

FPT AI Agents

RELEASE NOTE & HƯỚNG DẪN SỬ DỤNG

Tháng 03/2026

MỤC LỤC

- 1. Tính năng mới 3
 - 1.1. Công cụ tích hợp sẵn (Built-in Tool) 3
 - 1.1.1. Thêm công cụ từ Kho công cụ 3
 - 1.1.2. Sử dụng công cụ trong "Danh sách nghiệp vụ" 4
 - 1.1.3. Sử dụng công cụ trong "Nhiệm vụ" 5
 - 1.1.4. Kiểm tra và vận hành 6
 - 1.2. Tích hợp LINE 7
 - Bước 1: Tạo LINE Official Account 7
 - Bước 2: Cấu hình LINE Webhook 8
 - Bước 3: Cấu hình thông tin LINE trên AI Agent và kết nối 11
 - Bước 4: Kiểm thử tích hợp 12
 - 1.3. Sử dụng Projects trên My Agents 14
 - 1.3.1. Tạo mới Project 14
 - 1.3.2. Sử dụng Project 16
 - 1.3.3. Thao tác với Project 16
 - 1.4. Sử dụng nhiệm vụ trong một nhiệm vụ khác 18
 - 1.5. Phân loại chủ đề câu hỏi (Topic Classification) 19
 - 1.5.1. Cấu hình chủ đề 19
 - 1.5.2. Báo cáo phân loại câu hỏi QnA theo chủ đề 20
- 2. Tính năng cập nhật 21
 - 2.1. Lịch sử nhiệm vụ – Tracing Task 21
 - 2.2. Bật/Tắt quy trình nghiệp vụ 25
 - 2.3. Lọc metadata theo Issue_Date 27
 - Hướng dẫn cấu hình lọc issue_date 27
 - 2.4. Cập nhật Node LLM: Thinking Mode & Completion / Chat Mode 29
 - 2.4.1. Chế độ Completion / Chat Mode 29
 - 2.4.2. Thinking Mode (Chế độ tư duy) 34

1. Tính năng mới

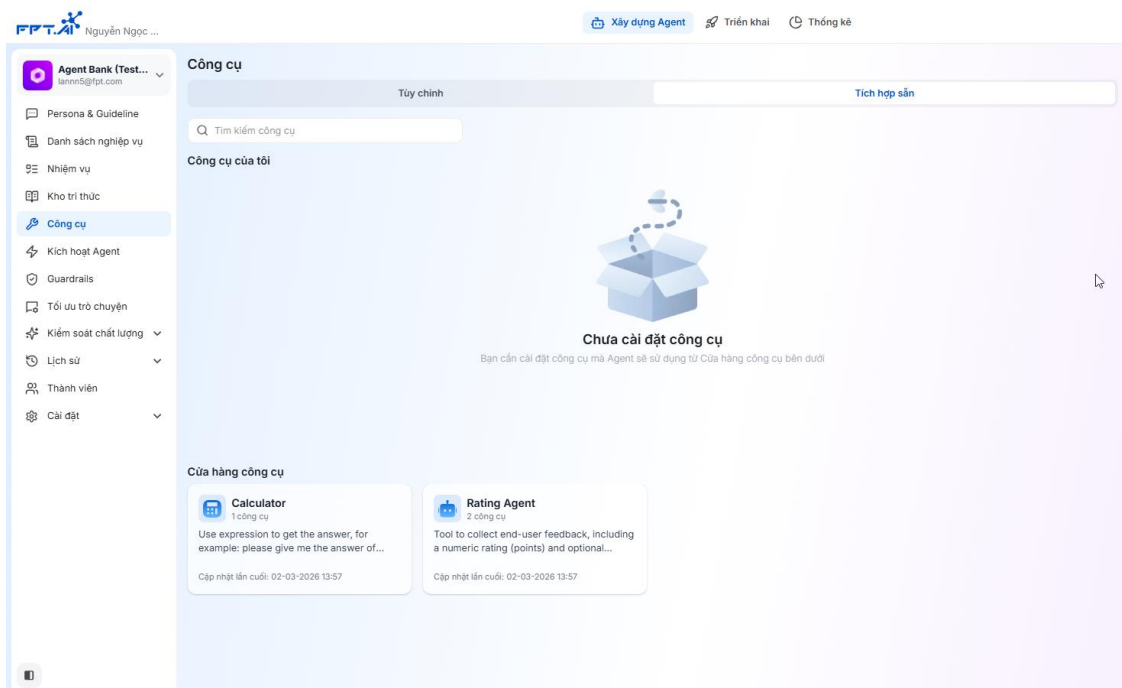
1.1. Công cụ tích hợp sẵn (Built-in Tool)

Công cụ tích hợp sẵn (Built-in Tool) là tập hợp các công cụ được FPT AI Agents tích hợp sẵn vào hệ thống, cho phép AI Agent thực hiện các tác vụ phổ biến một cách nhanh chóng mà không cần cấu hình kỹ thuật phức tạp.

Tính năng này cho phép người dùng:

- Khám phá và sử dụng các công cụ từ kho lưu trữ (Tool Store) của FPT AI Agents.
- Dễ dàng sử dụng các công cụ có sẵn trong quy trình nghiệp vụ / Nhiệm vụ của AI Agent.

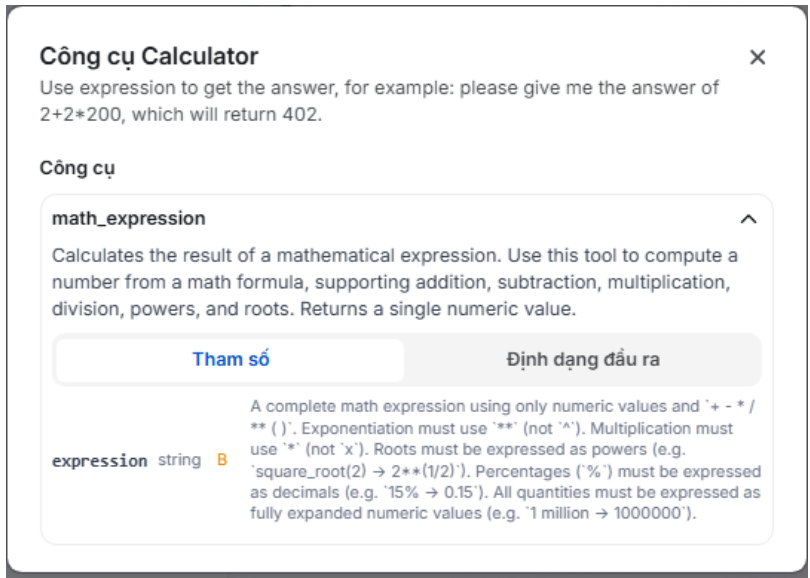
1.1.1. Thêm công cụ từ Kho công cụ



Giao diện Kho công cụ (Tool Store)

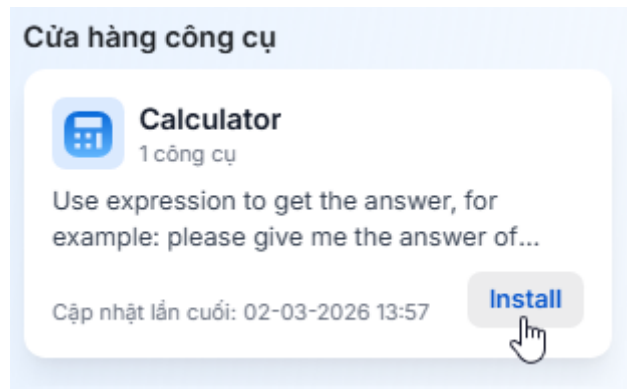
Để sử dụng một công cụ có sẵn, thực hiện các bước sau:

- **Bước 1:** Truy cập menu "Công cụ" từ thanh menu bên trái giao diện quản lý Agent.
- **Bước 2:** Chọn tab "Tích hợp sẵn".
- **Bước 3:** Tại mục "Cửa hàng công cụ", hệ thống hiển thị các công cụ tích hợp sẵn. Người dùng có thể xem chi tiết từng công cụ bằng cách nhấn vào tên công cụ (ví dụ: "Calculator" – dùng để tính toán công thức toán học, kết quả trả về dạng số).



Chi tiết công cụ Calculator

- **Bước 4:** Nhấn hover vào công cụ trong Cửa hàng → chọn "Install". Hệ thống đưa công cụ lên danh sách "Công cụ của tôi". Nếu công cụ yêu cầu API Key, hệ thống sẽ yêu cầu nhập trước khi sử dụng.



Cài đặt công cụ

1.1.2. Sử dụng công cụ trong "Danh sách nghiệp vụ"

Sửa nghiệp vụ
✕

Cập nhật nghiệp vụ theo yêu cầu mới.

Câu mẫu ⓘ
Mẹo: Mỗi câu hỏi mẫu cần được tách bằng dấu xuống dòng.

Gửi 40 triệu trong 2 tháng thì lãi bao nhiêu

44/8000

Hoàn thành

Mục tiêu

Mục tiêu mà quy trình nghiệp vụ này hướng đến 0/255

Chiến lược *

ReAct (Suy luận & Hành động) ▼

Hướng dẫn *

Mẹo: Hướng dẫn nên được mô tả rõ ràng bằng cách đánh số thứ tự, hoặc gạch đầu dòng...

tính lãi suất tiết kiệm của user theo công thức sau Tiền lãi = Số tiền gửi x lãi suất 6%/năm / 12 x Số tháng gửi
 Nếu thiếu số tháng hoặc thiếu số tiền gửi thì cần hỏi lại khách hàng cung cấp thông tin rồi mới thực hiện tính

Nhiệm vụ ⓘ

Chọn nhiệm vụ ▼

Công cụ ⓘ

math_expression x
▼

Custom
Built in

☰
Calculator
☰

✓
math_expression
☰

☰
Rating Agent
☰

☰
rating

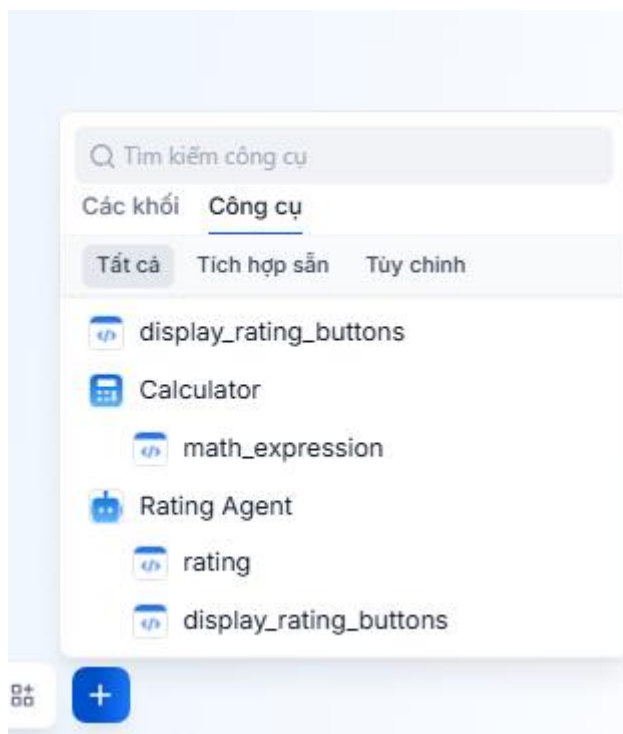
☰
display_rating_buttons

Gắn Built-in Tool vào Nghiệp vụ

Sau khi thêm công cụ vào danh sách Công cụ của tôi, gắn công cụ vào nghiệp vụ cụ thể:

- **Bước 1:** Truy cập menu "Danh sách nghiệp vụ".
- **Bước 2:** Chọn nghiệp vụ cần cấu hình hoặc nhấn "Tạo nghiệp vụ".
- **Bước 3:** Tại trường "Công cụ" > tab "Built in", nhấn vào dropdown list và chọn công cụ vừa cài đặt.
- **Bước 4:** Nhấn "Lưu" để hoàn tất cấu hình nghiệp vụ.

1.1.3. Sử dụng công cụ trong "Nhiệm vụ"



Gắn Built-in Tool vào Nhiệm vụ

- **Bước 1:** Truy cập menu "Nhiệm vụ".
- **Bước 2:** Chọn nhiệm vụ cần cấu hình hoặc nhấn "Tạo" để tạo nhiệm vụ mới.
- **Bước 3:** Nhấn "Thêm khối" → chọn tab "Công cụ" → tab "Tích hợp sẵn" → chọn công cụ.

1.1.4. Kiểm tra và vận hành

Sau khi hoàn tất gắn công cụ vào luồng nghiệp vụ, kiểm tra tại khu vực "Preview and Test":

- Nhập câu hỏi hoặc yêu cầu liên quan đến chức năng của công cụ.
- Quan sát phản hồi của Agent để đảm bảo Agent đã gọi đúng công cụ và trả về kết quả chính xác.

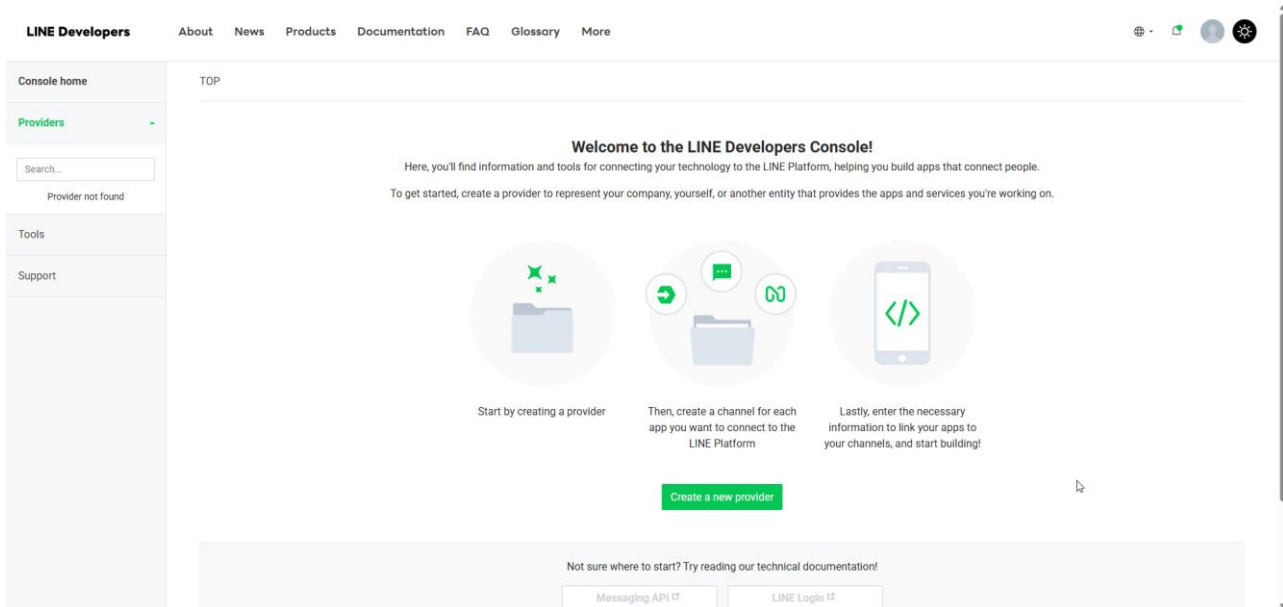
1.2. Tích hợp LINE

Sau khi xây dựng xong AI Agent, người dùng có thể tích hợp Agent lên kênh LINE để khách hàng tương tác.

