

Al Teamspace
A better way to boost productivity

Content

Promptech	3
Al Teamspace and Prompt Engineering	4
Enterprise Solutions	5
Enterprise Offers	6
Model Pricing	8
GPU Requirement and Pricing for Open-Source Models	10
Workshops and Online Courses	10
Calculation Example	11



Promptech is an AI Teamspace and Prompt Engineering Platform.

Our web application empowers users to engineer prompts, tap into diverse Large Language Models (LLMs), and collaborate on Al projects with their teams.



Key Features:

Page 3

- Prompt Engineering: Craft precise prompts tailored to your Al needs, enhancing model performance and output accuracy.
- LLM Access: Gain access to an array of powerful Large Language Models, ensuring you have the right tools for your projects.
- Team Collaboration: Work together with your team members, driving innovation and efficiency in AI development.
- Al Access Management: Easily manage access permissions, ensuring secure and controlled AI utilization across different teams.
- Enterprise IP Protection: Safeguard your intellectual property with our privately deployed LLM, ensuring your Al projects remain confidential and secure.

What can you do with Promptech?



Streamline

Streamline tasks with Al assistants, optimizing efficiency within your team's workflow.



Productivity

Collaborate in a teamspace brainstorm collectively, leveraging top large language models.



Safeguard

Manage access permissions, controlled Al utilization across teams, enhancing IP Protection

Prompt Engineering

Craft prompts, create AI agents, make edits, and iteratively track versions.

- Generate new prompts, make edits, and track versions effortlessly.
- Utilize prompt templates with parameters to tailor prompts to specific needs.
- Test various prompts and make datadriven decisions.
- Access a public prompt library for inspiration across different tasks.
- Adjust prompts on the fly without needing to write code or redeploy.



Prompt Engineering:

Prompt engineering involves guiding a Large Language Model (LLM) to produce desired outputs by crafting tailored prompts, ensuring precise and desired responses. It is a strategic approach essential for optimizing the performance and functionality of LLMs in various applications.

Al Teamspace

Unify Collaboration and Innovation with AI Teamspace: Access Multiple Models, Streamline Feedback.

- Unite your team in a shared workspace with access to numerous LLM models.
- One Platform, Multiple Models: Access GPT-4, Claude-3, Mistral-Large, and more.
- Keep everyone on the same page by ensuring they are working with the latest prompt versions.
- Log requests and monitor usage history for transparency and accountability.
- Implement user management and governance with distinct access levels for Developers, Members, Guests, and Admins.



Al Teamspace:

An Al Teamspace is a collaborative workspace where teams can access and utilize advanced Large Language Models collectively. It serves as a centralized platform for teams to collaborate, experiment with different LLMs, and streamline the development and deployment of Al solutions.

Enterprise-ready, Accessible To All

Al solution for all businesses, ensuring scalability and inclusivity across diverse structures and demographics.

- Ensure transparency and accountability across your enterprise.
- Benefit from dedicated technical support and flexible hosting options.
- Access exclusive features like Azure OpenAl and Isolated Meta Llama.
- Benefit from enhanced data privacy and IP protection for your enterprise
- Enjoy comprehensive functionality immediately available to your team.



Al Integration:

Bring generative AI to your team and seamlessly integrate it into your daily workflows for enhanced productivity. Experience the power of AI-driven creativity in your team's routine operations.

Empowering Industries with Tailored AI Solutions

Promptech pioneers a transformative approach to industry challenges through bespoke AI solutions, revolutionizing operations across diverse sectors. Without relying on pre-made templates, Promptech empowers organizations to craft custom Co-pilots and AI assistants, tailored precisely to their unique needs. From finance to marketing, legal services to telecommunications, Promptech's adaptive capabilities serve as a versatile virtual assistant, driving efficiency, innovation, and competitive advantage.

With a commitment to tailored solutions, Promptech serves as a versatile virtual assistant, offering support in areas such as financial advice and accounting, marketing strategies, legal services, property management, telecommunications, hospitality, insurance, education, and HR consulting. By facilitating tasks ranging from data analysis and strategy development to customer service automation and regulatory compliance, Promptech ensures that businesses stay ahead of the curve, driving sustainable growth and success in today's competitive environment.

Get in touch with us today to schedule your free consultation and discover how Promptech can revolutionize your industry with customized Al solutions.

