
70B parameters model, outperforming Gemini Pro 1.5
— Apr —
DeepSeek V2 🐋 deepseek
Phi-3

— May —
Gemini 1.5 Pro Gemini

🐦 Qwen **Qwen 2**
— Jun —
Claude 3.5 Sonnet:
Set new benchmark for AI coding

Ø Grok **Grok-2**
— Aug —

— Sep —
Qwen 2.5 🐦 Qwen2.5

Open AI o1:
First reasoning-focused model tier optimized for advanced capabilities; powerful but costly and energy-intensive
— Nov —
Grok-3 Ø Grok

— Dec —
DeepSeek V3 🐋 deepseek
Phi-4
Nova 1

Peak Reasoning, Fully Multimodal Models, First AI Agents

2025

🐋 deepseek **DeepSeek Janus Pro 7B**
🐦 Qwen2.5 **Qwen 2.5-VL**
— Jan —
Open AI Operator:
Autonomous AI agent capable of browsing and performing actions online

DeepSeek R1:
Cost-efficient open-source LLM optimized for high performance and efficiency, driving market shift toward low-cost AI
— Feb —
GPT-4.5:
Advanced pattern recognition and reduced hallucinations

— Mar —
Command A
Gemma 3 ✧ Gemma 3

LLaMA 4:
First major LLM with a 10M token context window
— Apr —
Qwen 3 🐦 Qwen

🐋 deepseek **DeepSeek R1-0528**
— May —
Gemini 2.5 Flash:
Real-time multimodal AI optimized for latency and throughput

— Jun —
OpenAI o3-pro:
Most advanced OpenAI model, set a new benchmark for clarity, precision and instruction-following

Gemini 2.5 Pro:
Google's most advanced release, leading benchmarks with enhanced reasoning and audio support
Claude 4 ✴ Claude

ChatGPT Agent:
Performs tasks on behalf of users; surpassed all models on Humanity's Last Exam (HLE)
— Jul —
Grok-4:
Multi-agent model, achieves state-of-the-art performance on advanced reasoning benchmarks

Phi-4-mini-flash

- The timeline traces AI's evolution from 2022's specialized models like DALL·E 2 and Stable Diffusion to data-efficient large language models (LLMs)
- In 2023, generative AI saw mass adoption with ChatGPT and GPT-4 's multimodality. Open-source models like LLaMA 1 made advanced AI research widely accessible
- By 2024, AI advanced to real-time multimodal interaction and advanced reasoning with GPT-4o, Gemini, and Grok 3. Models like o1 and LLaMA 3.2 brought reasoning and vision capabilities to mobile devices
- In 2025, fully multimodal and reasoning-centric models like o3 Pro, Gemini 2.5 Pro, and Grok-4 emerged while OpenAI Operator marks a shift toward autonomous AI agents