Bước 1: Tạo LINE Official Account

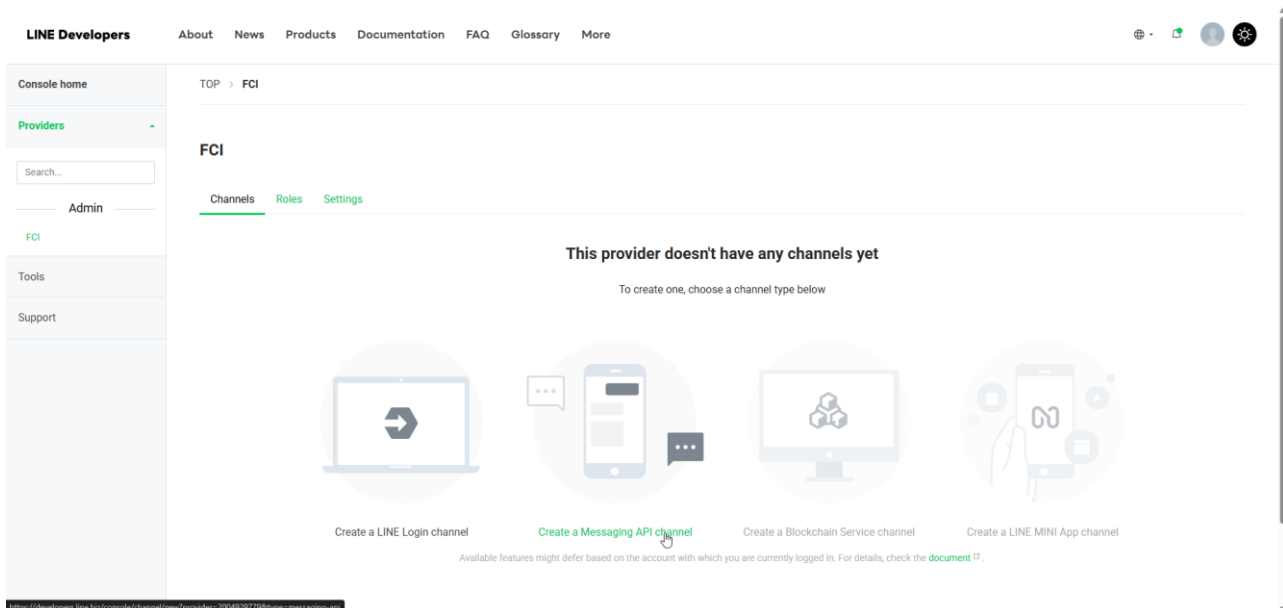
Để tích hợp AI Agent với LINE, người dùng cần có Tài khoản chính thức (Official Account) của kênh LINE. Nếu chưa có, thực hiện theo các bước sau:

1. Đăng nhập vào Developers Console: <https://developers.line.biz/console/>



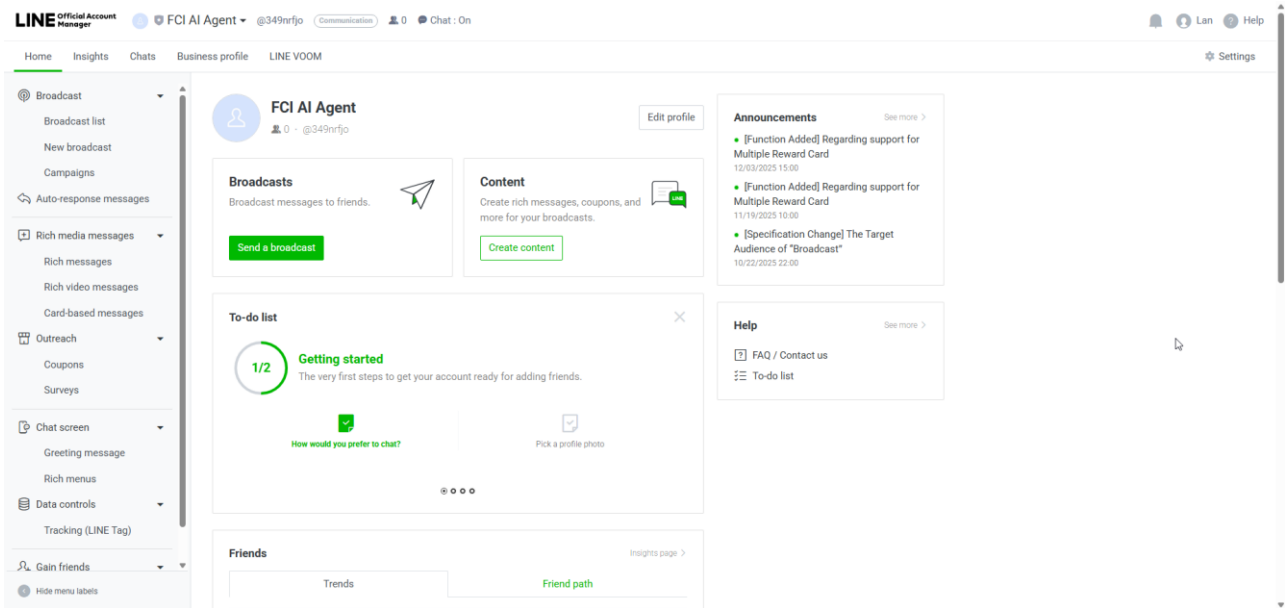
Màn hình LINE Developers Console

2. Tại màn hình chính, nhấn "Creating a new provider" và nhập tên nhà cung cấp.
3. Tại trang Providers > Provider vừa tạo > nhấn chọn "Create a Messaging API channel".



Tạo Messaging API channel

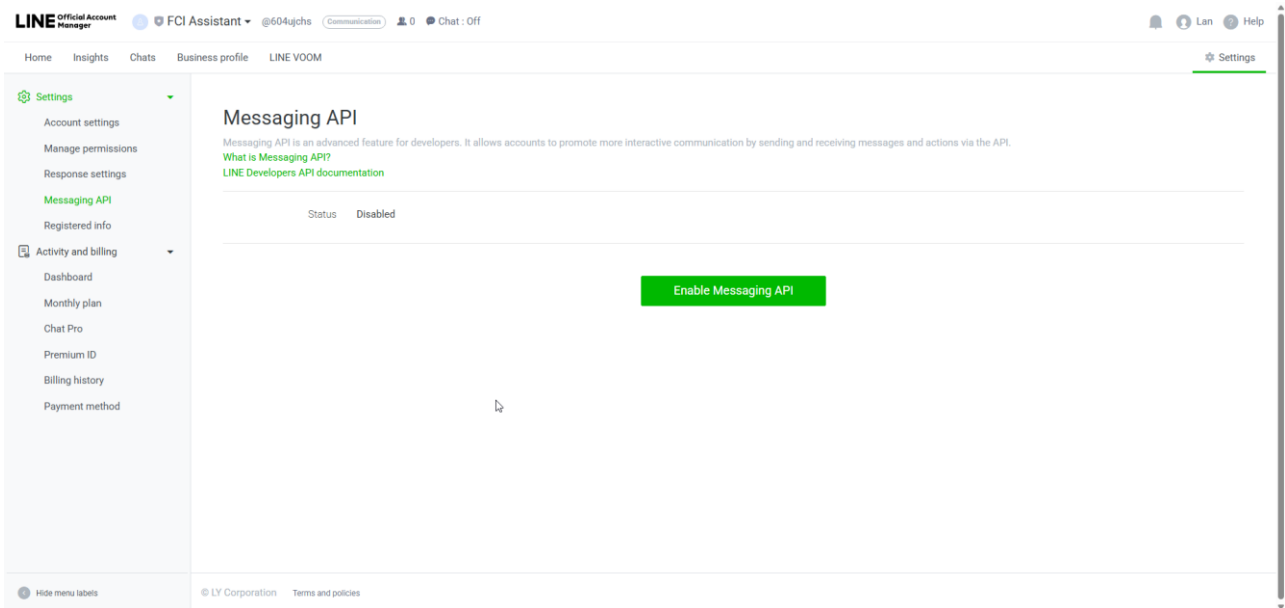
4. Nhấn nút "Create a LINE Official Account" để tạo OA mới.
5. Nhập thông tin theo yêu cầu. Sau khi tạo OA thành công, hệ thống chuyển hướng sang LINE Official Account Manager (<https://manager.line.biz/account/>).



