

Meet Ava: the Whatsapp Agent

Lesson 1: Course Overview



MIGUEL OTERO PEDRIDO

FEB 05, 2025

SI

87

5

10

What happens when [two ML Engineers](#) with a love for sci-fi movies team up? 😊

You get **Ava**, a **Whatsapp agent** that can engage with users in a **realistic way**, inspired by the great film [Ex Machina](#). Ok, let's be real, you won't be building a full sentient robot in this project, but you will enjoy some pretty interesting Whatsapp conversations. I can assure you that! 😊

[**Check the code here!**](#) 🤖



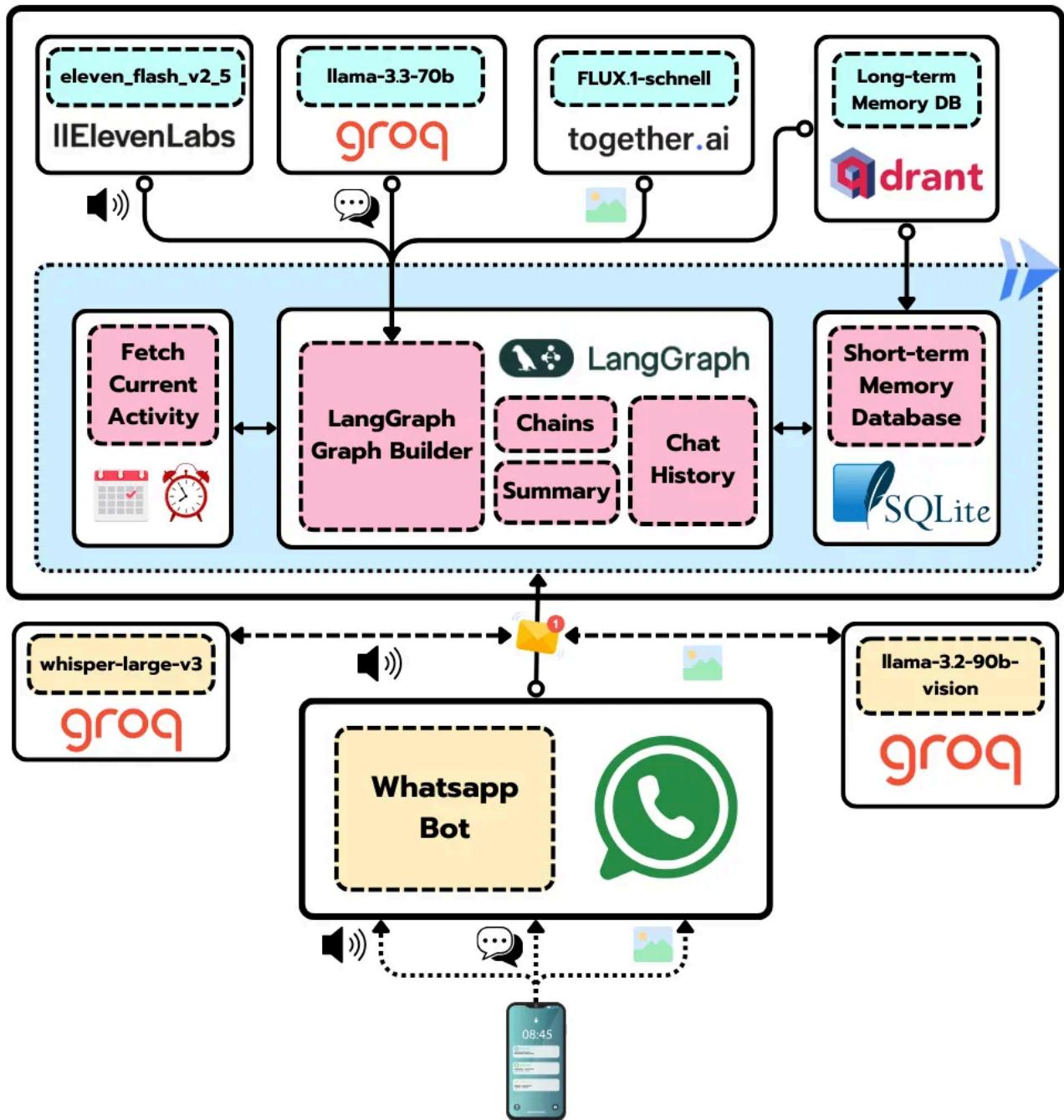
The magic of face swapping

This course is divided into **six lessons**:

 **Lesson 1:** Project overview **Lesson 2:** Ava's brain is just a graph **Lesson 3:** Unlocking Ava's memories **Lesson 4:** Giving Ava a Voice **Lesson 5:** Ava learns to see **Lesson 6:** Ava installs Whatsapp

Today, we'll start with the **first lesson** - a general introduction to the project and its core components.

