



**COMMUNITY DAY**

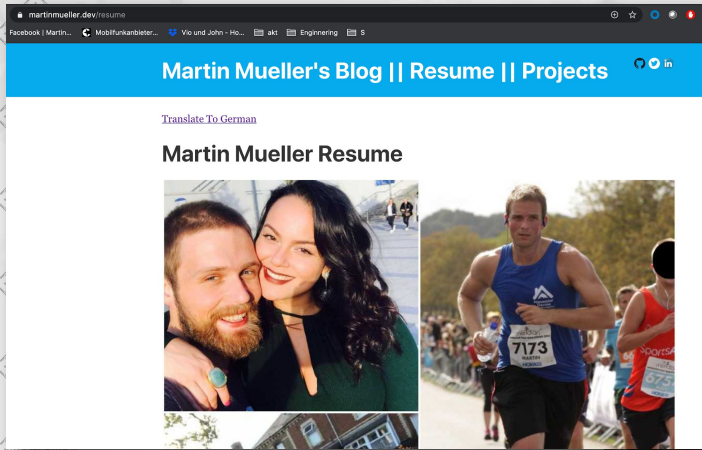
# Run your AI LangChain on AWS

Martin Müller



# Who am I?

MartinMueller.dev/Resume



linkedin.com/in/martinmueller88



mmuller88



martinmueller.dev



office@martinmueller.dev



## Summary

office@martinmueller.dev

Passionate Full Stack Developer. I like to have a broad collection of skills. From developing the backend and frontend applications written in Java, JavaScript, or any other language, to developing a cost-optimized cloud infrastructure solution. An enthusiastic agile player with strong communication skills. Looking for taking on new challenges and learning new technologies.

"What I don't know, I will learn"

## Highlights

- AWS Professional Solution Architect certified  [pic](#)
- AWS Community Builder
- TypeScript, Java, JavaScript, Angular, React
- AWS Lambda, EC2, CDK, DynamoDB and many more
- Maven, Gradle, SQL, ElasticSearch
- Docker, Kubernetes, Helm
- Spring, GitHub (Actions), Travis, GitLab, BitBucket

## Experience

AWS Community Builder  
Mar 2022 - current

AWS Consultant  
Nov 2020 - current

- Helping startups into the AWS Cloud
- Leveraging AWS CDK to manage AWS resources

Senior Cloud Architect - [www.kreuzwerker.de](http://www.kreuzwerker.de), GER  
Jan 2022 - May 2022

- Together with the customer building awesome AWS cloud solutions

## Co-Founder and CTO at [hallocasa.com](https://hallocasa.com)

HalloCasa

Propiedades Agentes Precios [Iniciar Sesión](#) [Español](#)

Anna Schmidt

El Creador de Página Web Inmobiliaria más Fácil

EMPIEZA YA!

Mira el Video: Solo necesitas 5 Minutos..

Book an Appointment! +54234234325 vewihi5319@iucake.com

Juan



# COMMUNITY DAY

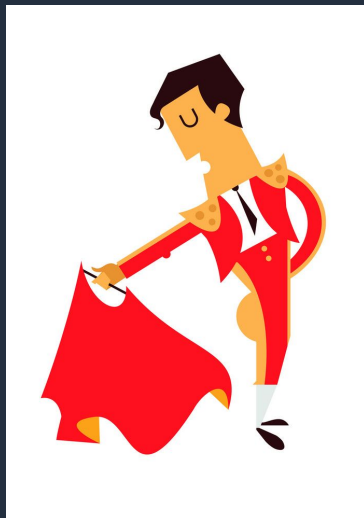
## Agenda

**01**

**What is  
LangChain?**

**03**

**LangChain  
Demo**



**05**

**Reflecting & QA**

**02**

**Why run it  
on AWS?**

**04**

**Local LLM**

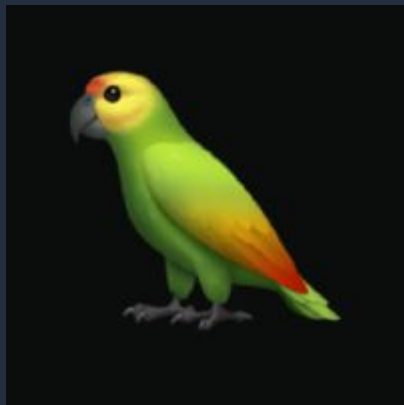
# Disclaimer!