LINE Official Account Manager

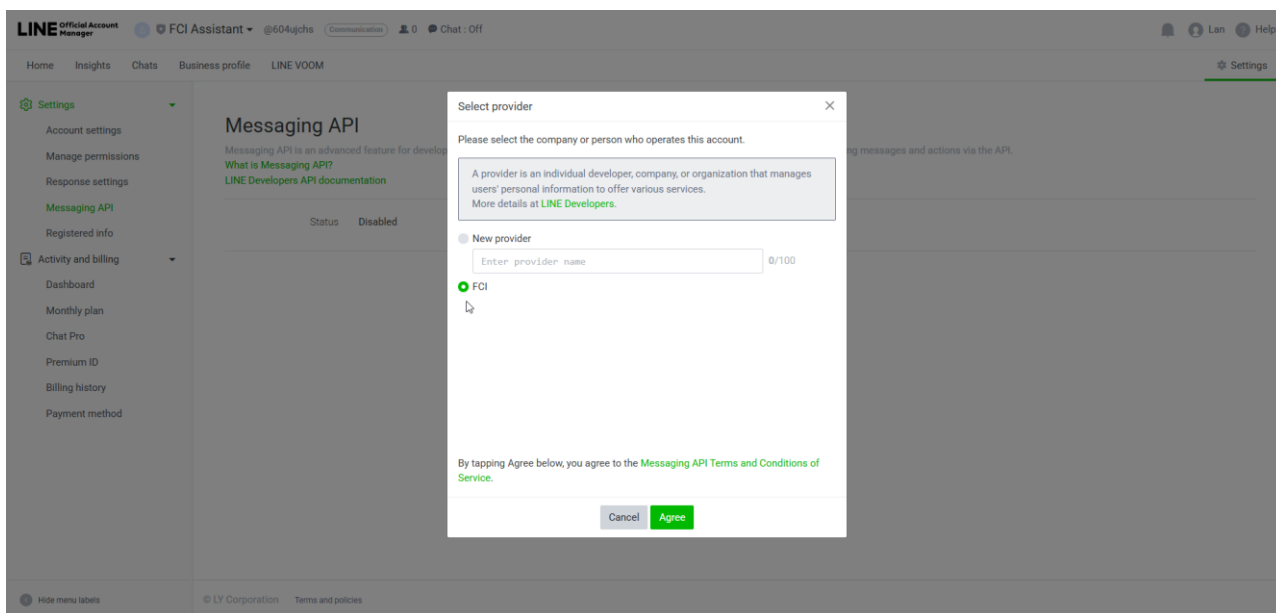
Bước 2: Cấu hình LINE Webhook

6. Chọn tab "Settings" > "Messaging API". Trạng thái đang là Disable → nhấn "Enable Messaging API".



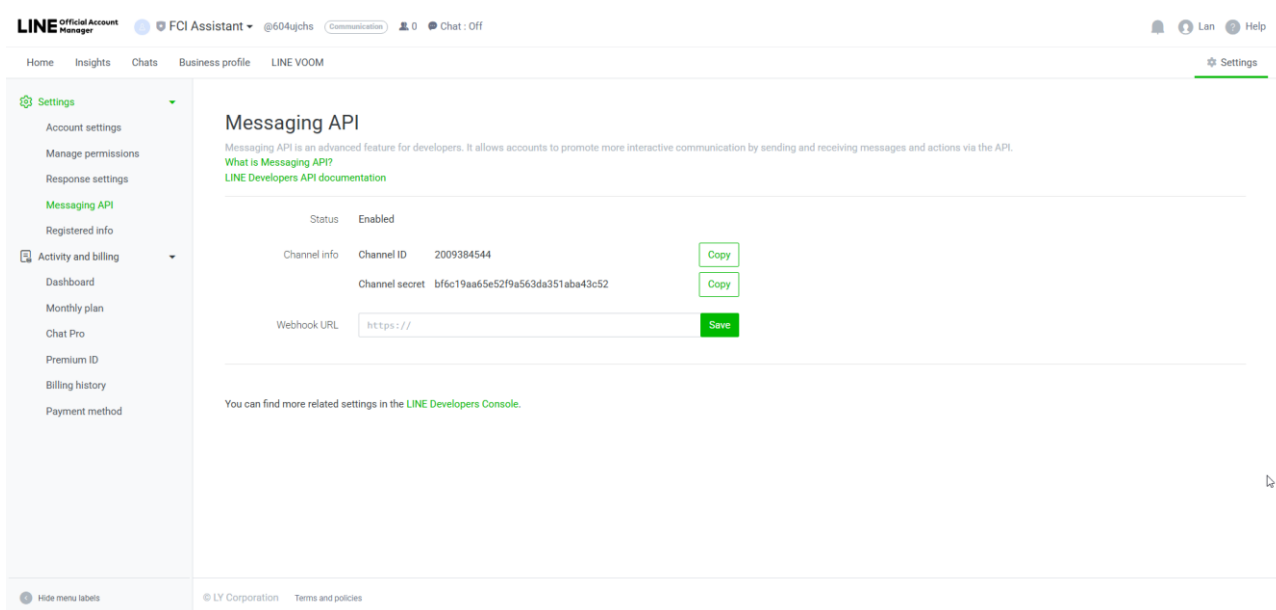
Bật Messaging API

7. Chọn Provider để vận hành account, đồng ý ở các màn hình tiếp theo.



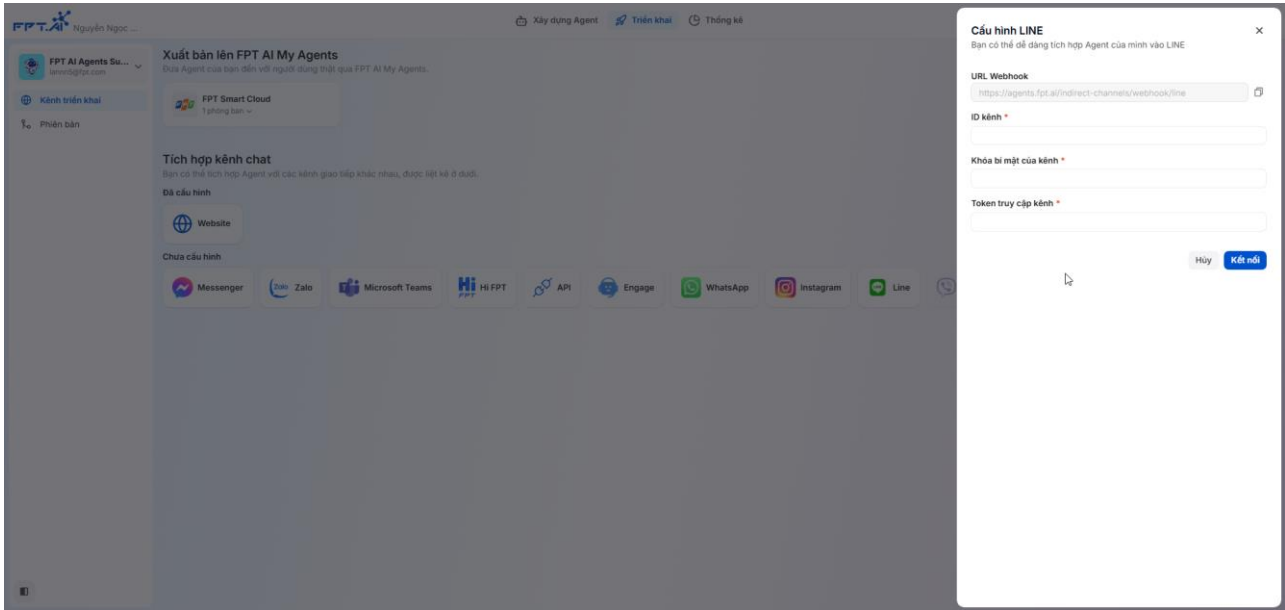
Chọn Provider

8. Sau khi enable, hệ thống LINE yêu cầu nhập Webhook URL từ AI Agent và hiển thị Channel ID, Channel secret.



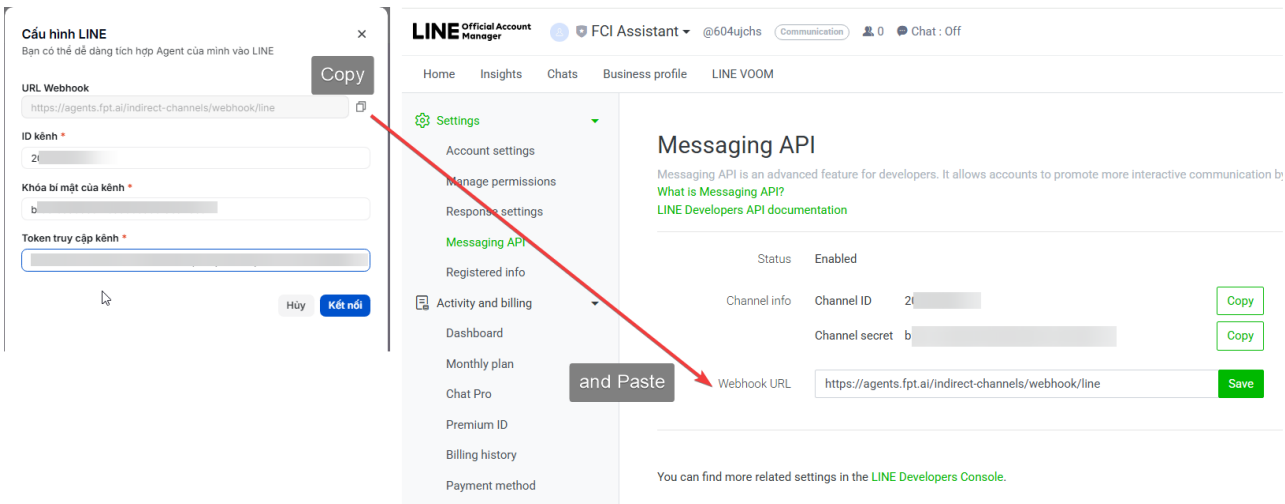
Thông tin Webhook & Channel

9. Truy cập agents.fpt.ai > tab Triển khai > Tích hợp kênh chat > LINE.



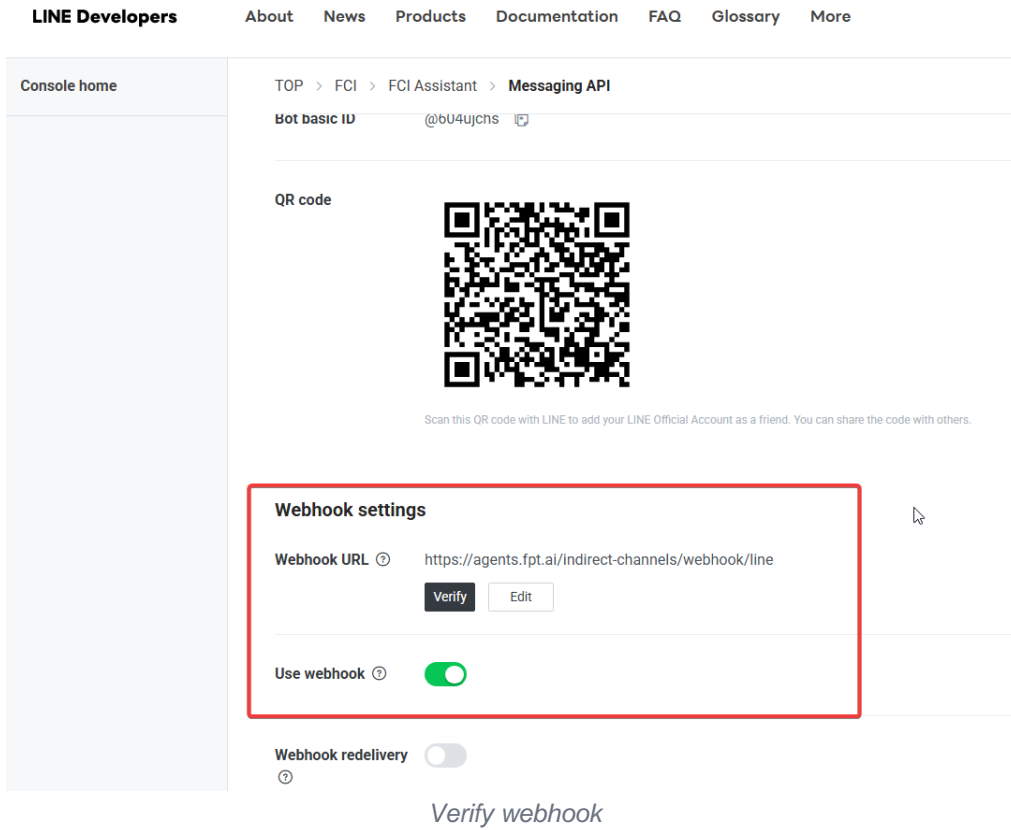
Cấu hình LINE trên AI Agent

10. Copy URL Webhook từ AI Agents và dán vào trường Webhook URL của LINE Official Account Manager.



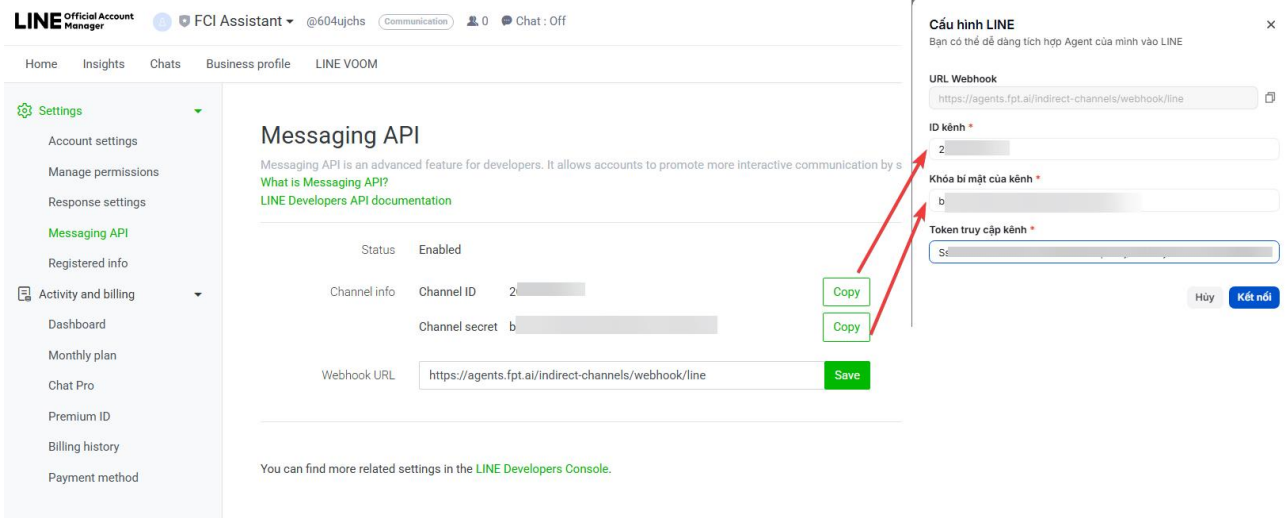
Dán Webhook URL

11. Verify webhook: Truy cập LINE Developers Console > chọn Provider > tab Messaging API > Webhook settings > nhấn "Verify" > bật "Use webhook".

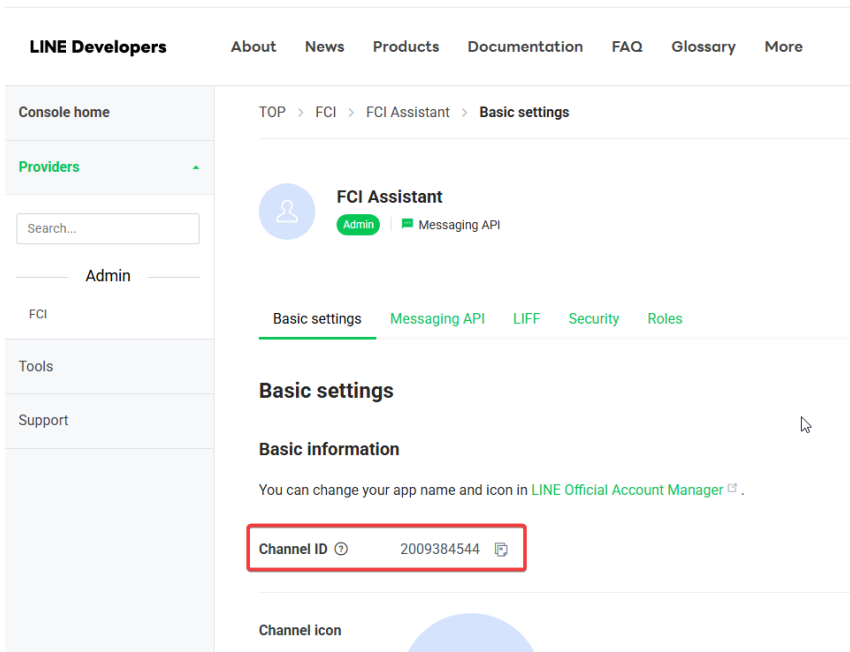


Bước 3: Cấu hình thông tin LINE trên AI Agent và kết nối

12. Copy Channel ID và Channel secret từ LINE Official Account Manager (Settings > Messaging API hoặc Developers Console > Basic settings) và dán vào trường tương ứng trên AI Agent.

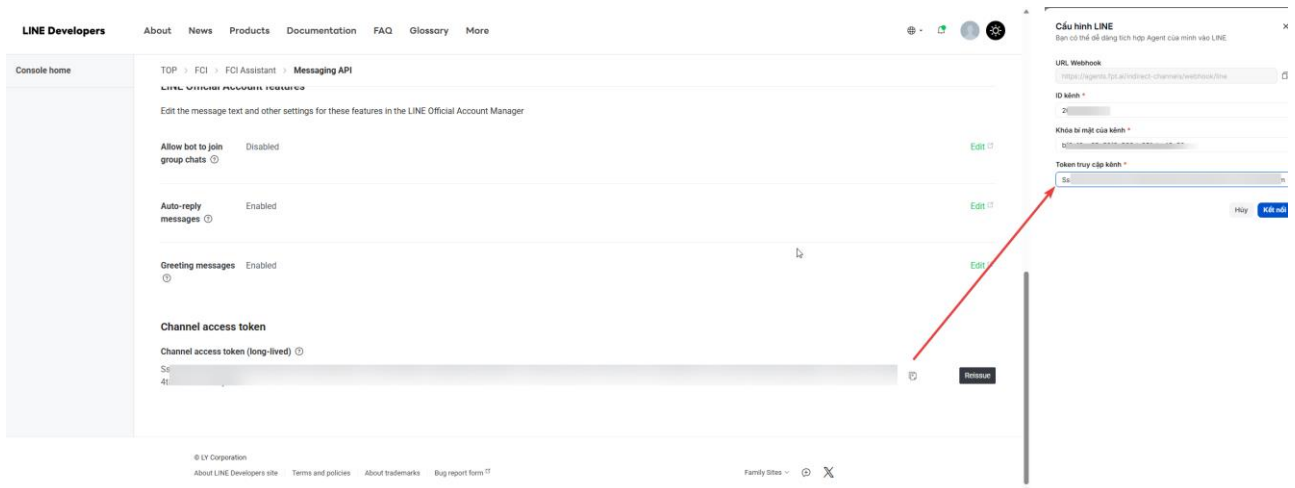


Điền Channel ID & Channel secret



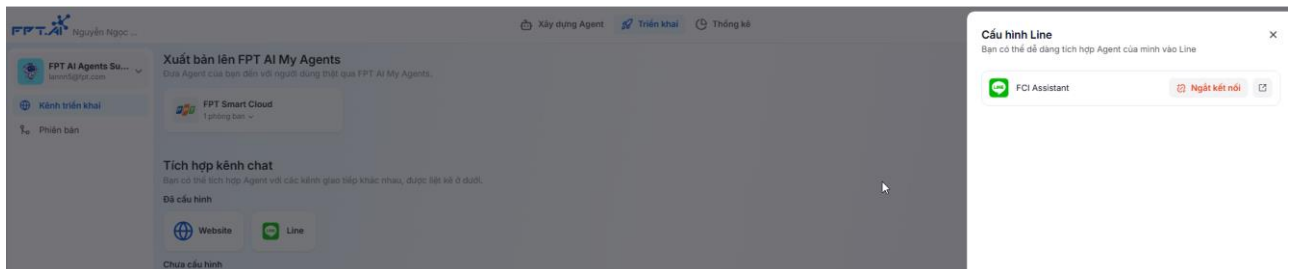
Vị trí Channel ID & secret trên LINE

13. Truy cập LINE Developers Console > Provider > tab Messaging API > cuối trang > "Channel access token (long-lived)" > nhấn "Issue" > copy token và dán vào "Token truy cập kênh" trên AI Agent.



Lấy Channel access token

14. Nhấn "Kết nối". Hệ thống thông báo tích hợp thành công và hiển thị tên OA LINE đã tích hợp.



Kết nối thành công

Bước 4: Kiểm thử tích hợp

Quét mã QR của LINE OA và kiểm thử tương tác với Agent:

LINE Developers About News Products Documentation FAQ Glossary More

Console home

Admin

FCI

Tools

Support

TOP > FCI > FCI Assistant > **Messaging API**


Basic settings **Messaging API** LIFF Security Roles

Messaging API settings

Bot information

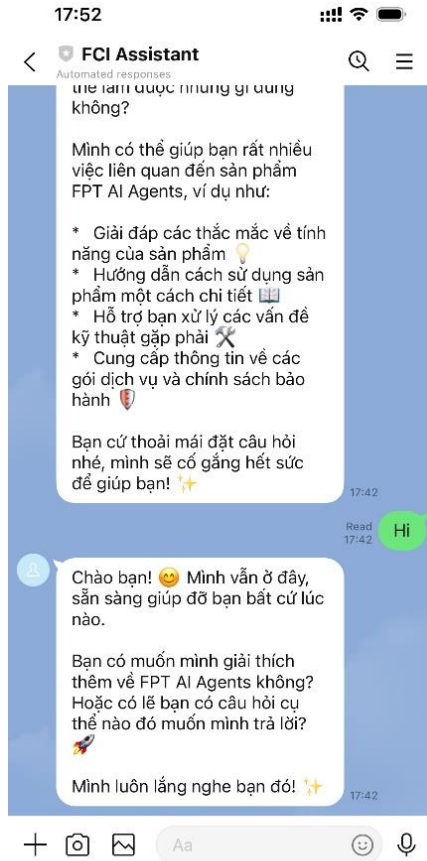
Bot basic ID @604ujchs

QR code



Scan this QR code with LINE to add your LINE Official Account as a friend. You can share the code with others.