Project Overview



Project Overview

Ava is a "Whatsapp Agent", meaning it will interact with you through this app. But it won't just rely on "regular" text messages, it will also **listen to your voice notes** (even if you are **one of those** people 😊) and **react to your pictures**.

And that's not all ... Ava can also respond with its own voice notes and images of whatever it's up to - yes, Ava has a life beyond talking to you, don't be such a narcissist! 😂



Jesús in Westworld mode, messing with Ava's mind

At this point, you might be wondering:

What kind of system have we implemented to handle multimodal inputs / outputs coherently?

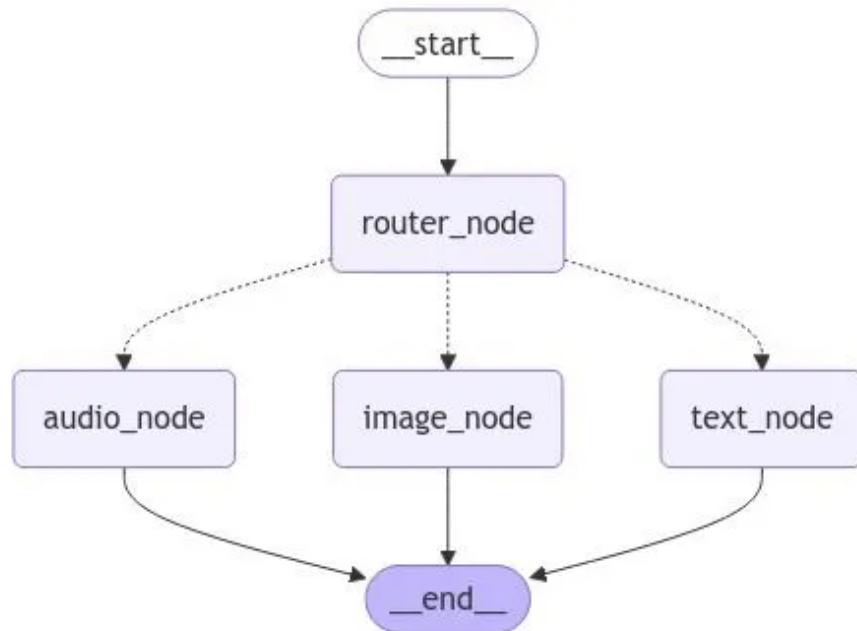
The short answer: Ava's brain is **just a graph** ... a [LangGraph](#) (sorry, I couldn't resist).

Ava's Graph

Your brain is made up of neurons, right? Well, Ava's brain is made up of **LangGraph nodes** and **edges** - one for the processing images, another for listening to your voice, another for fetching relevant memories, and so on.

At its core, **Ava is simply a graph with a state**. This state maintains all the key details of the conversation, including shared information (text, audio or images), current activities, and contextual information.

This is exactly what we'll explore in **Lesson 2**, where you'll learn how **LangGraph** can be used to build agentic design architectures, such as **the router**.



Ava will determine the type of output based on your input

Ava's memory

An Agent **without memory** is like talking to the **main character of "Memento"** (if you haven't seen that film... seriously, what are you doing with your life?).

Ava has **two types of memory**:

