

Broadband Company – Case study

About the Client

The client is a network infrastructure company that is at the forefront of connecting New Zealanders to the world. It built, and now operates the majority of the country's fiber internet network, ensuring high-speed, reliable and world-class internet connections for homes and businesses.

Client Challenge

Client architecture repository contains a complex technology landscape and poses challenges for new staff, who find it difficult to navigate and slow to answer their queries.

With consideration to the architecture documentation challenges, the client sought to investigate generative AI solutions to foster efficiency within their CTO Function. Specifically, they wanted to leverage generative AI to enable architects to access and understand information more quickly – improving productivity, efficiency, and speed up decision-making processes, with the goal of achieving the following:



Quantify the benefits of an aggregated model over and above using embedded AI assisted tools in source platforms.



The benefits of using additional capability to aggregate and present information from multiple sources to enable other use cases to be quantified.



An understanding of the associated operating model and capability needed to manage a generative Al capability (model establishment, training, operate, tuning, environment management, business and technology capability and skills).

Deloitte Solution

Deloitte built a generative Al solution on Amazon Bedrock to unify and operationalize the software architecture documents. Specific functions included:



Information Retrieval:

Receives a query in the form of text input searching architecture documents.



Drafted Response:

Provides an answer in a conversational format with links to appropriate source documents.



Validation, and Storage of Human Feedback:

Architects quickly cross-check to make sure information is correct, ask follow-ups, and rate the quality of first response generated by the tool. This feedback is saved so that it can be incorporated into an improved model in the future.

More specifically, the solution uses the following AWS services to perform these functions:

Amazon S3 and Amazon Kendra are used to store and index software architecture documents

Amazon Lex is used as the conversational Al interface for a chatbot

Amazon Bedrock is used as the API for LLMs to perform the vector search on Kendra and provide responses AWS Lambda is used in conjunction with Langchain for the orchestration

www.www.mannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenthermannenth

Results and Benefits

The generative AI based solution on AWS aggregated architectural documents across multiple sources and formats into a unified, searchable database. This eliminated the previous challenges of manually reading, and processing siloed documentation. Architects gained the ability to quickly search keywords and concepts, accessing required information in seconds rather than hours.

By quantifying the improvements, it was found that the Al-assisted aggregation and insight generation led to a 75% boost in architect productivity. Additional use cases became possible as well, such as automating cross-reference checks and identifying relationships across domains. The operating model study determined that while initial investment was required for capability building and model development, the long-term costs were surpassed by efficiency gains. The tuned generative models were able to operate with high accuracy once established, requiring only periodic reviews. With the proper environment and multidisciplinary team in place leveraging the technology, significant benefits were achieved which previously required substantial manual effort.

In summary, aggregating architectural documentation with generative AI delivered faster access to information, improved productivity and efficiency, enabled additional use cases, and ultimately boosted speed of decision making once adopted with the proper operating model.

About Deloitte

Deloitte is one of the largest professional services firms in the world and a leader in digital transformation strategy. Through a network of more than 330,000 professionals, industry specialists, and an ecosystem of alliances, Deloitte assists clients in turning complex business issues into opportunities for growth. They are a strategic global systems integrator with over 17,000 certified AWS practitioners across the globe. They continue to raise the bar through participation in the AWS Competency Program with 25 competencies including Machine Learning.

Contact us:

Damian Harvey

Partner - Consulting, Cloud and Engineering Lead

Deloitte

dharvey@delotte.co.nz

Kylie Bryant

Partner - Consulting

Deloitte

kybryant@deloitte.co.nz

Deloitte.

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see www.deloitte.com/about for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms. Please see www.deloitte.com/us/about for a detailed description of the legal structure of Deloitte LLP and its subsidiaries. Certain services may not be available to attest clients under the rules and regulations of public accounting.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. Thi publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

Copyright © 2023 Deloitte Development LLC. All rights reserved

