

All your favorite parts of Medium are now in one sidebar for easy access.

Okay, got it

Profile

Stories

Stats

Following

Python in Plain En...

Towards AI

Stackademic

AI Advances

Artificial Intelligen...

Data Science Colle...

Google Cloud - Co...

The Pythoners

Level Up Coding

Generative AI

▼ More

AI/ML & GenAI Solutions ...



What is the role of an AI Solutions Architect in 2025?



Toby Fotherby

[Follow](#)

3 min read · Dec 19, 2024

22

...

This is a question that I am asking myself as I kick-off my next online training program called *AI Solution Architect*; a live sixteen-part program to help IT and software developers transition into AI solutions roles. I also happen to be an AI/ML Specialist Solution Architect with AWS, collaborating with customers on their journey, and considering how their roles are evolving.

The last couple of years have been pivotal for AI practitioners. The advances in tooling — LLMs, multi-modal models, no-code/low-code AI/ML, and AI as a service — has changed what is possible, viable and expected. Along with this, the role of the AI Solution Architect is changing. For those transitioning to this role and those consolidating their skills, considering the expectations, opportunities and nuances will help you move in the right direction.

The foundation of AI Solutions Architect (AI SA) role is very much that of the standard Software Solutions Architect role. For instance, looking at some of the core behaviors, there is much similarity:

- Designing and developing comprehensive solutions that align with business requirements, best practices and technical constraints
- Evaluating and selecting appropriate technologies, tools, and frameworks to build robust and scalable solutions
- Collaborating with product stakeholders and providing technical leadership to ensure effective communication and alignment
- Staying up-to-date with emerging technologies and industry trends to continuously improve and optimize solutions

Looking for differences helps kick-start a more specific definition the AI Solution Architect role. As an AI SA, central concerns include:

- Designing and developing AI solutions and patterns that align with business objectives and requirements
- Assessing and selecting appropriate AI technologies, frameworks, and tools for projects in focus
- Overseeing the integration of AI systems and patterns with existing infrastructure and applications
- Ensuring the scalability, performance, security, and maintainability of AI solutions

Attending to these concerns can be particularly challenging:

- AI technologies are evolving extremely rapidly
- Key solution aspects, such as acceptable quality, latency, and cost, often lack clarity during scoping and design
- Assessing and communicating what is possible, design options, challenges, and risks, etc.

This suggests a particular requirement for soft-skills, along with specific AI skills, to bridge the gap for stakeholders and to facilitate decision making.

It may seem as if the AI SA role needs so many skills that it must be a Unicorn: It's impossible! More reasonably, perhaps it's more like a juggler who is able to focus and deep-dive on the specifics of their domain.

All your favorite parts of Medium are now in one sidebar for easy access.



AI Solution Architect: Juggler (generated by Stable Diffusion)

How is AI Solution Architecture in 2025 different?

Changes in state of the art:

- Smarter high-end models
- Smarter light-weight model
- Agent based reasoning and application frameworks
- RAG 2.0
- Multi-modal models and solution patterns

Change in the application space:

- New domains adopting Generative AI
- Evolution of stakeholder understanding and expectations
- Application and domain specific challenges

With the scope of the AI SA practice, the scope of tools, and the scope of the business domains, all expanding, there is an increasing need for specialization with the AI SA role.

Opportunities for AI SA Specialism

Closing Thoughts

All your favorite parts of Medium are now in one sidebar for easy access.

The industry needs more AI Solution Architects. Skilled software professionals and product developers are in an excellent position to transition into AI and build on their domain and industry knowledge. With AI being more powerful and easier to apply than ever before, 2025 is a great time start. And for those who are already in the ranks as an AI SA, 2025 is excellent time for further specialization.

AI

Generative Ai Tools

Solution Architect



Published in AI/ML & GenAI Solutions Architecture

49 followers · Last published Dec 29, 2024

Follow

Patterns, tools and practices for AI solutions. This publication is designed for solutions architects and AI practitioners who are shaping the future of intelligent systems. To share expert articles, deep dives, and critical thinking on architecture and practice.



Written by Toby Fotherby

20 followers · 3 following

Follow

No responses yet



Witek ten Hove

What are your thoughts?

More from Toby Fotherby and AI/ML & GenAI Solutions Architecture

The screenshot shows the 'Model evaluation' section of the AI/ML & GenAI Solutions Architecture interface. It includes a sidebar with categories like 'Imported models', 'Playgrounds', 'Builder tools', 'Safeproofs', 'Inference and Assessment', and 'Cross-region Inference'. The main area has tabs for 'Process' and 'Kubernetes tasks'. Under 'Process', there's a 'Model evaluation' section with a 'Create' button and a 'How it works' section. The 'How it works' section describes two options: 'Automatic' (using a pre-trained model as a judge) and 'Human' (using real people to evaluate responses). There are also sections for 'Model evaluations' and 'Model metrics'.

In AI/ML & GenAI Solutions Archi... by Toby Foth...

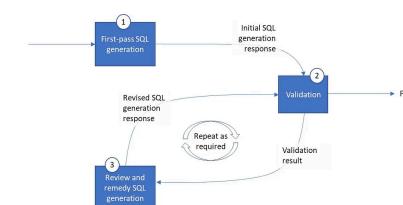
Efficient Generative AI model evaluation using Amazon Bedrock...

Authors: Toby Fotherby and Atul Varshneya

Dec 29, 2024

6

+



In AI/ML & GenAI Solutions Archi... by Atul Varsh...

Agentic Reflection for Natural Language to SQL Generation

Authors: Atul Varshneya, Toby Fotherby, Shweta Keshavanarayana

Dec 20, 2024

5

+

[See all from Toby Fotherby](#)[See all from AI/ML & GenAI Solutions Architecture](#)

All your favorite parts of Medium are now in one sidebar for easy access.

Recommended from Medium



Aarefa Bhurka

Local LLM Deployment in 2025: Run Powerful AI Models on Your...

In 2025, deploying Large Language Models (LLMs) locally is not just a research...

Aug 1

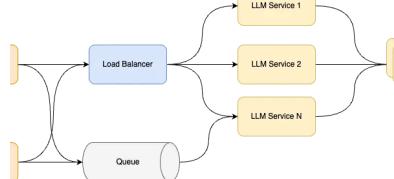


In Level Up Coding by Fareed Khan

Building a Self-Improving Agentic RAG System

Specialist agents, multi-dimensional eval, Pareto front and more.

Nov 14



Shukri Shukriev

LLM-as-Microservice: Integration Patterns and Trade-offs

Introduction

Jun 15

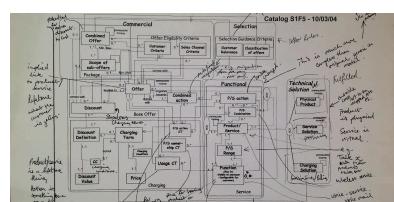


Rohit Sarkar

What I Really Learned About AI Agents.

12 Months, 300 Agents, and a Few Headaches Later

Sep 12



Marcin Ros

From Static Diagrams to AI Architects: The Future of IT...

In this article I touch the subject of IT Architecture design method—particularly...

Sep 1



In Towards AI by Alpha Iterations

Agentic AI Project: Build a Multi-Agent system with LangGraph an...

This is an end to end project on building a multi-agent insurance support system using...

Nov 12



[See more recommendations](#)

All your favorite parts of
Medium are now in one
sidebar for easy access.