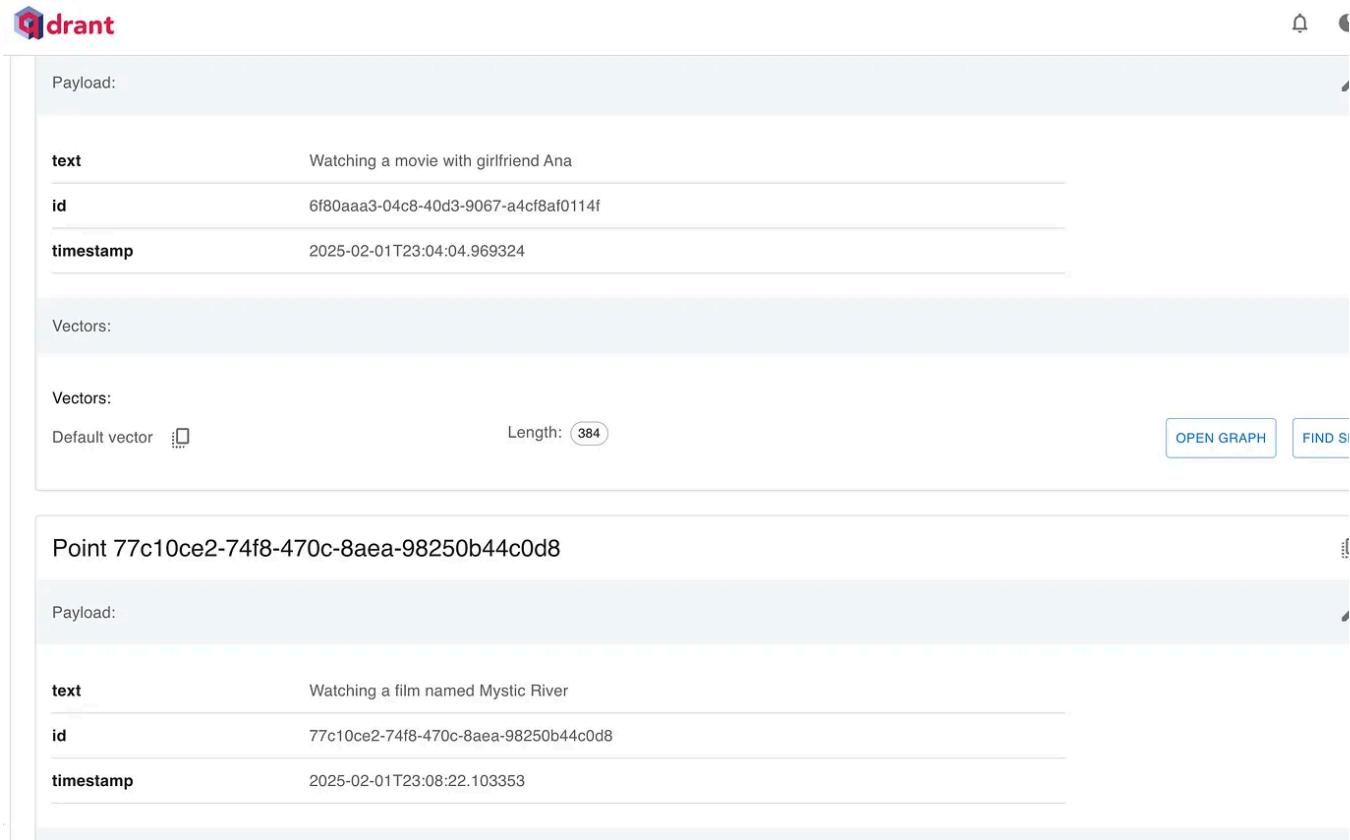
◆ Short term memory

The usual - it stores the sequence of messages to maintain conversation context. In our case, we save this sequence in [SQLite](#) (we are also storing a summary of the conversation, but that's for future lessons 😊).

◆ Long term memory

When you meet someone, you don't remember **everything** they say; you retain only the **key details**, like their **name**, **profession**, or **where they're from**, right?. That's exactly what we wanted to replicate with [Qdrant](#) - extracting relevant information from the conversation and storing it as embeddings.

Don't worry because we'll cover the memory modules in [Lesson 3](#).



The screenshot shows the Qdrant application interface. It displays two conversation payloads with their corresponding vectors.

Payload 1:

text	Watching a movie with girlfriend Ana
id	6f80aaa3-04c8-40d3-9067-a4cf8af0114f
timestamp	2025-02-01T23:04:04.969324

Vectors:

- Default vector: Length: 384
- Buttons: OPEN GRAPH, FIND SIMILAR

Payload 2:

text	Watching a film named Mystic River
id	77c10ce2-74f8-470c-8aea-98250b44c0d8
timestamp	2025-02-01T23:08:22.103353

Capturing relevant facts about the conversation (e.g. watching Mystic River with my girlfriend)

◆ Ava's senses

Real Whatsapp conversations aren't limited to just text. Think about it - do you remember the last cringe GIF your mom sent you last week? Or that neverending audio note from your high school friend? Exactly. We need both images and audio.

To make this possible, we've selected the following tools.

◆ Text

Both Jesús and I are [Groq](#) fans (if you chat with Ava, ask about its job, you might be surprised). That's why we are using **Groq models** for all text generation. Specifically, we've chosen **llama-3.3-70b-versatile** as our core LLM.

◆ Images

The image module handles **two tasks: processing user images and generating ones** (take a look at the image below).

- For **image “understanding”**, we're using Groq's **llama-3.2-90b-vision-preview**.
- For **image generation**, **black-forest-labs/FLUX.1-schnell-Free** using [Toge AI](#).