Mã QR LINE OA



Kiểm thử trên điện thoại

1.3. Sử dụng Projects trên My Agents

Tính năng Projects cho phép tạo các không gian làm việc riêng để tổ chức và lưu trữ dữ liệu theo từng mục đích. Projects gắn với từng AI Agent.

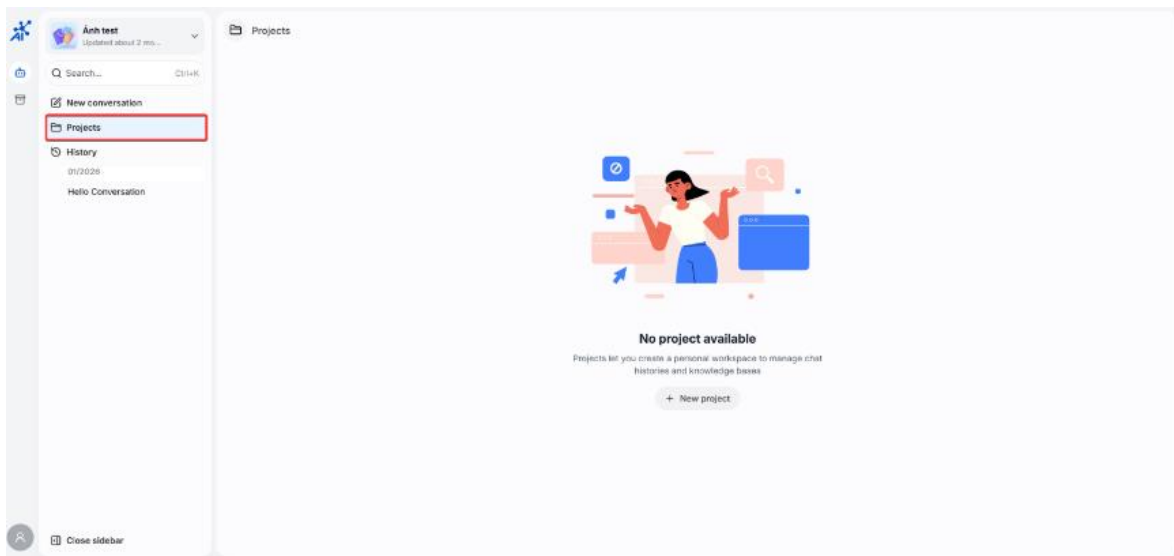
Với Projects, người dùng có thể:

- Tập trung các cuộc trò chuyện liên quan vào cùng một nơi.
- Dễ dàng theo dõi và tìm kiếm lại nội dung theo từng mục đích sử dụng.

i Use case tiêu biểu: Tạo Project cho công việc (ví dụ: lập báo cáo theo tháng, quý) hoặc phục vụ học tập, mỗi Project tương ứng với một chủ đề cụ thể.

1.3.1. Tạo mới Project

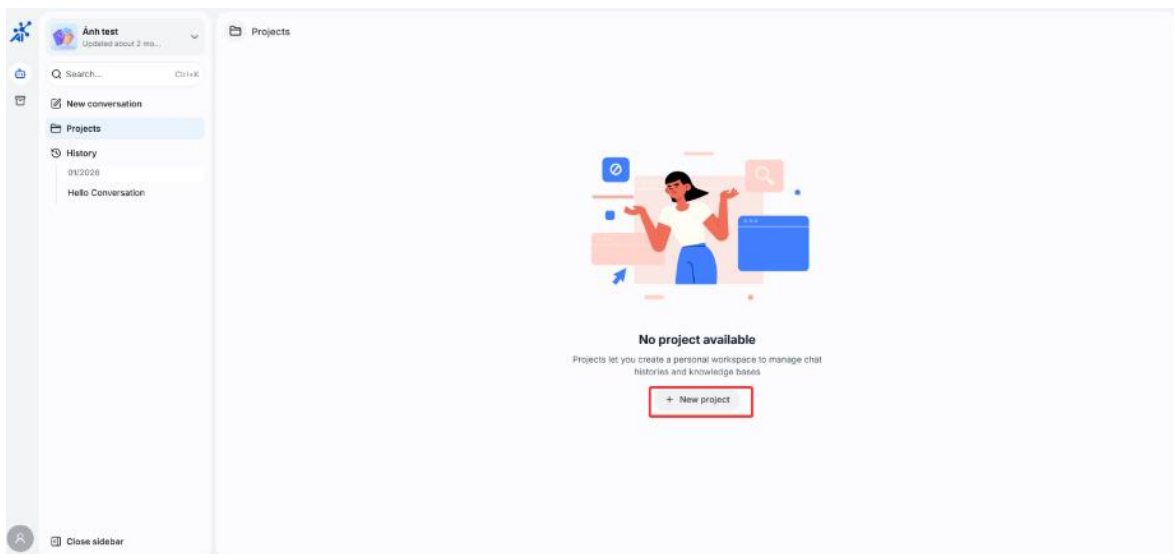
Bước 1. Truy cập một AI Agent, chọn mục Project.



Mục Project trong AI Agent

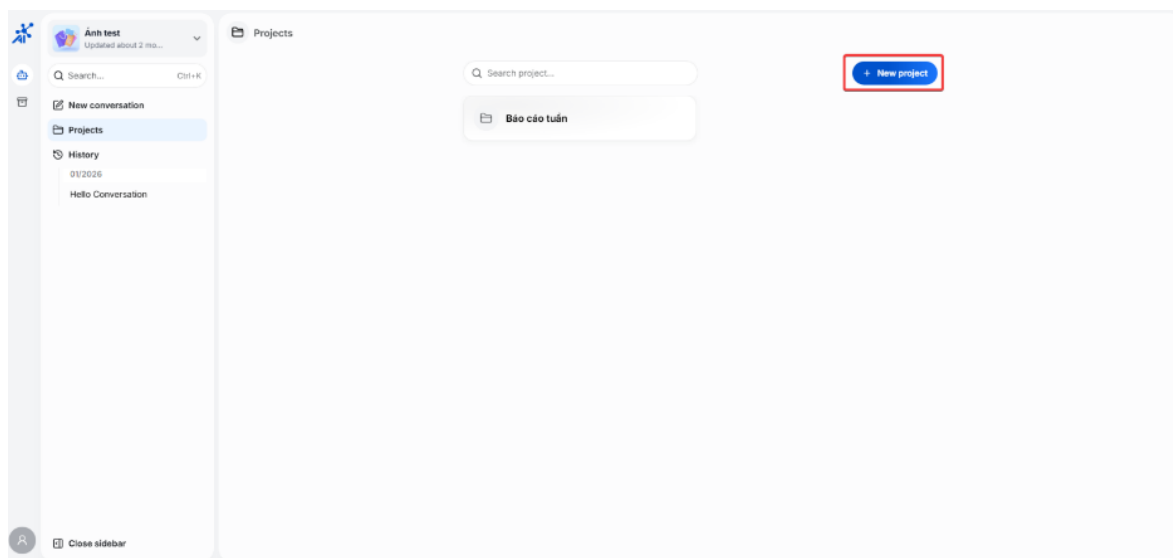
Bước 2. Chọn + New project.

Nếu chưa có project nào, chọn + New project ở giữa màn hình:



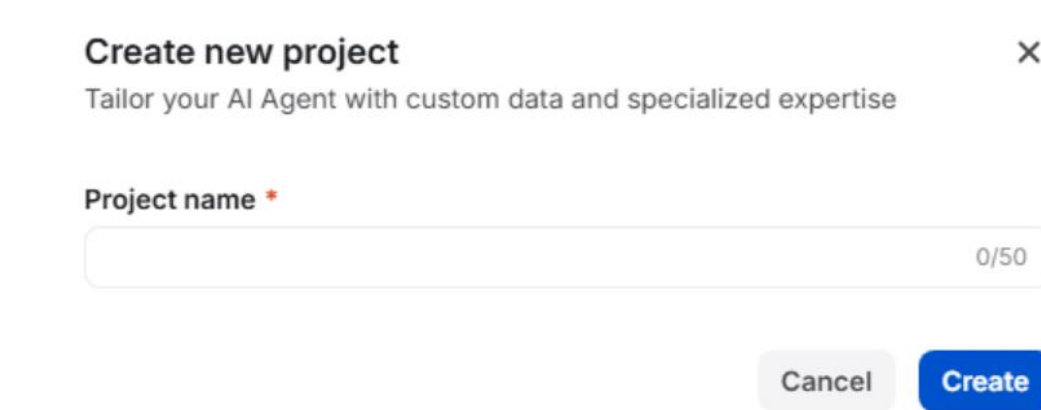
Tạo project đầu tiên

Nếu đã có project, chọn + New project ở góc trên bên phải:



Thêm project mới

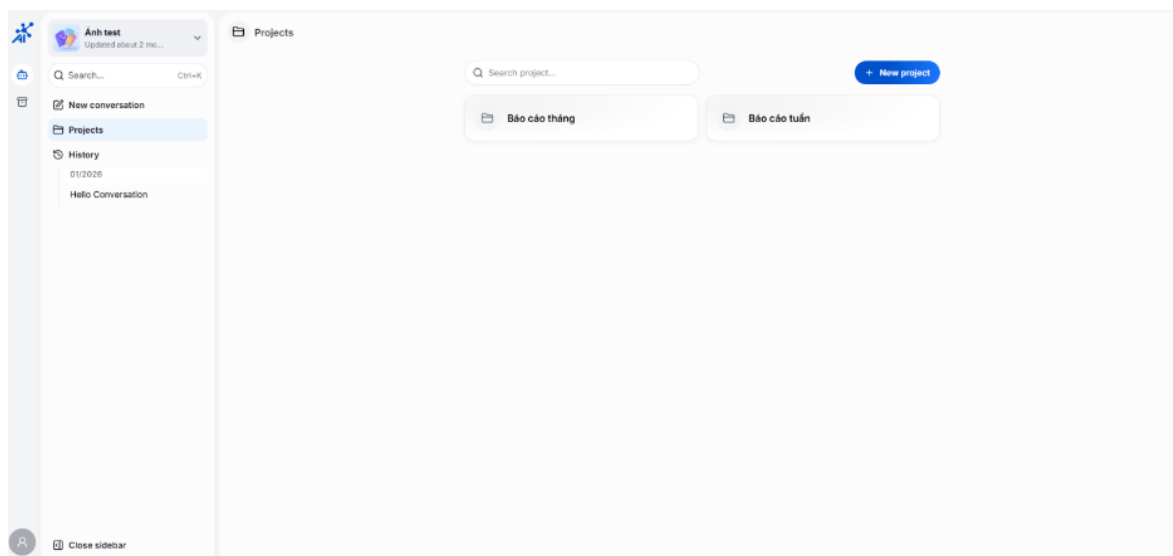
Bước 3. Điền Tên cho project và nhấn Create để xác nhận.



Đặt tên Project

⚠️ Tên project không được trùng nhau.

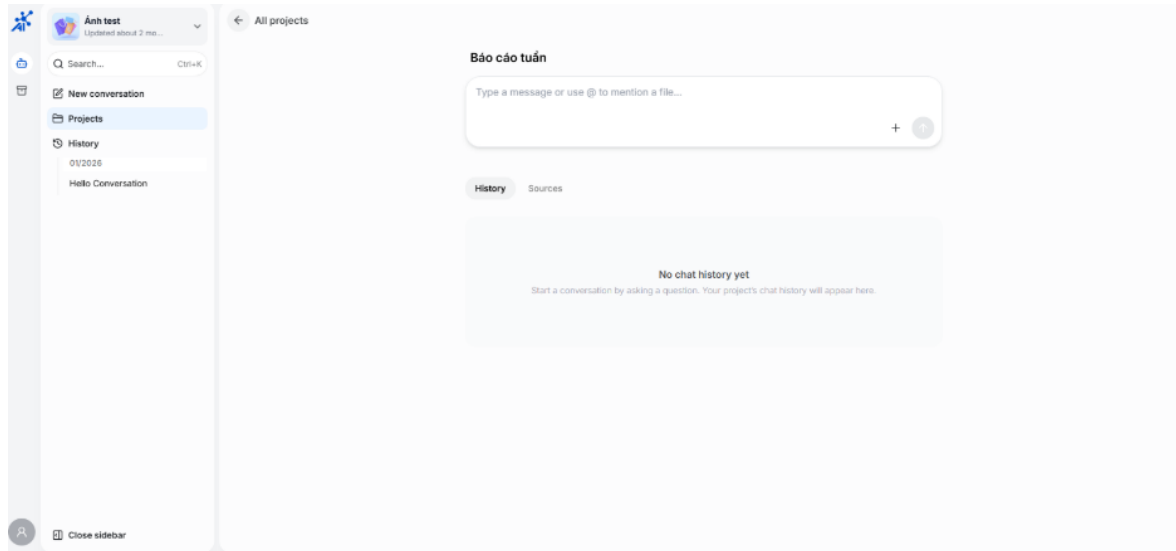
Project sau khi tạo thành công hiển thị trong danh sách:



Danh sách Project

1.3.2. Sử dụng Project

Trong màn danh sách Project, chọn Project muốn sử dụng:

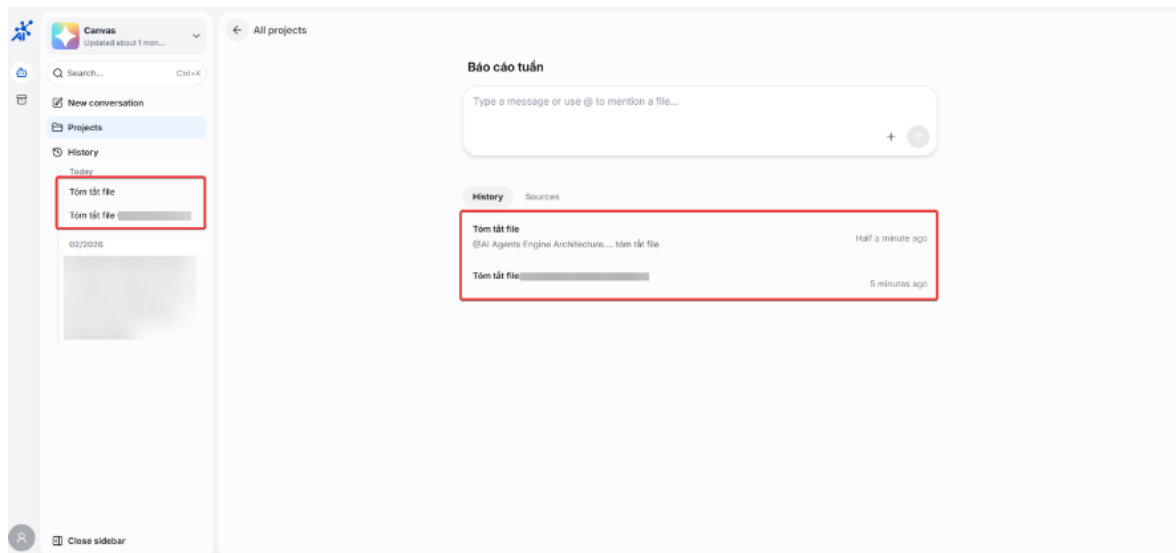


Chọn Project

Người dùng có thể tương tác trong Project như luồng My Agents thông thường:

- Hỏi đáp.
- Đề cập tệp từ My Storage (mention file).
- Tải tệp và hỏi đáp với tệp.

