**GIỚI THIỆU RASA PRO**

# **1. Giới thiệu chung**

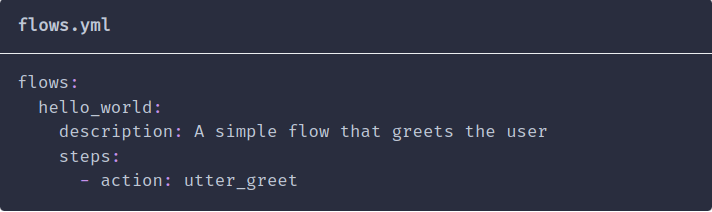
* Rasa Pro là phiên bản mở rộng của Rasa Open Source, tích hợp với **CALM** (Conversational AI with Language Models), mang đến một giải pháp toàn diện cho việc phát triển các trợ lý ảo, kết hợp với các tính năng phân tích, bảo mật giúp doanh nghiệp giám sát, theo dõi, và phân tích hiệu suất của các trợ lý ảo mà họ phát triển.
* CALM là một phương pháp xây dựng AI hội thoại dựa trên Mô hình Ngôn ngữ Lớn (LLM), được phát triển bởi Rasa sau nhiều năm hỗ trợ các đội ngũ doanh nghiệp tạo ra các trợ lý tương tác với khách hàng. Với CALM, bạn sẽ có được sự kết hợp hoàn hảo giữa tốc độ, tính tổng quát và độ lưu loát của các LLM thế hệ mới, cùng với tính bền vững, độ tin cậy, khả năng kiểm soát và sửa lỗi chatbot.

# **2. Các thành phần cơ bản**

Trong Rasa Pro bao gồm ba thành phần chính:

* **Business Logic (Logic Kinh Doanh):** được triển khai dưới dạng các **flows** (luồng hội thoại). Một flow mô tả quy trình mà trợ lý AI của bạn có thể xử lý, bao gồm thông tin cần lấy từ người dùng, dữ liệu cần truy xuất từ API hoặc cơ sở dữ liệu, và logic phân nhánh dựa trên thông tin thu thập được. Một flow chỉ mô tả logic mà trợ lý của bạn sẽ làm theo, chứ không phải tất cả các kịch bản hội thoại có thể xảy ra.

**Vd:**

****- Tên Flow: “hello\_world”

- Description: mô tả cách hoạt động của flow

- Steps: các hành động cụ thể được định nghĩa trước mà trợ lý AI thực hiện trong quá trình xử lý một yêu cầu hoặc tương tác với người dùng. Ở ví dụ này, trợ lý AI sẽ thực hiện hành động “utter\_greet” , tức là gửi một lời chào đến người dùng.

* **Dialogue Understanding (Hiểu Hội Thoại):** giúpgiải thích ý định của người dùng. Sử dụng LLM tạo ra các lệnh để điều hướng cuộc trò chuyện theo logic kinh doanh và ngữ cảnh. Các lệnh này bao gồm bắt đầu và kết thúc các flow, lấy các thông tin cần thiết từ người dùng, …
* **Automatic Conversation Repair (Sửa Chữa Hội Thoại Tự Động):** Sửa chữa hội thoại tự động xử lý tất cả các tình huống khi cuộc hội thoại diễn ra không theo kịch bản đã định.

**Vd:**- Trợ lý của bạn yêu cầu một địa chỉ email, nhưng người dùng lại nói điều gì đó khác.

- Người dùng cắt ngang flow hiện tại và chuyển sang chủ đề khác.

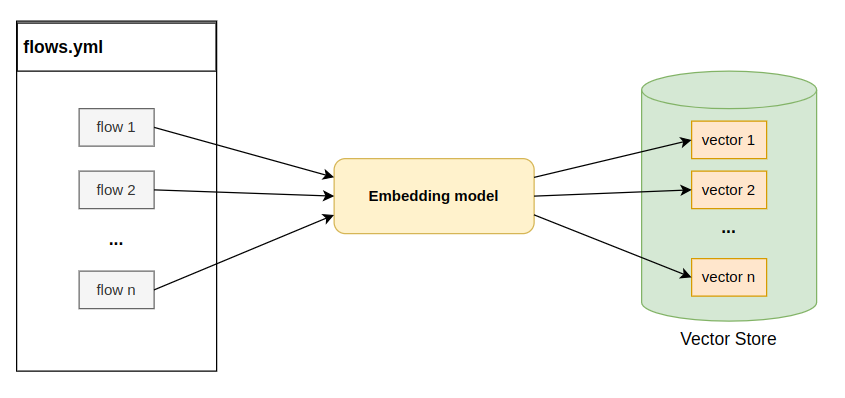
- Người dùng thay đổi ý kiến về điều gì đó họ đã nói trước đó.