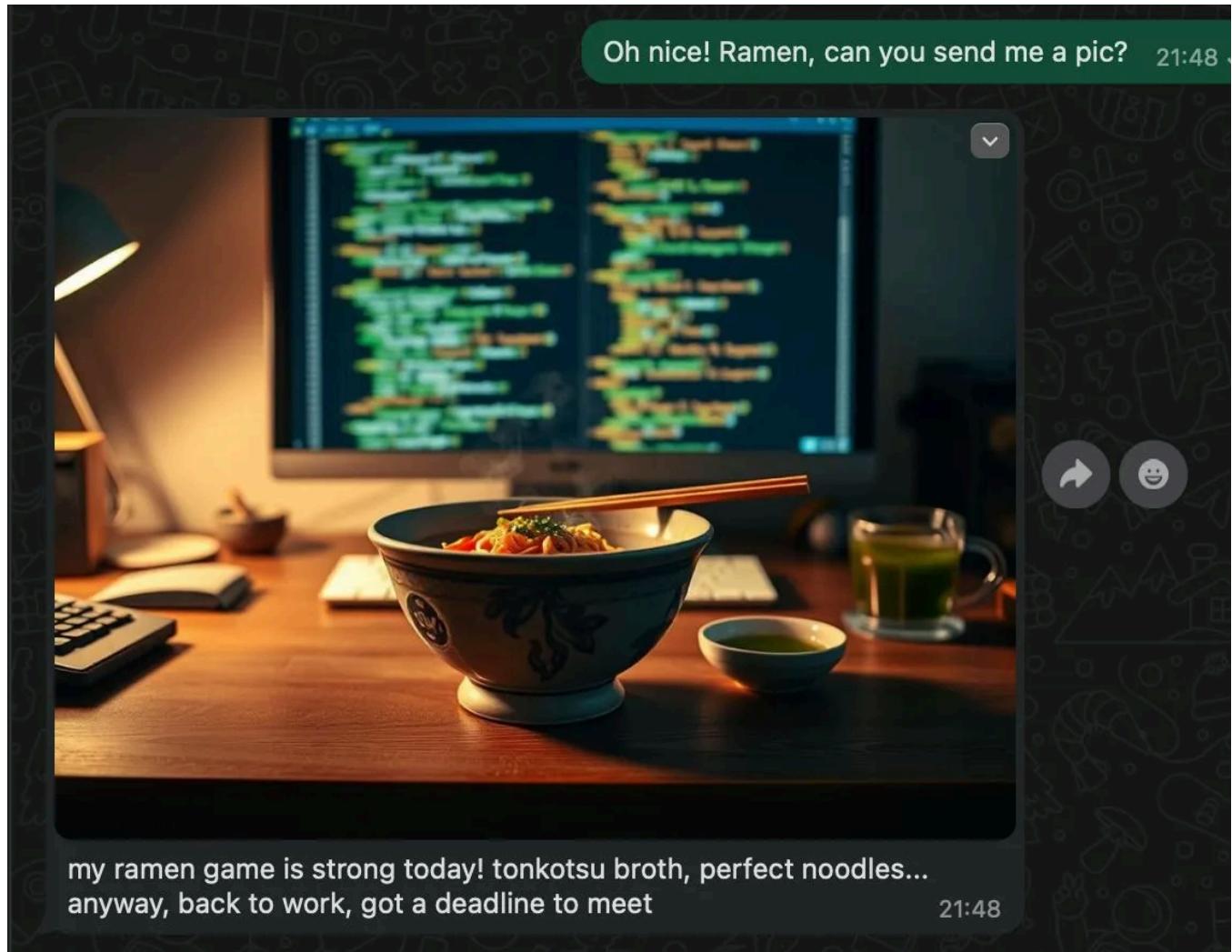
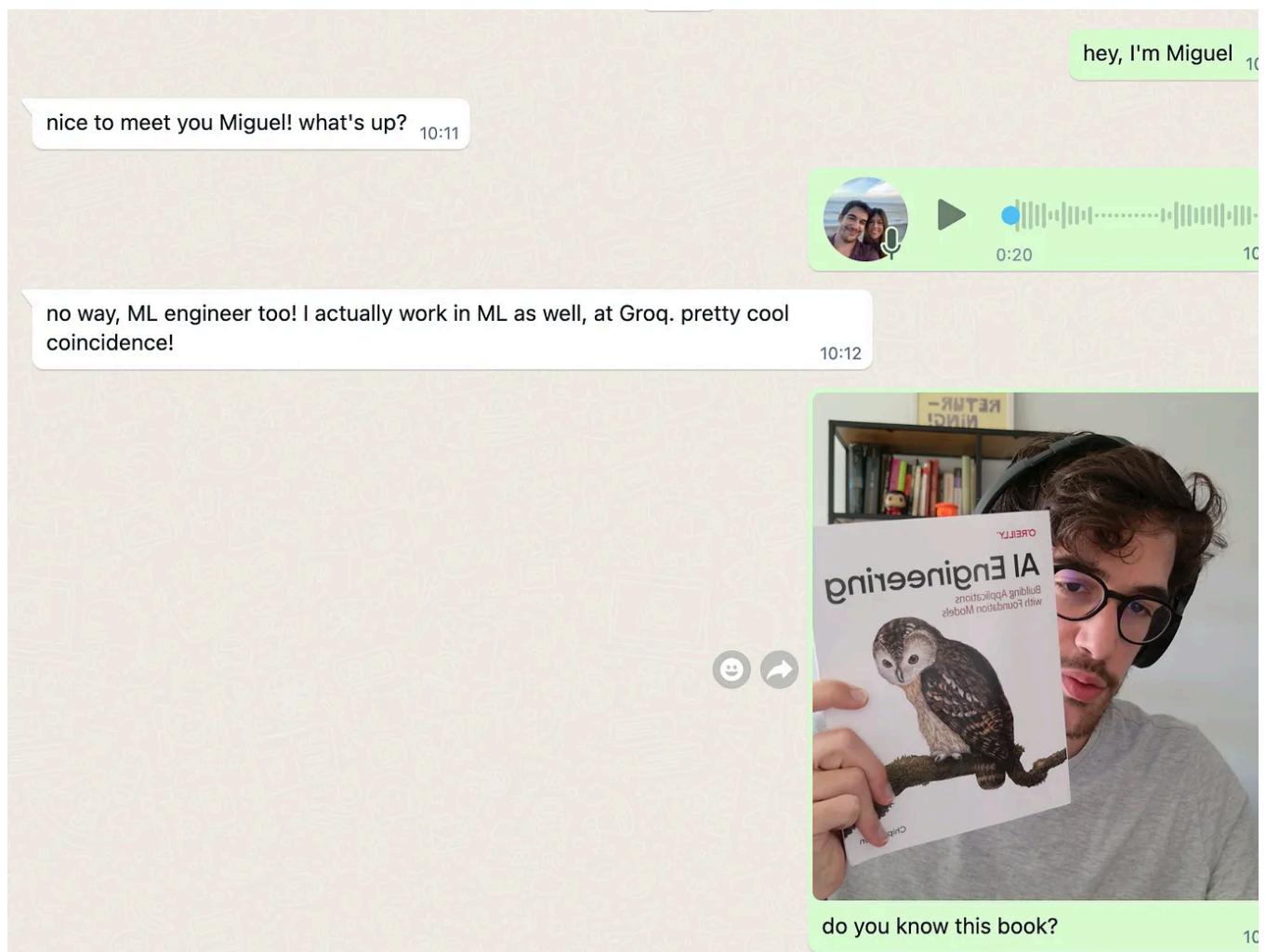