Lịch sử các cuộc trò chuyện hiển thị trong mục History ở trong và ngoài project:

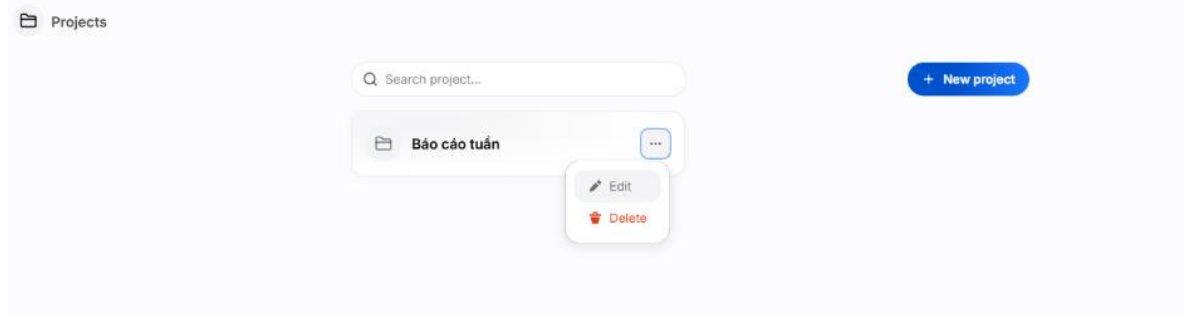


Lịch sử hội thoại trong Project

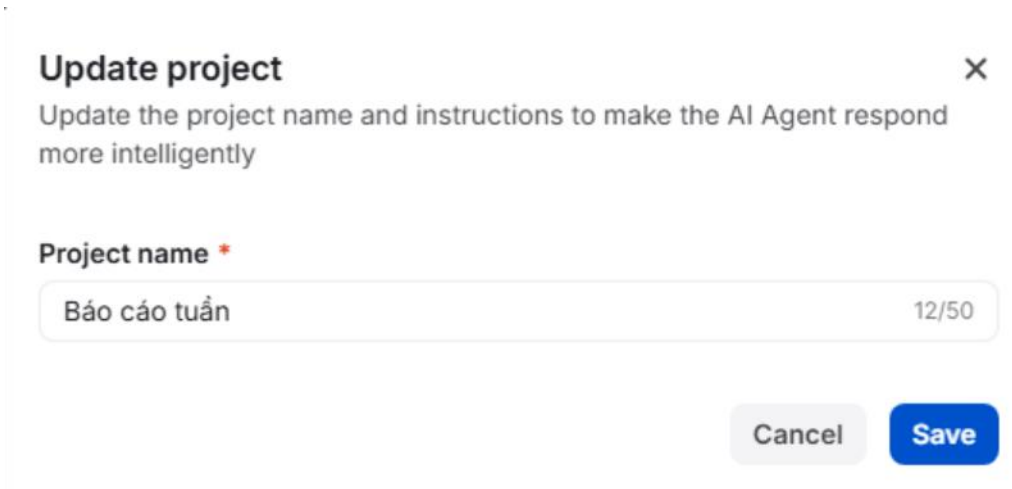
1.3.3. Thao tác với Project

Chỉnh sửa tên project:

Di chuyển chuột đến Project trong danh sách → chọn (...) → nhấn Edit → chỉnh sửa tên → nhấn Save.



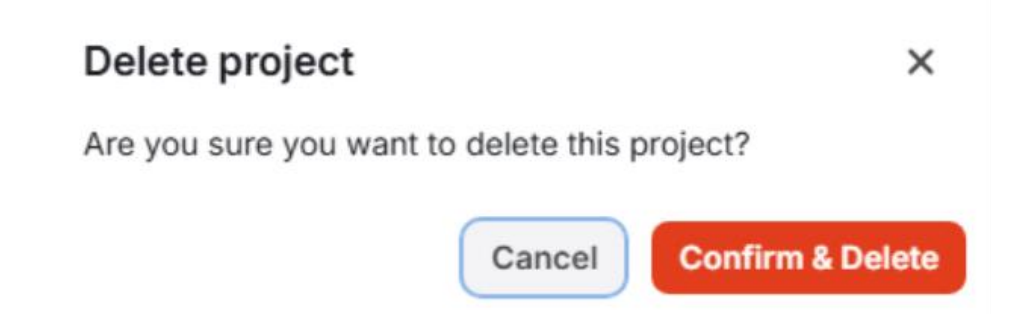
Menu thao tác Project



Chỉnh sửa tên Project

Xóa project:

Di chuyển chuột đến Project → chọn (...) → nhấn Delete → trong popup xác nhận, chọn Confirm & Delete.



Xác nhận xóa Project

1.4. Sử dụng nhiệm vụ trong một nhiệm vụ khác

Trong chức năng "Nhiệm vụ", người dùng có thể gọi tới một nhiệm vụ (Task) khác như một công cụ để thực hiện workflow hoàn chỉnh.

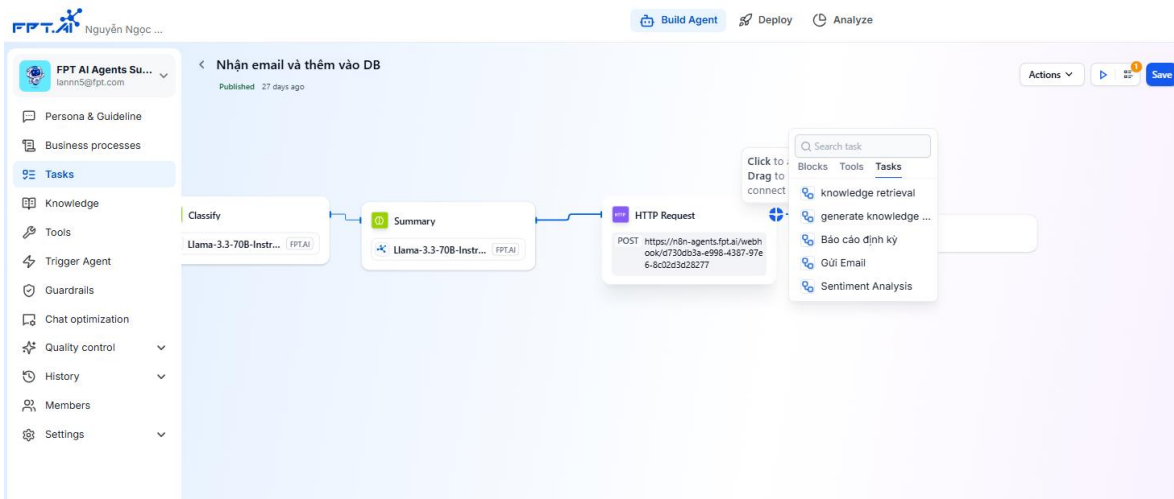
Mục đích của việc này nhằm:

- Tái sử dụng các task nhỏ mà không cần xây dựng trùng lặp.
- Với các task chứa nhiều node, việc tách nhỏ giúp quản lý dễ dàng hơn và dễ truy vết khi thực hiện.

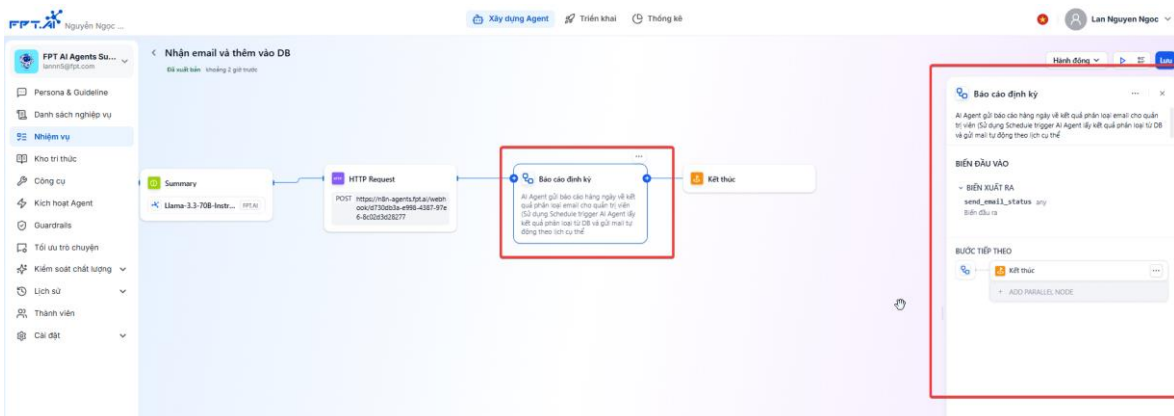
i Để thực hiện, bạn cần có nhiệm vụ chính và một nhiệm vụ cần gọi tới từ nhiệm vụ chính đó.

Bước 1: Truy cập vào nhiệm vụ chính, tại node cần gọi sang nhiệm vụ khác:

- Nhấn nút "Thêm".
- Chọn tab "Nhiệm vụ".
- Chọn nhiệm vụ cần gọi tới.



Thêm block Nhiệm vụ



Chọn nhiệm vụ cần gọi

Bước 2: Kiểm tra nhiệm vụ vừa cài đặt, đảm bảo hệ thống chạy full workflow.

1.5. Phân loại chủ đề câu hỏi (Topic Classification)

Tính năng Phân loại chủ đề cho phép quản trị viên định nghĩa các chủ đề để Agent tự động phân loại câu hỏi QnA của khách hàng. Từ đó, hệ thống thống kê và hiển thị báo cáo tại Analyze > Performance.

Mục tiêu: Giúp quản trị viên dễ dàng cấu hình chủ đề và nắm bắt xu hướng câu hỏi của khách hàng.

1.5.1. Cấu hình chủ đề

Danh sách chủ đề được quản lý tại giao diện Cấu hình chủ đề. Để truy cập:

- Truy cập vào Cài đặt.
- Trong menu Cài đặt, tìm và chọn submenu "Cấu hình chủ đề".

Giao diện hiển thị danh sách chủ đề với các cột: Tên chủ đề/Chủ đề phụ, Mô tả, Ví dụ, Cập nhật lúc và Actions (Sửa, Xóa). Danh sách mặc định sắp xếp theo Last updated (mới → cũ). Người dùng có thể tìm kiếm chủ đề theo tên (không phân biệt hoa/thường và dấu tiếng Việt).

Tạo mới chủ đề

Nhấn nút "Thêm chủ đề". Hệ thống hiển thị popup yêu cầu chọn loại form:

Loại form	Mô tả
Chỉ gồm chủ đề	Cho phép nhập chủ đề chính duy nhất (1 cấp)
Chủ đề & chủ đề phụ	Cho phép nhập chủ đề chính và nhiều chủ đề con (2 cấp)

Tab 1 – Chỉ gồm chủ đề (phân loại 1 cấp):

Trường	Bắt buộc	Ví dụ
Tên chủ đề	Có	"Thẻ ATM", "Vay tín chấp"
Mô tả	Có	Mô tả ngắn nội dung chủ đề
Ví dụ	Không	"Thẻ ATM là gì?", "Làm thế nào để phát hành thẻ ATM?"

Tab 2 – Chủ đề & chủ đề phụ (phân loại 2 cấp):

Trường	Bắt buộc	Ví dụ
Tên chủ đề	Có	Tên chủ đề chính
Tên chủ đề phụ	Có	"Thủ tục cấp thẻ ATM"
Mô tả	Có	Mô tả nội dung cho từng chủ đề phụ
Ví dụ	Không	Câu mẫu liên quan đến chủ đề phụ, mỗi câu một dòng

i Nhấn "+ Thêm chủ đề phụ" để khai báo thêm nhiều chủ đề phụ trong cùng một lần tạo.

Sau khi điền đầy đủ thông tin, nhấn "Tạo" để lưu hoặc "Hủy bỏ" để đóng popup.

Chỉnh sửa và Xóa chủ đề

- **Sửa:** Nhấn nút "Sửa" tại cột Actions để cập nhật thông tin. Việc chỉnh sửa không ảnh hưởng đến dữ liệu thống kê lịch sử đã ghi nhận.
- **Xóa:** Nhấn nút "Xóa" tại cột Actions. Hệ thống hiển thị popup xác nhận trước khi xóa vĩnh viễn.

1.5.2. Báo cáo phân loại câu hỏi QnA theo chủ đề

Sau khi cấu hình chủ đề, xem thống kê tại Analyze > Performance. Giao diện bổ sung hai thành phần:

Thành phần	Mô tả
Biểu đồ QnA Topics (Treemap)	Hiển thị số lượng và tỷ lệ % cuộc hội thoại theo từng chủ đề. Câu hỏi không khớp chủ đề nào được gom vào nhóm Others.
Bảng QnA Subtopic Details	Hiển thị chi tiết số lượng và tỷ lệ % theo từng chủ đề phụ. Có thể nhấn chọn một chủ đề trên biểu đồ để lọc bảng tương ứng.

⚠ Một cuộc hội thoại có thể được ghi nhận vào nhiều chủ đề. Ví dụ: khách hàng vừa hỏi về thẻ, vừa hỏi về vay → hệ thống ghi nhận 1 lượt cho chủ đề "Thẻ" và 1 lượt cho chủ đề "Vay".

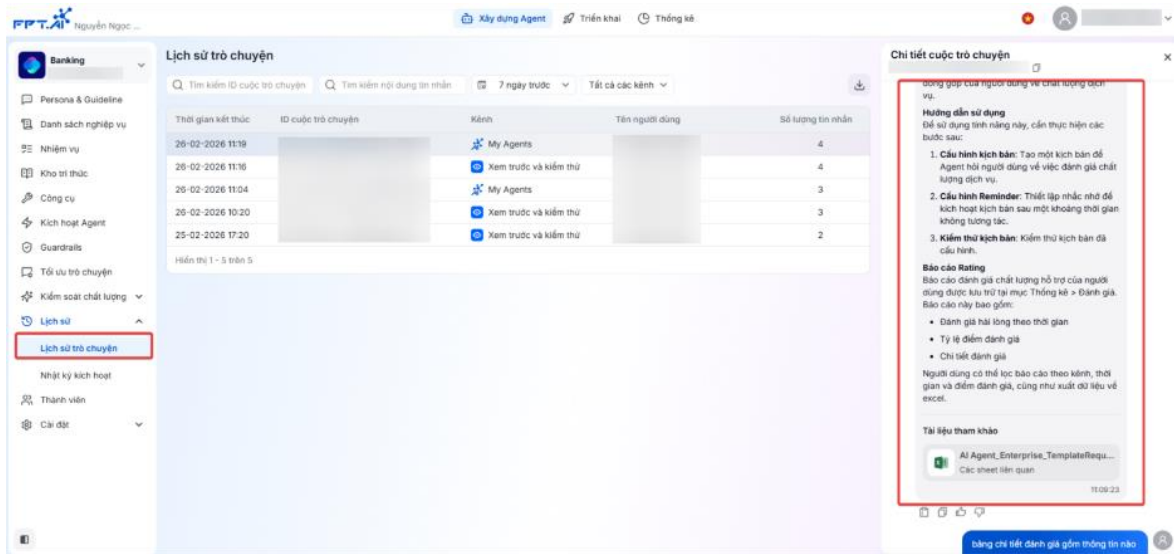
2. Tính năng cập nhật

2.1. Lịch sử nhiệm vụ – Tracing Task

Trong quá trình tương tác với người dùng cuối, Agent thực hiện nhiều Nhiệm vụ (Task) và mỗi Task đi qua các khối xử lý khác nhau. Tính năng Tracing Task trong mục Lịch sử trò chuyện cho phép xem lại toàn bộ luồng thực thi của nhiệm vụ trên thực tế.