- I'm not a Data Engineering or ML expert! I'm an AWS DevOps Full-stack engineer!
- Still looking for nice use cases (clients) to solve problems with it :)





**COMMUNITY DAY**



# **01 What is LangChain?**

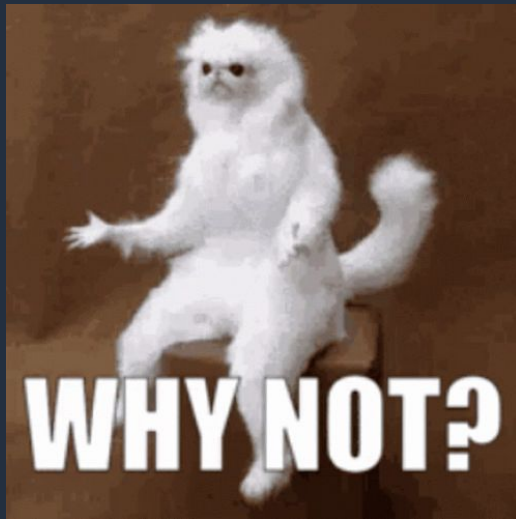


# What is LangChain?

- Framework / Library / SDK for interacting with (multiple) LLMs **programmatically**
  - I love the abstraction on top of like the OpenAI API
  - Extends the possibilities you can do with LLMs by a lot!
- Python library <https://github.com/hwchase17/langchain>
- TypeScript library <https://github.com/hwchase17/langchainjs>
- Has cool features like:
  - **Chains** - Sequence of LLM calls. Even across multiple LLMs.
  - **Memory** - Persisting state (like conversations) between calls of a chain/agent.
  - **Agents** - Involve an LLM making decisions about which Actions to take.
  - **Vector Machines** - “Train” the LangChain LLM with your specific data



# COMMUNITY DAY

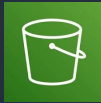


## 02 Why run it on AWS?

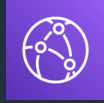


# Why running it on AWS?

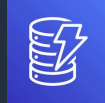
- Your laptop isn't production ready! You will need an external service like AWS to run your AI app.
- AWS offers an amazing set of Serverless services you easy can integrate like:



S3



Cloudfront



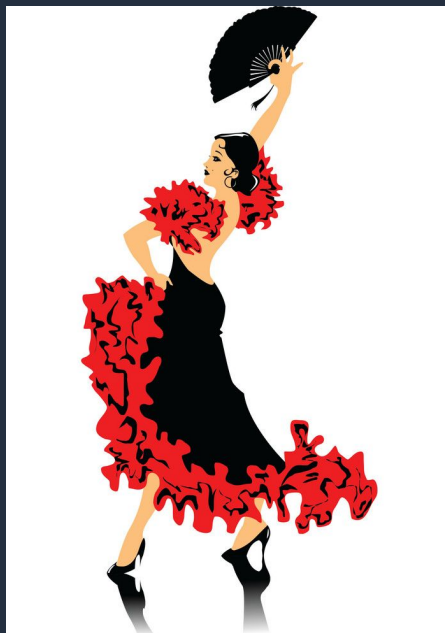
DynamoDB



RDS



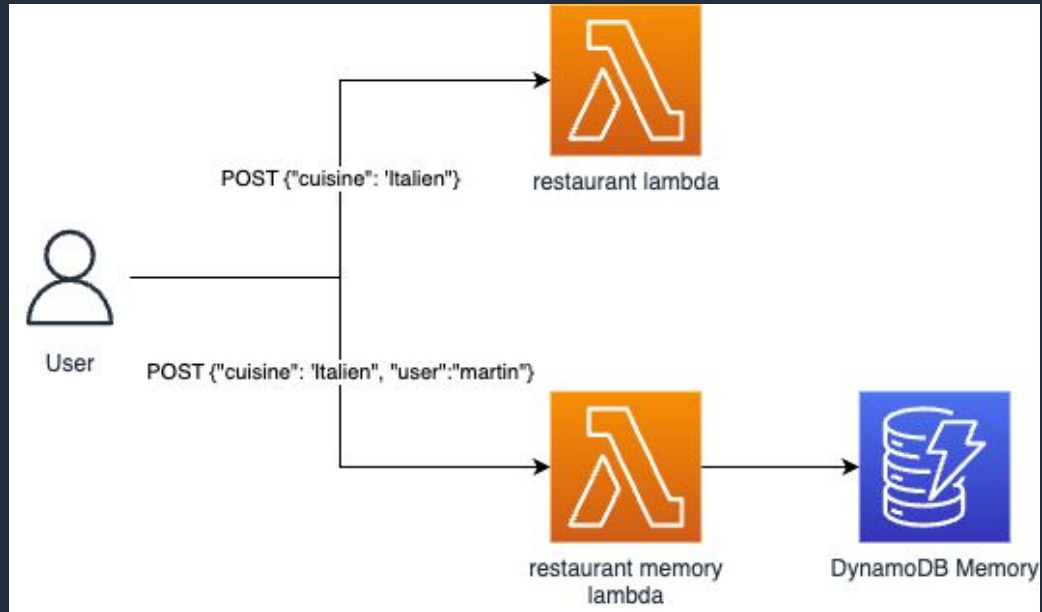
**COMMUNITY DAY**



## **03** LangChain Demo

# LangChain Demo

<https://github.com/mmuller88/aws-ts-langchain>



# LangChain Demo

<https://github.com/mmuller88/aws-ts-langchain>



Prompt 1: I want to open a restaurant for {cuisine} food. Suggest a fancy name for this.

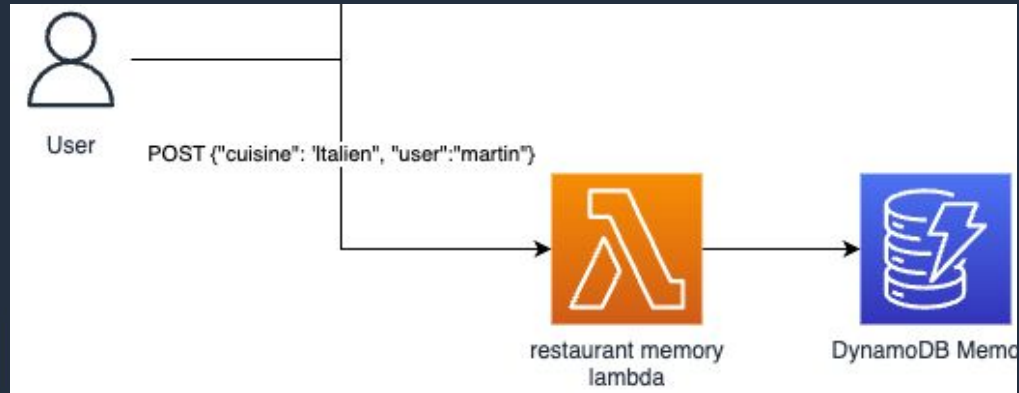
Prompt 2: Suggest some menu items for {restaurant\_name}. Return it as a comma-separated string.

Those two prompts are chained and will return the name for the restaurant and suggestions for menu items.

# LangChain Demo

<https://github.com/mmuller88/aws-ts-langchain>

Prompt: I want to open a restaurant for {cuisine} food. Suggest a fancy name for this.





# COMMUNITY DAY



## 04 Local LLM



# Local LLM

- Don't want to send your precious data to an external LLM? Try a local LLM on a Lambda.
- Download your Model on <https://huggingface.co/models>
- Why else running a local LLM?
  - Optimised on certain use cases like Llama chat or instructions for coding
  - Perhaps less hallucinating as data is more specialised
  - Cost reductions

```
• → aws-ts-langchain git:(main) x URL=https://vwoljomx7sjgdyziy5eiybk5oy0tsvrz.lambda-url.us-east-1.on.aws/prompt?cuisine=Italian
curl $URL | jq .
% Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
           %    71    71    0      0      0     0    0  --:--:--  0:11:12  --:--:--   16
{
  "returnmsg": " \"Il Piacere\" - It means \"The Pleasure\" in Italian."
}
```

(llama2-13b.Q4\_0.gguf)



# COMMUNITY DAY

Need help with your AI work loads?

Contact me :) !



## 04 QA