Promptech Team info@promptech.ai



Enterprise Offers - Usage Pricing and Plans

Only available at DEMO DAYS

Contact us for enterprise pricing tiers and promotions. info@promptech.ai

Enterprise Offers

Page 6

Enterprise Offer 1: Visionary

+50 Monthly Users

- Unlimited Workspace
- Unlimited Prompt Cards
- Unlimited Guests
- Unlimited Request History
- Unlimited Tokens
- Max Length: 16K & 32K
- Azure OpenAl Private Model
- Fully Isolated Meta Llama and Mistral open-source models with unlimited usage at no additional cost
- Can share VM/GPU costs with others
- New Zealand Data Residency
- 2 Free workshops per year: Leveraging LLM for Business Success (valued at \$2800)
- 5 hours of free consultations per year (valued at \$1250)
- Free Udemy course for all users
- Free OpenAl Tokens (Promotion ends 10 Sep 2024)

NZ\$5 member/month (Full Year, Billed Monthly)
Pay for what you use (USD to NZD conversion applies)
Pay for VM/GPU (See Table 2, USD to NZD conversion applies)

Enterprise Offer 2: Illuminate

+20 Monthly Users

- Unlimited Workspace
- Unlimited Prompt Cards
- Unlimited Guests
- Unlimited Request History
- Unlimited Tokens
- Max Length: 16K & 32K
- Azure OpenAl Private Model
- New Zealand Data Residency
- 1 Free Workshop per year: Leveraging LLM for Business Success (valued at \$1400)
- 2 hours of free consultations per year (valued at \$500)
- Free Udemy course for all users
- Free OpenAl Tokens (Promotion ends 10 Sep 2024)

NZ\$9 member/month (Full Year, Billed Monthly)
Pay for what you use (USD to NZD conversion applies)

Enterprise Offer 3: Perspective

+10 Monthly Users

- Unlimited Workspace
- Unlimited Prompt Cards
- Unlimited Guests
- Unlimited Request History
- Max Length: 16K & 32K
- ½M Tokens / Member / Month
- Purchase Token Packages
- Discounted workshops: Leveraging LLM for Business Success (NZ\$1400 \$300/ workshop)
- Free Udemy course for all users
- Free OpenAl Tokens (Promotion ends 10 Sep 2024)

NZ\$25 member/month

Full Year, Billed Monthly

OpenAl

	Input	Output	Context Window	Max Output
gpt-4o *	\$5.00	\$15.00	128k	4k
gpt-4-turbo **	\$10.00	\$30.00	128k	4k
gpt-4 ***	\$30.00	\$60.00	8k	8k
gpt-3.5-turbo-16k ***	\$0.50	\$1.50	16k	16k
gpt-3.5-turbo ***	\$1.50	\$2.00	4k	4k

^{*} Training Data Up to Oct 2023

Anthropic

	Input	Output	Context Window	Max Output
claude-3-opus *	\$15.00	\$75.00	200k	4k
claude-3-sonnet *	\$3.00	\$15.00	200k	4k
claude-3-haiku *	\$0.25	\$1.25	200k	4k
claude-2.1 **	\$8.00	\$24.00	200k	4k

Google

	Input	Output	Context Window	Max Output
Gemini 1.0 Ultra	TBD	TBD	1M	8k
Gemini 1.5 Pro	\$7.00	\$21.00	1M	8k
Gemini Pro	\$0.50	\$1.50	32k	2k

^{**} Training Data Up to Dec 2023

^{***} Training Data Up to Sep 2021

^{*} Training Data Up to Aug 2023 ** Training Data Up to Early 2023

Mistral

	Input	Output	Context Window	Max Output
mistral-large	\$4.00	\$12.00	32k	32k
mistral-small	\$1.00	\$3.00	32k	32k
mixtral-8x22b *	Unlimited ¹	Unlimited ¹	64k	64k
mixtral-8x7b *	Unlimited ¹	Unlimited ¹	32k	32k
mixtral-7b *	Unlimited ¹	Unlimited ¹	32k	32k

^{*,1} open-source

Meta

	Input	Output	Context Window	Max Output
Llama 3 - 70B *	Unlimited ¹	Unlimited ¹	8k	8k
Llama 3 - 8B *	Unlimited ¹	Unlimited ¹	8k	8k
Llama 2 - 70B *	Unlimited ¹	Unlimited ¹	4k	4k
Llama 2 - 13B *	Unlimited ¹	Unlimited ¹	4k	4k
Llama 2 - 7B *	Unlimited ¹	Unlimited ¹	4k	4k