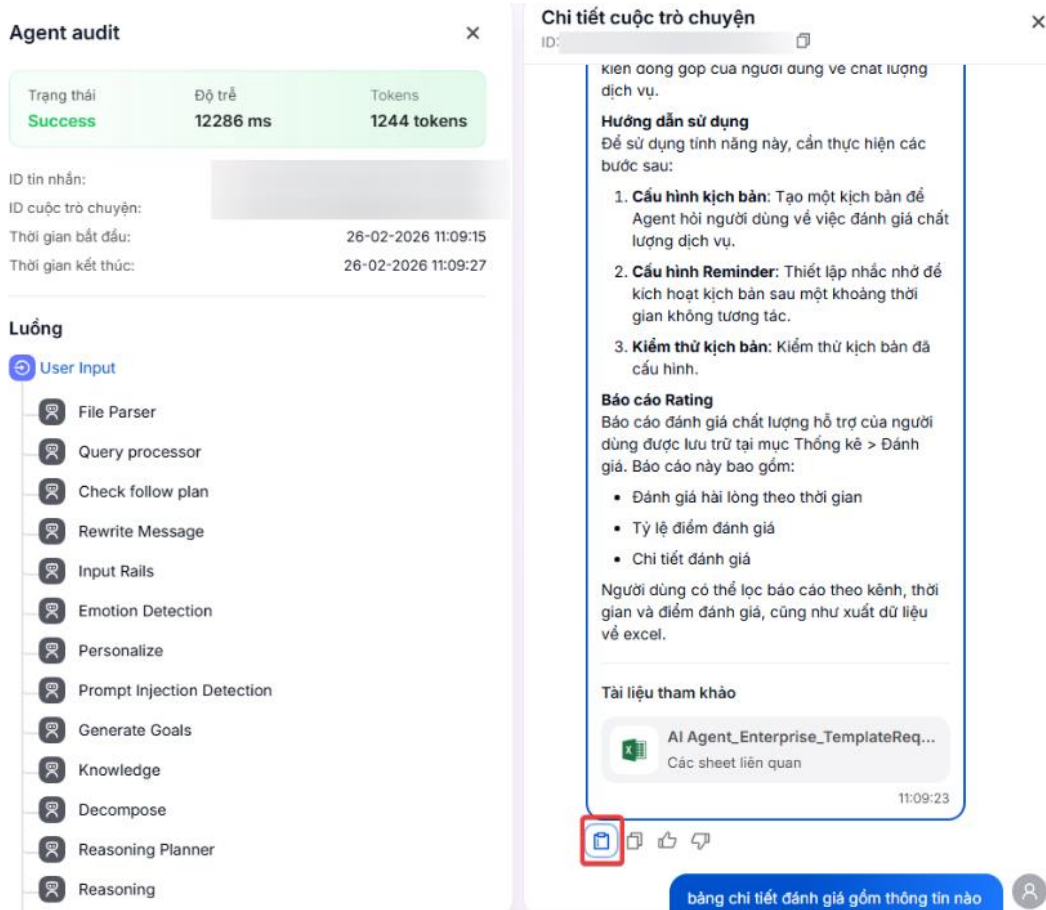
Bước 1. Truy cập mục Lịch sử trò chuyện.

Trong phần Xây dựng Agent, truy cập Lịch sử > Lịch sử trò chuyện, chọn câu trả lời cần xem để mở Tracing Task.



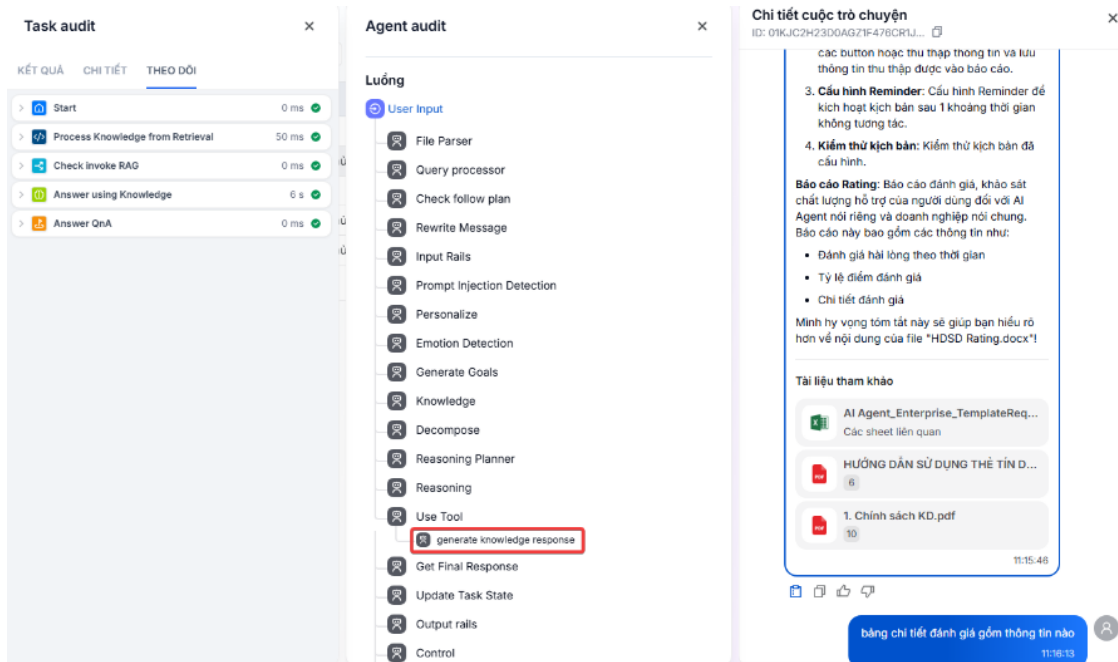
Màn hình Lịch sử trò chuyện

Bước 2. "Chọn Audit" phía dưới câu trả lời.



Nút Audit dưới câu trả lời

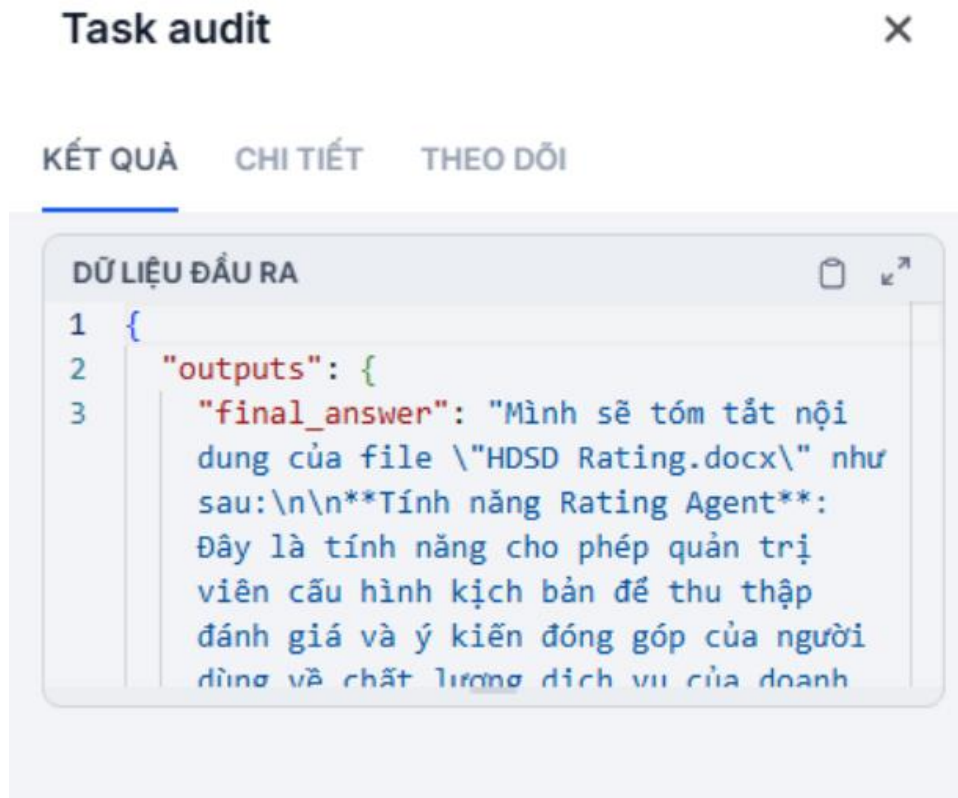
Bước 3. "Di chuyển xuống bước "Use tool", chọn nhiệm vụ muốn xem Tracing."



Danh sách nhiệm vụ trong Use tool

Trong Task Audit hiển thị 3 tabs:

Tab	Nội dung hiển thị
Kết quả	Dữ liệu đầu ra của Task.
Chi tiết	Thời gian thực hiện, trạng thái, dữ liệu đầu vào, tổng số token sử dụng.
Theo dõi	Các khối thực thi nhiệm vụ. Chọn từng khối để xem đầu vào/đầu ra chi tiết.



Tab Kết quả

Task audit ✕

KẾT QUẢ **CHI TIẾT** THEO DÕI

TRẠNG THÁI	THỜI GIAN THỰC THI	TỔNG SỐ TOKEN
● SUCCESS	6 s	0 tokens

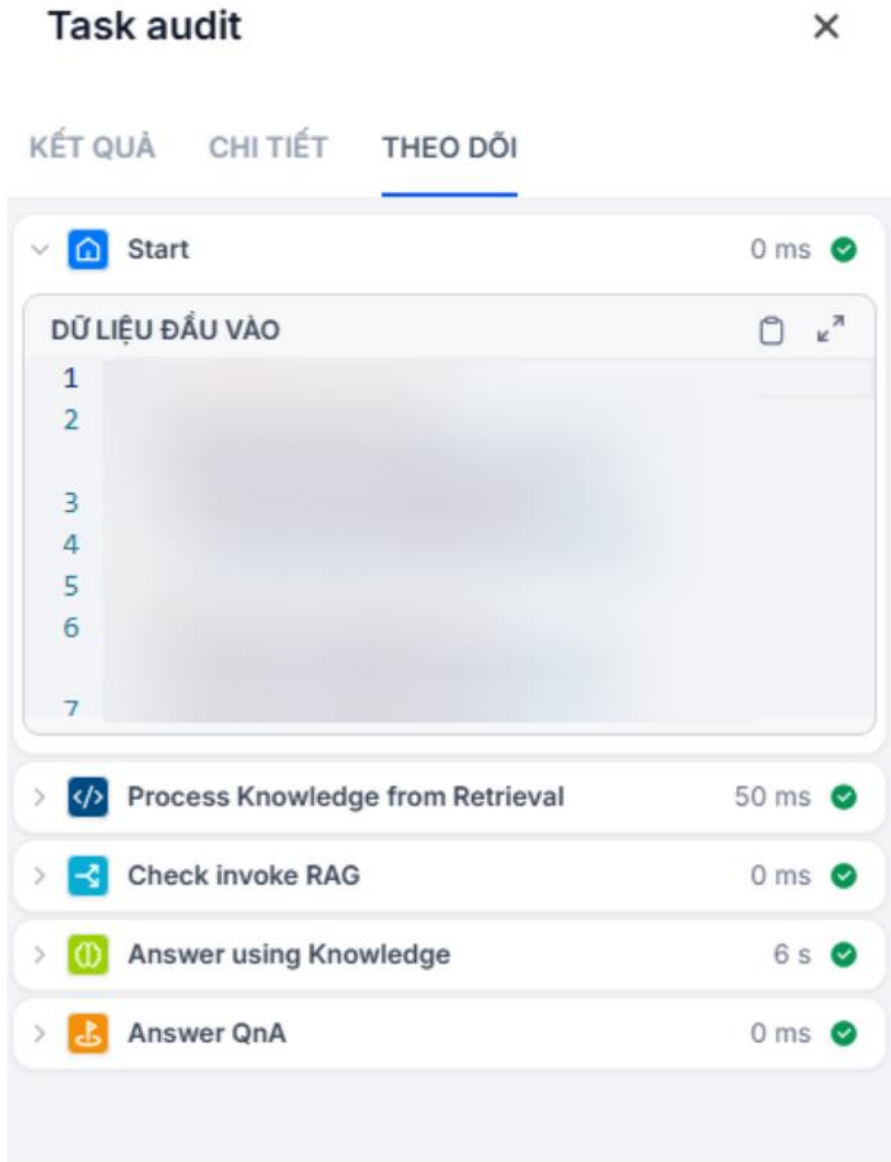
DỮ LIỆU ĐẦU VÀO 🗑️ ↗️

```
1 {  
2  
3  
4  
5  
6  
7
```

DỮ LIỆU ĐẦU RA 🗑️ ↗️

```
1 {  
2   "outputs": {  
3     "final_answer": "Mình sẽ tóm tắt nội  
dung của file \"HSDS Rating.docx\  
như sau:\n\n**Tính năng Rating  
Agent**:\n\nĐây là tính năng cho phép  
quản trị viên cấu hình kịch bản để  
thu thập đánh giá và ý kiến đóng góp  
của người dùng về chất lượng dịch vụ
```

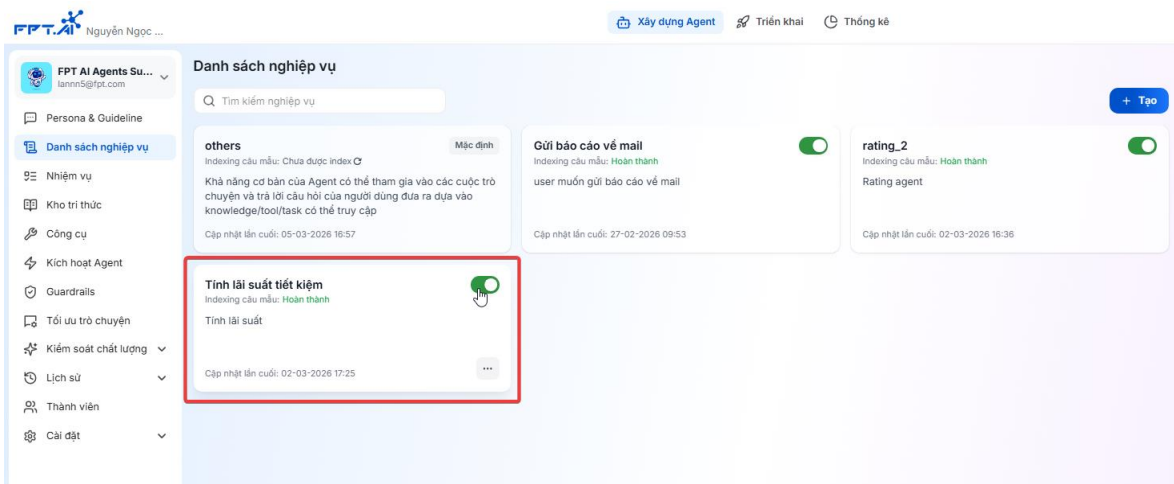
Tab Chi tiết



Tab Theo dõi

2.2. Bật/Tắt quy trình nghiệp vụ

Người dùng có thể bật/tắt từng quy trình nghiệp vụ để thao tác thuận tiện trong quá trình cấu hình.