Image generation example. Turns out Ava loves ramen

◆ Audio

The audio module needs to take care of **TTS (Text-To-Speech)** and **STT (Speech To-Text)**.

- For **TTS**, we are using [Elevenlabs](#) voices.
- For **STT**, [whisper-large-v3-turbo](#) from [Groq](#).



Ava listens to my voice note, where I'm introducing myself as an ML Engineer

We'll cover the **audio module** in **Lesson 4** and the **image module** in **Lesson 5!**

And that's all for today! As you can see, this is a very complete course, so we hope you're excited to get started with it! Remember, **Lesson 2** will be available next

Wednesday, February 12th. Every lesson (including this one) comes with a complementary video on [Jesús Copados' YouTube channel](#).

We strongly recommend exploring both resources (written lessons and video lessons) to maximize your learning experience! 😊

WhatsApp AI Agent Tutorial 1: Meet Ava | What is an AI Agent?



📱 Happy Whatsapping! 📱

Thanks for reading The Neural Maze! Subscribe for free to receive new posts and support my work.



87 Likes • 10 Restacks

← Previous

Next →

Discussion about this post

[Comments](#) [Restacks](#)

Write a comment...



A Pox on Both Your Houses. 13 Feb

Liked by Miguel Otero Pedrido

OK, you whetted my appetite....but now I am disappointed. It is now Feb 13. Did I miss the next installment somewhere?

LIKE (1) REPLY

1 reply by Miguel Otero Pedrido



Serghei Dragan 7 Feb

Liked by Miguel Otero Pedrido

Good one and Funny Miguel 😊 ! Can't wait to see what's coming next.

LIKE (1) REPLY

1 reply by Miguel Otero Pedrido

3 more comments...