^{*,1} open-source

Azure OpenAl

	Input	Output	Context Window	Max Output
gpt-4-turbo *	\$10.00	\$30.00	128k	4k
gpt-4 *	\$30.00	\$60.00	8k	8k
gpt-3.5-turbo-16k *	\$0.50	\$1.50	16k	16k

 $^{{}^{\}star}\,\mathsf{Azure}\,\mathsf{security}; \mathsf{Your}\,\mathsf{prompts}\,\mathsf{(inputs)}\,\mathsf{and}\,\mathsf{completions}\,\mathsf{(outputs)}\,\mathsf{are}\,\mathsf{NOT}\,\mathsf{used}\,\mathsf{to}\,\mathsf{improve}\,\mathsf{OpenAl}\,\mathsf{models}.$

	Suitable for	Pay-as-you- go	1 Year Savings Plan
ND96asr A100 v4 (8x A100 NVlink)	Llama 3 – 70B, Llama 2 – 70B, mixtral-8x22b	\$27.1970/ hour	\$22.6197/ hour
NC96ads A100 v4 (4x A100)	Llama 3 – 70B, Llama 2 – 70B, mixtral-8x7b	\$14.6920/ hour	\$12.2296/ hour
NC48ads A100 v4 (2x A100)	Llama 2 – 13B, mixtral-7b	\$7.3460/ hour	\$6.1148/ hour
NC24ads A100 v4 (1x A100)	Llama 3 – 8B, Llama 2 – 7B	\$3.6730/ hour	\$3.0574/ hour

Full-Day Workshops

Workshop: Leveraging LLM for Business Success

LLM Architecture and Prompt Engineering:

Learn the architecture of Large Language Models (LLMs) with a focus on prompt engineering. Effectively craft prompts to your desired outcome.

Identifying Use Cases for LLM across various teams:

Learn how LLMs can be used across various departments. Streamline tasks, generate content, enhance decision-making, and unlock valuable insights from data.

Tailored Solutions for Enterprise:

Hands-on consultation session. Enterprises will work closely with industry experts to identify LLM integration opportunities and benefits within their operations.

Free on-demand Udemy course for all attendants:

Mastering Large Language Models: No-Code Prompt Engineering On-demand video Full lifetime access **ûdemy**

NZ\$1400 NZ\$ 300 The following examples explain how to calculate monthly usage of Enterprise Offer 1 on monthly usage with fully isolated Meta Llama and Mistral opensource models with unlimited usage at no additional cost.

Assumptions:

1 token ~ ¾ words

User cost (fixed cost):

200 users NZ\$ 5 per users / month

Total User Cost: $200 \times 5 = NZ$ 1000$

Token and GPU cost (variable cost):

200 users

Model: gpt-4-turbo

80 request per users / month

200 input tokens (~ 150 words) / request 500 output tokens (~ 375 words) / request

Cost: $200 \times 80 \times ((200 \times (US\$10/1,000,000)) + (500 \times (US\$10/1,000,000)))$

(US\$30/1,000,000)) = 32 + 240 = US\$272

Model: llama 3 - 8B

Unlimited request per users / month

Up to 8k input tokens (~ 6000 words) / request **Up to 8k output tokens** (~ 6000 words) / request

Cost: **Free**

GPU: NC24ads A100 v4 (1x A100)

Plan type: 1 Year Savings Plan

Average total hours / day: 10 hours

Cost: $30 \text{ days } \times 10 \times 3.0574 =$ **US\$ 917.22**

You can share VM/GPU costs with other enterprises.

Total monthly cost:

For 200 Users: NZ\$ 1000 + US\$ 272 + US\$ 917.22 ≈ NZ\$ 2,981

≈NZ\$ 15 per user

The following examples explain how to calculate monthly usage of Enterprise Offer 2 on monthly usage.

Assumptions:

1 token ~ 4 chars in English 1 token ~ ¾ words

User cost (fixed cost):

20 users NZ\$ 9 per users / month Total User Cost: 20 x 9 = **NZ\$ 180**

Token and GPU cost (variable cost):

20 users

Model: gpt-4-turbo 80 request per users / month 200 input tokens (~ 150 words) / request 500 output tokens (~ 375 words) / request Cost: 50 x 80 x ((200 x (US\$10/1,000,000)) + (500 x (US\$30/1,000,000)) = 3.2 + 24 = **US\$ 27.2**

Total cost:

For 20 Users: NZ\$ 180 + US\$ 27.2 ≈ NZ\$ 225 ≈ NZ\$ 11 per user

Contact us for enterprise pricing tiers and promotions.
www.promptech.ai

info@promptech.ai



Notes

promptech