Toggle bật/tắt quy trình nghiệp vụ

⚠ Lưu ý khi tắt quy trình nghiệp vụ: • Các chức năng khác đang sử dụng quy trình này cũng không hoạt động (tại Kích hoạt Agent, Nhắc nhở, Tối ưu trò chuyện, Kiểm soát chất lượng > Kiểm thử tự động). • Với người dùng cuối: nếu đang thực hiện nghiệp vụ mà bị tắt, Agent vẫn thực hiện hết luồng cho tới khi hoàn thành (tránh gián đoạn). Nếu chưa bắt đầu, Agent sẽ không kích hoạt nghiệp vụ đã tắt.

2.3. Lọc metadata theo Issue_Date

Trong quá trình thực thi nhiệm vụ tại các khối Truy xuất tri thức (Knowledge Retrieval) và Tra cứu tri thức (Knowledge Lookup), hệ thống cho phép lọc dữ liệu dựa trên metadata của tài liệu.

Ngoài các trường đã hỗ trợ (file_name, chunk_type, original_url), hệ thống bổ sung thêm trường:

- **issue_date** – Ngày ban hành / ngày có hiệu lực của tài liệu.

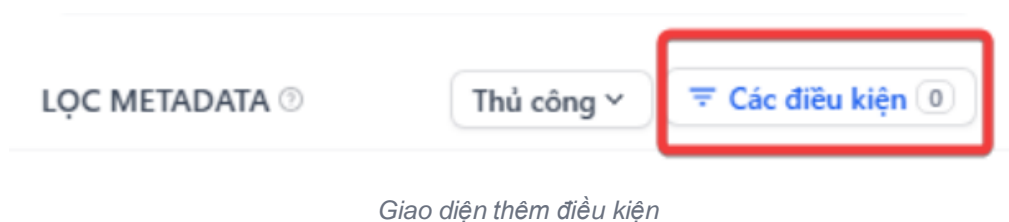
i Use case: Đối với các văn bản pháp luật hoặc tài liệu có ngày ban hành, người dùng có thể cấu hình điều kiện lọc để chỉ lấy các tài liệu có ngày ban hành sau một mốc thời gian xác định, đảm bảo Agent sử dụng phiên bản văn bản mới nhất còn hiệu lực.

Hướng dẫn cấu hình lọc issue_date

Bước 1. Truy cập khối Truy xuất tri thức hoặc Tra cứu tri thức cần lọc. Trong phần Lọc metadata, chọn "Thủ công".

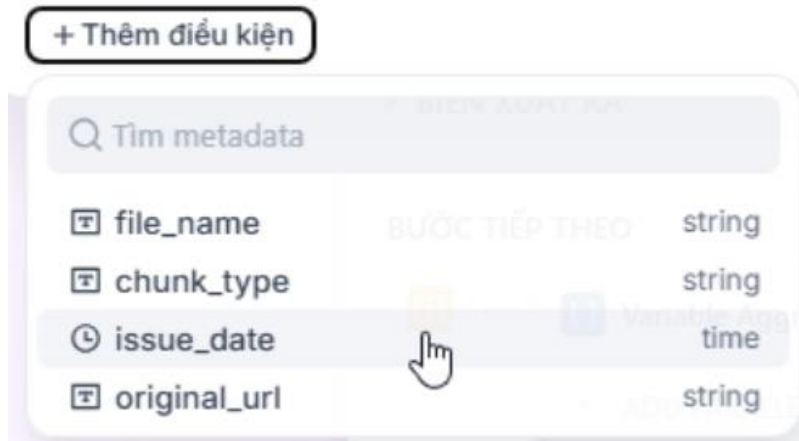


Bước 2. "Chọn "Các điều kiện" để hiển thị giao diện thêm metadata."



Bước 3. "Nhấn "Thêm điều kiện" và chọn "issue_date"."

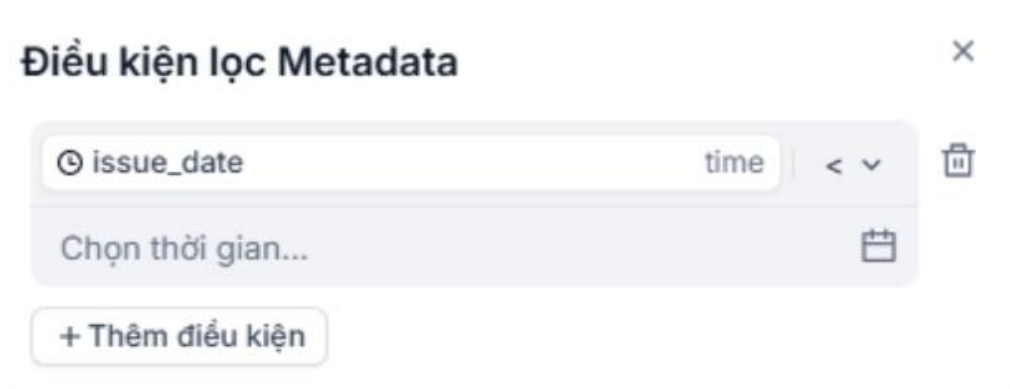
Điều kiện lọc Metadata



Chọn trường issue_date

Bước 4. Chọn điều kiện lọc và giá trị so sánh.

Điều kiện	Mô tả
Nhỏ hơn	Lọc tài liệu có issue_date nhỏ hơn giá trị chọn. Người dùng chọn giá trị thời gian trong phần "Chọn thời gian", chính xác đến phút.
Lớn hơn	Lọc tài liệu có issue_date lớn hơn giá trị chọn. Chọn giá trị thời gian tương tự.
Tồn tại	Lọc các tài liệu có lưu trường issue_date trong metadata.
Không tồn tại	Lọc các tài liệu không có trường issue_date trong metadata.



Cấu hình điều kiện lọc theo thời gian

2.4. Cập nhật Node LLM: Thinking Mode & Completion / Chat Mode

Phiên bản mới bổ sung hai cải tiến quan trọng cho Node LLM trong Task:

Tính năng	Mô tả
Completion / Chat Mode	Node LLM hiển thị giao diện cấu hình prompt phù hợp với từng loại model (Completion hoặc Chat).
Thinking Mode	Cho phép kích hoạt chế độ tư duy sâu (reasoning) cho các model hỗ trợ, với tùy chọn tách kết quả thinking ra riêng.

Hai tính năng này giúp:

- Tối ưu cách sử dụng Node LLM theo đúng đặc tính của từng model.
- Cải thiện độ ổn định và chất lượng câu trả lời.
- Hỗ trợ các luồng hội thoại phức tạp với Chat mode.
- Khai thác năng lực lập luận của các model hỗ trợ reasoning.

2.4.1. Chế độ Completion / Chat Mode

Node LLM hiển thị giao diện cấu hình prompt khác nhau tùy theo chế độ hoạt động của model đã chọn. Chế độ này được cấu hình tại màn hình Quản lý Mô Hình (trường Chế độ Completion).

Chỉnh sửa GLM-4.7
✕

Cập nhật mô hình đã chọn

Tên mô hình *

API Key *

API Endpoint *

Chế độ Completion *

Chat
▼

Độ dài ngữ cảnh *

Số token đầu ra tối đa *

Gọi hàm *

Not supported
▼

Hỗ trợ streaming *

Use
▼

Hỗ trợ thị giác *

Not supported
▼

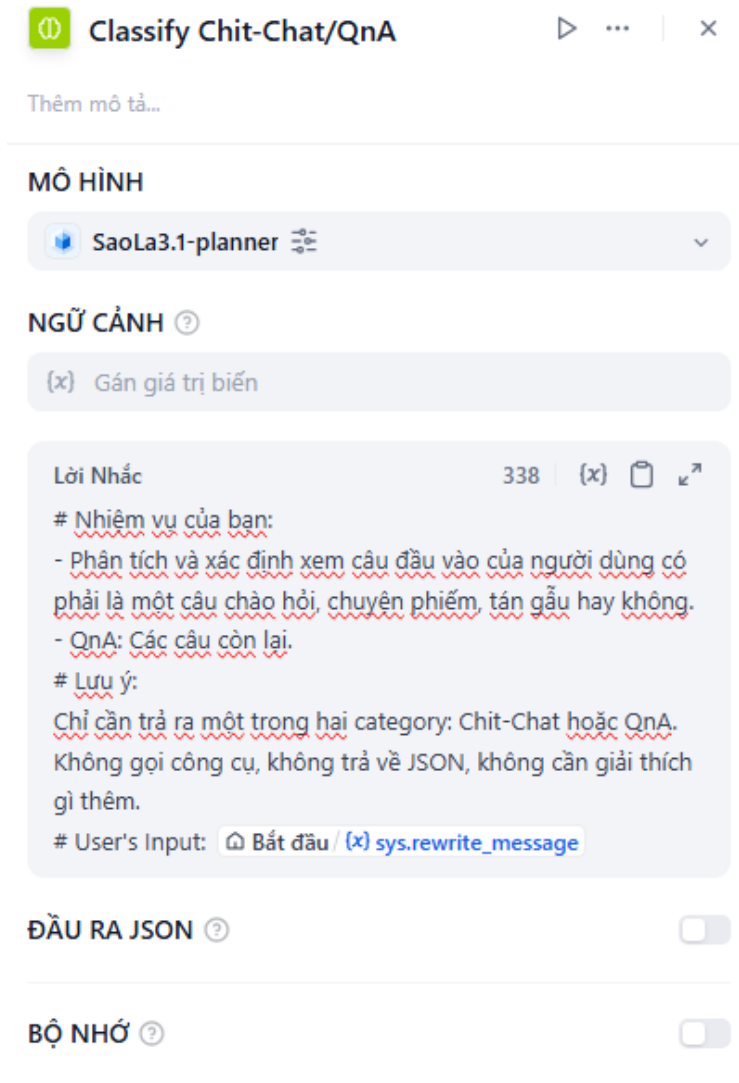
Hủy bỏ
Lưu

Hình 1 – Cấu hình Chế độ Completion tại Quản lý Mô Hình

Completion Mode

Khi model được cấu hình là Completion mode, toàn bộ prompt được nhập trong một ô Lời nhắc duy nhất (giống giao diện cũ).

- **Lời nhắc:** Nhập toàn bộ nội dung hướng dẫn và câu hỏi vào đây.



Hình 2 – Node LLM với Completion mode

Chat Mode



Khi model được cấu hình là Chat mode, giao diện chia thành các mục tin nhắn riêng biệt:

Mục	Bắt buộc	Mô tả
System message	Có (mặc định)	Xác định vai trò của AI, mục tiêu và quy tắc trả lời (tone, format, giới hạn). Kiểm soát cách AI phản hồi nhất quán.
User message	Không	Chứa yêu cầu, câu hỏi hoặc thông tin đầu vào từ người dùng.
Assistant message	Không	Lưu lại phản hồi trước đó của AI. Dùng để duy trì ngữ cảnh hoặc mô phỏng lịch sử hội thoại.

LLM 2 ▶ ... | ✕



Thêm mô tả...

MÔ HÌNH

 **Llama-3.3-70B-Instruct** FPT.AI  ▼

NGỮ CẢNH ?

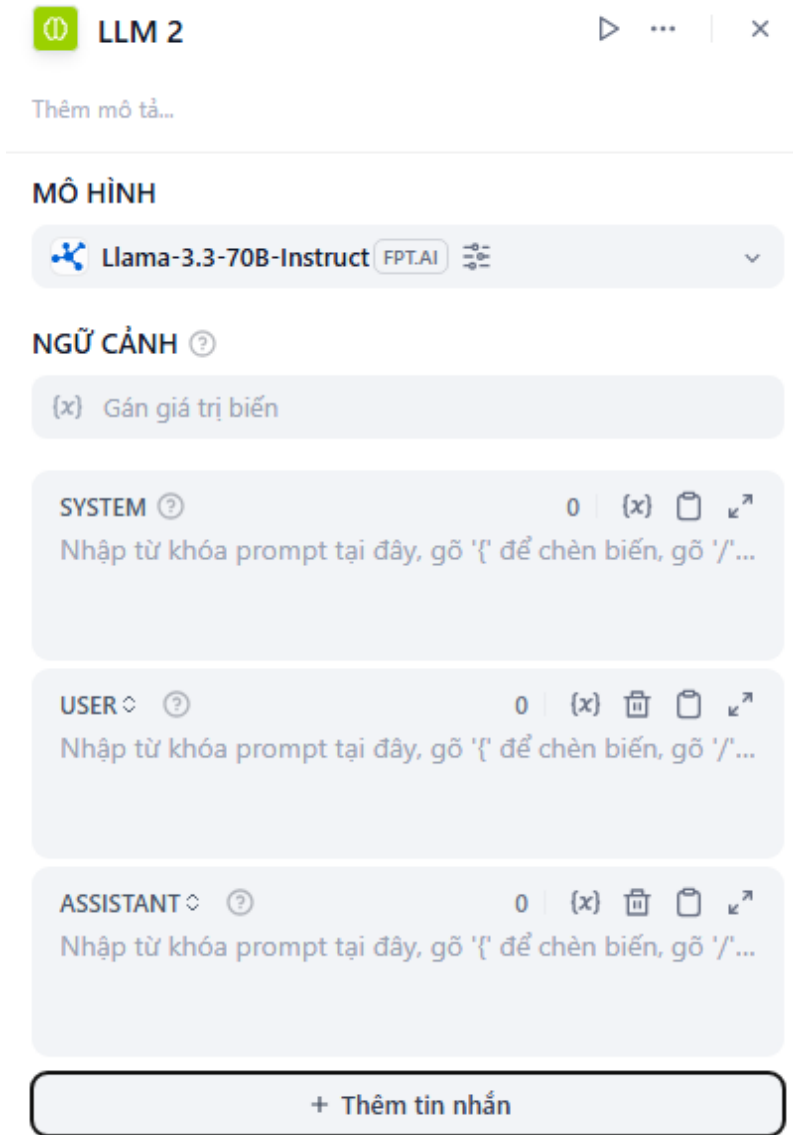
{x} Gán giá trị biến

SYSTEM ? 0 {x}  

Nhập từ khóa prompt tại đây, gõ '{' để chèn biến, gõ '/' ...