Những trường hợp này và nhiều tình huống khác sẽ được xử lý tự động, và bạn hoàn toàn có thể tùy chỉnh.

# **3. Luồng hoạt động**

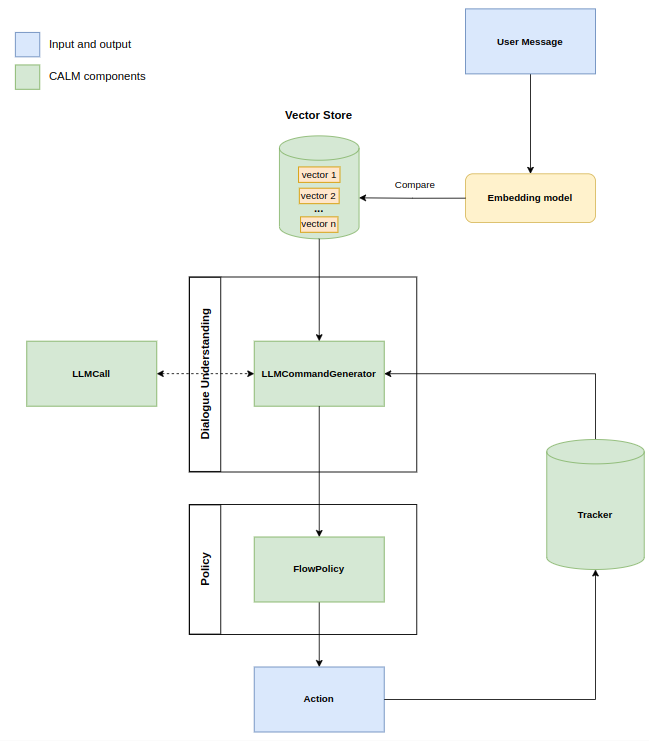
**Training**

* Trong quá trình training, tất cả các flow được chuyển đổi thành vector thông qua embedding model và được lưu trữ trong vector store.



**Inference**

* Trong quá trình inference, ngữ cảnh của cuộc hội thoại hiện tại sẽ được chuyển đổi thành một vector và so sánh với các flow đã được lưu trữ trong vector store. Quá trình so sánh này giúp xác định những flow có độ tương đồng cao nhất với ngữ cảnh hiện tại của cuộc hội thoại. Những flow phù hợp nhất sau đó sẽ được đưa vào prompt của LLMCommandGenerator**.** Điều này đảm bảo rằng các lệnh (command) được tạo ra sẽ dựa trên những flow phù hợp và liên quan nhất, giúp trợ lý AI phản hồi một cách chính xác và hiệu quả, ngay cả khi ngữ cảnh cuộc hội thoại phức tạp và đa dạng.

****

* Tóm tắt luồng hoạt động của quá trình inference:

1. Tin nhắn người dùng được chuyển qua Embedding model để tạo ra vector.
2. Vector này được so sánh với các vector trong Vector Store.
3. Những flow phù hợp nhất sau đó cùng với tin nhắn người dùng sẽ được đưa vào prompt của LLMCommandGenerator**.**
4. LLMCall sau đó xử lý prompt để tạo ra các lệnh.
5. Kết quả được chuyển đến FlowPolicy để quyết định hành động (Action) phản hồi người dùng.
6. Thông tin về hành động được ghi vào Tracker.
7. Tracker cung cấp dữ liệu ngữ cảnh cho LLMCommandGenerator trong các lượt tương tác tiếp theo.

# **4. Lời kết**

* Rasa Pro không chỉ đơn thuần là một công cụ nâng cấp từ Rasa Open Source, mà còn mang đến một bước tiến mới trong việc xây dựng các trợ lý AI mạnh mẽ và thông minh hơn. Với các tính năng tiên tiến như CALM, khả năng phân tích chuyên sâu, bảo mật vượt trội, và khả năng giám sát toàn diện, Rasa Pro cung cấp cho các doanh nghiệp một giải pháp toàn diện để nâng cao trải nghiệm người dùng và tối ưu hóa quy trình tương tác với khách hàng.
* Nếu bạn đang tìm kiếm một nền tảng linh hoạt, mạnh mẽ và có khả năng tùy chỉnh cao để phát triển các trợ lý AI, Rasa Pro chắc chắn là lựa chọn lý tưởng. Sự kết hợp giữa công nghệ hiện đại và tính năng doanh nghiệp sẽ giúp bạn không chỉ xây dựng mà còn quản lý và phát triển các trợ lý AI một cách hiệu quả và an toàn.

# **5. Tài liệu tham khảo**

Rasa Pro Documentation: <https://rasa.com/docs/rasa-pro/>