+ Thêm tin nhắn

Hình 3 – Node LLM Chat mode (mặc định: System message)

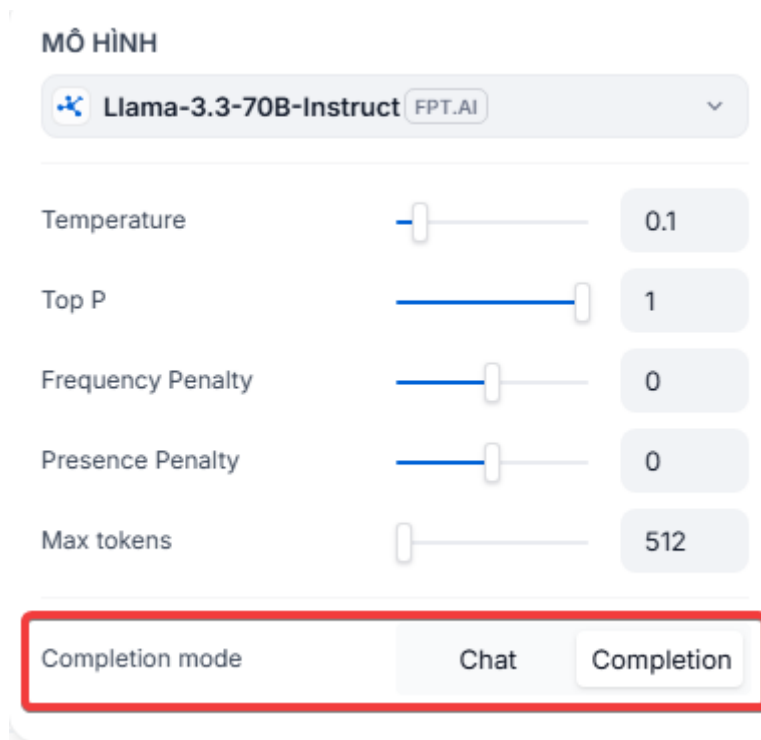


Hình 4 – Node LLM Chat mode (đã thêm User & Assistant message)

i Nhấn + Thêm tin nhắn để thêm User message hoặc Assistant message. Mặc định khi tạo mới Node LLM với Chat mode, chỉ có System message.

Xử lý Node LLM cũ (tạo trước 31/03/2026)

Với các Node LLM được tạo trước ngày 31/03/2026, hệ thống cho phép chuyển đổi thủ công giữa 2 chế độ để tương thích với model:



Hình 5 – Nút chuyển đổi chế độ trên Node LLM cũ

! Sau khi chuyển đổi chế độ trên Node LLM cũ, cấu hình prompt hiện tại có thể bị thay đổi. Hãy kiểm tra lại nội dung prompt trước khi lưu.

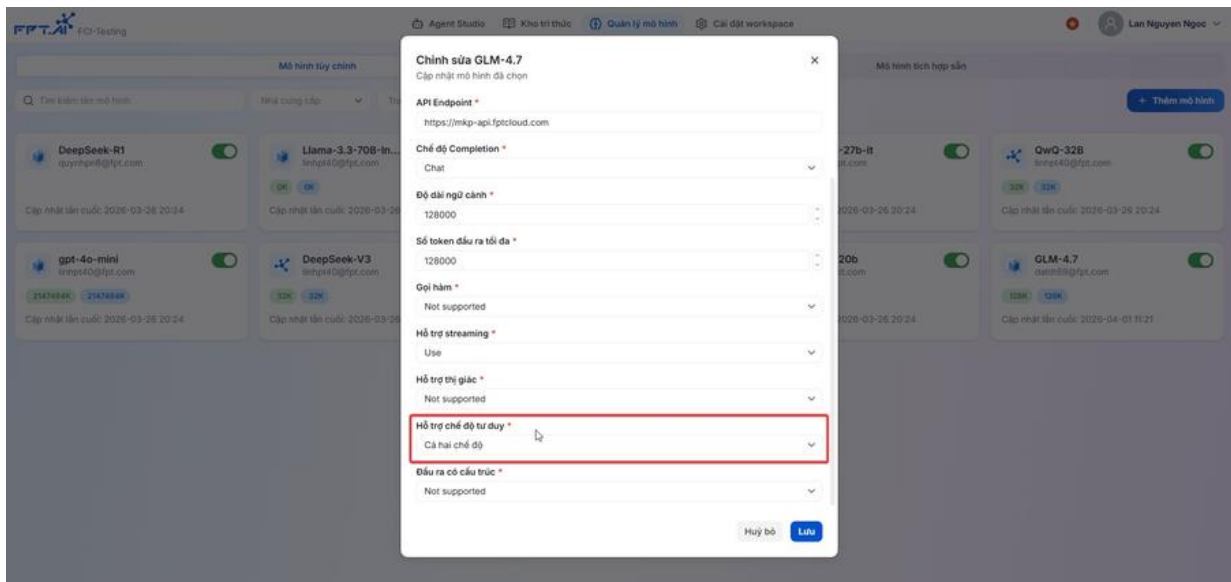
2.4.2. Thinking Mode (Chế độ tư duy)

Thinking Mode cho phép model thực hiện lập luận nội bộ (reasoning) trước khi trả lời. Tính năng này chỉ khả dụng trên các model hỗ trợ reasoning.

Cấu hình Thinking Support tại Quản lý Mô Hình

Trong form chỉnh sửa model (dạng OpenAI-API-Compatible), trường Hỗ trợ chế độ tư duy cho phép chọn một trong ba giá trị:

Giá trị	Ý nghĩa
Only non-thinking mode	Model chỉ chạy luồng thông thường, không có reasoning. (Mặc định khi tạo mới)
Only thinking mode	Model luôn thực hiện reasoning. Người dùng không cần cấu hình thêm tại Node.
Both modes	Model hỗ trợ cả hai chế độ. Người dùng tự bật/tắt Thinking mode tại Node LLM.



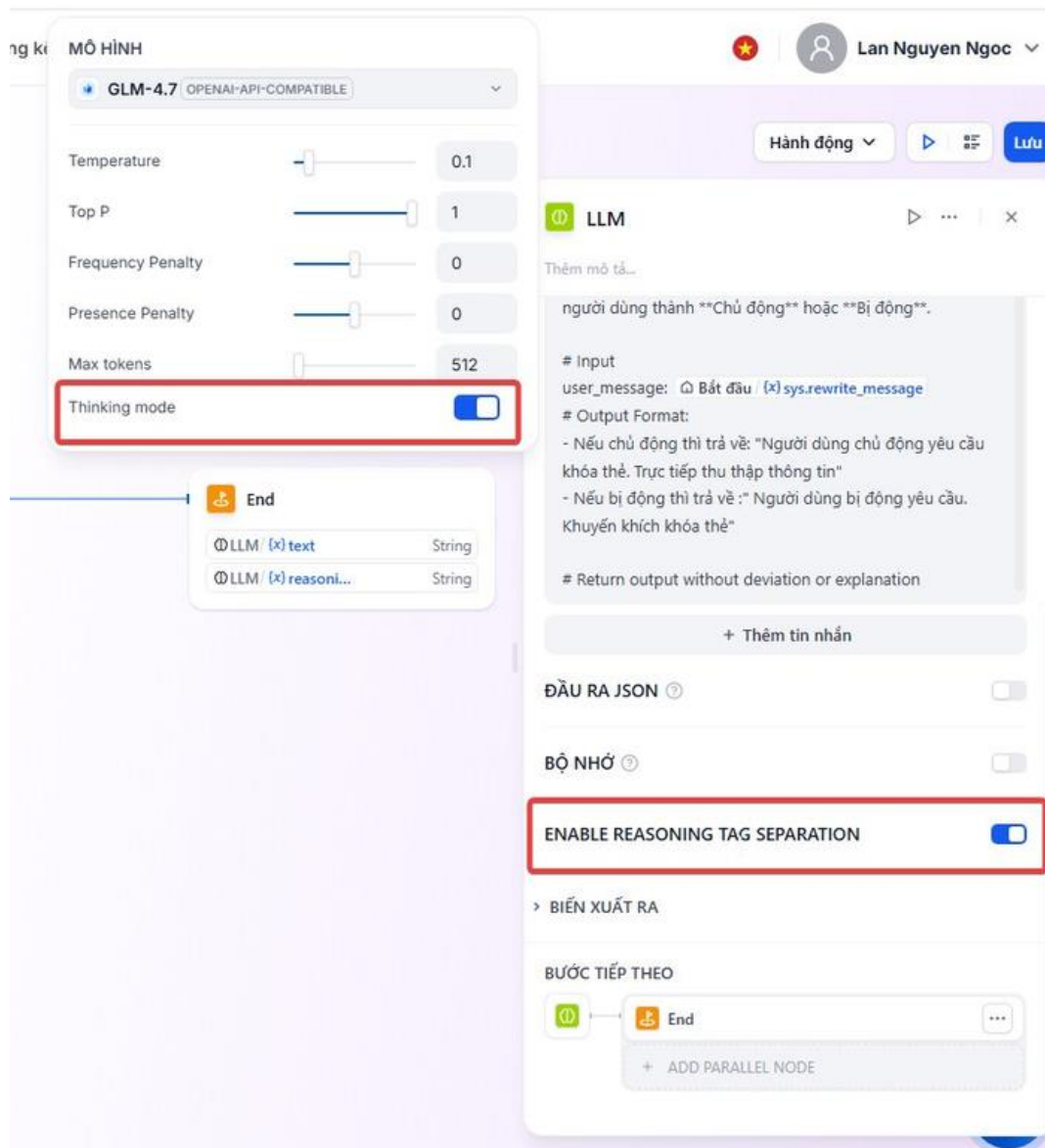
Hình 6 – Trường Hỗ trợ chế độ tư duy trong cấu hình model

- Giá trị mặc định khi tạo mới model là Only non-thinking mode. Chỉ chọn Both modes khi model thực sự hỗ trợ reasoning ở cả hai chế độ.

Tham số Thinking mode tại Node LLM

Khi chọn model trong Task, tham số Thinking mode xuất hiện hoặc ẩn tùy theo cấu hình model:

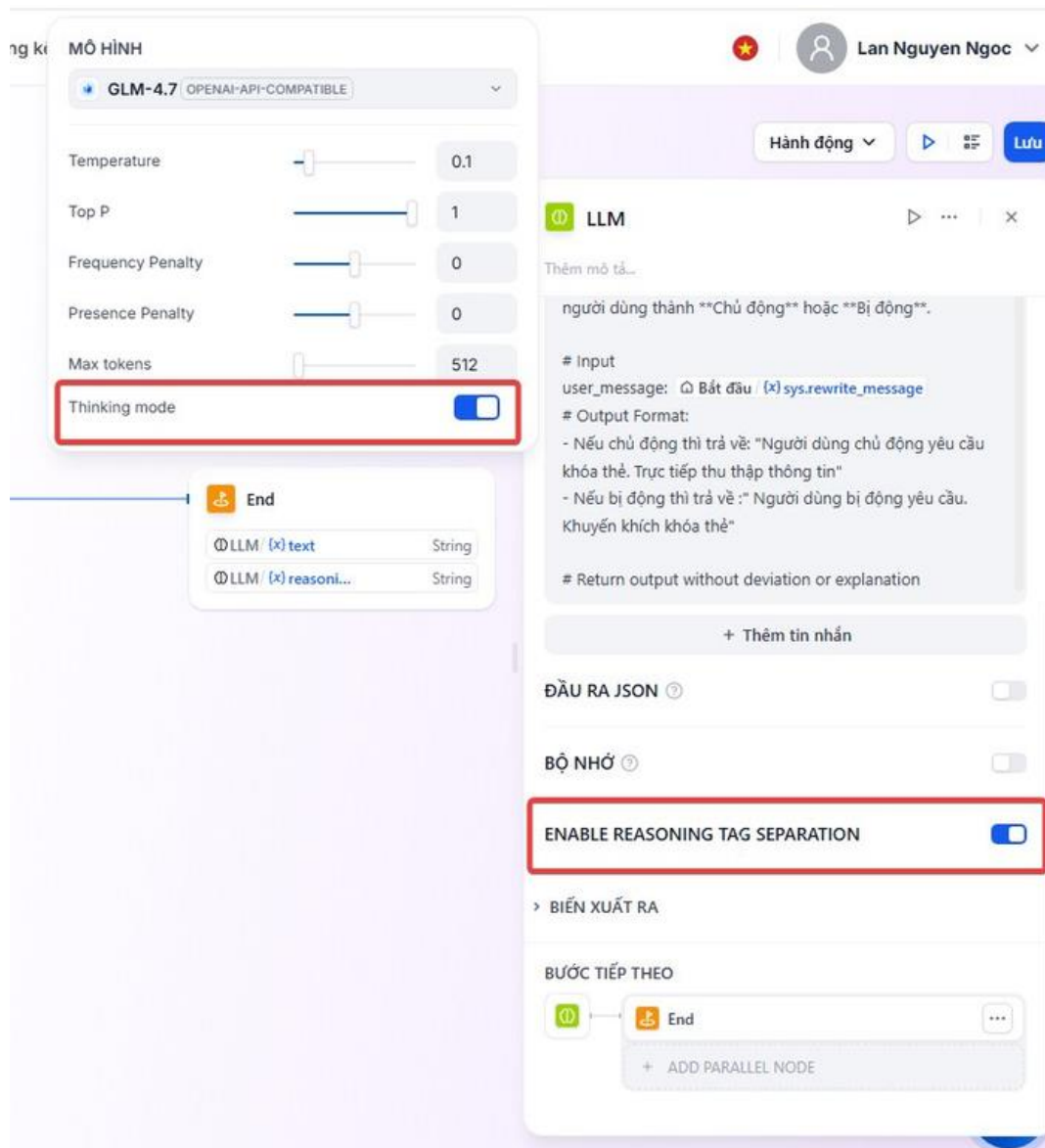
Cấu hình model	Hiển thị tại Node	Hành vi
Only non-thinking mode	Ẩn	Luôn chạy luồng thông thường.
Only thinking mode	Ẩn	Luôn chạy reasoning. Không cần người dùng thao tác.
Both modes	Hiển thị toggle Bật/Tắt (mặc định: Tắt)	Người dùng tự quyết định bật/tắt reasoning.



Hình 7 – Tham số Thinking mode tại Node LLM (model cấu hình Both modes)

Enable Reasoning Tag Separation

Khi Thinking mode đang BẬT (hoặc model cấu hình là Only thinking mode), tùy chọn bổ sung xuất hiện bên dưới mục MEMORY:



Hình 8 – Enable Reasoning Tag Separation tại Node LLM

i Bật Enable Reasoning Tag Separation khi bạn muốn xử lý riêng phần reasoning của model, ví dụ: log lại nội dung suy luận hoặc hiển thị bước suy nghĩ cho người dùng cuối.

Biến xuất ra của Node LLM khi dùng Thinking Mode

Khi Thinking mode được kích hoạt, Node LLM trả về thêm biến `reasoning_content` trong mục BIẾN XUẤT RA:

- **text** – Nội dung câu trả lời chính.
- **reasoning_content** – Nội dung lập luận của model (có giá trị khi Enable Reasoning Tag Separation = ON; bằng null khi OFF).

⚠ Biến `reasoning_content` có thể được sử dụng trong các bước tiếp theo của Task để kiểm tra logic suy luận hoặc ghi log. Nếu không cần, để mặc định OFF.