Data Innovation Hackathon

In collaboration with the Ministry of Justice and AWS

1 5:36

34% My 12





Case Study: Ministry of Justice Data Innovation Hackathon

Introduction:

The UK Ministry of Justice (MoJ) recently organised a two-part hackathon in collaboration with Cognizant and AWS to explore how AI, Big Data, and Large Language Models (LLMs) could be used to improve various aspects of their operations.

Day One of the hackathon took place on the 12th of March and it involved formulating the use case at the AWS office and Day Two took place on the 16th of April at the Cognizant office, where in the teams demonstrated their functional proof of concepts (PoCs) and presented to a judging panel comprising of a mix of MOJ, AWS and Cognizant technical leadership.

Over the course of this Hackathon, 70 colleagues from across digital, data platform, data engineering, analytics engineering, and data science at the Ministry of Justice we brought together, distributed into three teams and each team further paired with solution specialists from Cognizant and AWS. Working collaboratively, the teams co-developed project ideas and generated code base within the AWS platform, tools, and sandbox.

Use cases presented: The Hackathon focused on three main use cases:

1. Use of LLMs to understand, document, and improve codebases.

Challenge: Using LLMs to understand, document, and improve codebases was a significant challenge for the MoJ and its partners. These legacy codebases embody critical delivery of services and are often being difficult to maintain, understand, and extend due to legacy technology, lack of documentation, and departure of the original contributors. The technical debt associated with these codebases stifled innovation, increased maintenance costs, and posed a significant risk during migration.

Solution: To address this challenge, the team designed and developed a proof of concept (PoC)/pilot to explore and validate the application of LLMs to software development. The team used LLMs to explain code, produce documentation, suggest improvements, and translate code into another programming language following an AI augmented test driven development approach.

Business outcome: This approach reduced complexity in tackling legacy codes, maintenance costs, accelerate innovation, and mitigate risks across Data and Analytics engineering for the MoJ.

2. Improve methods for data ingestion and data sharing on the Analytical Platform.

Challenge: This use case aimed to address the diverse set of challenges that come with the initial ingestion of data onto the platform from across MoJ, adjacent bodies, and other organisations. These challenges included different data formats, security concerns, lack of consistent data standards, inability to access source systems, and challenges accessing data controlled by 3rd party organisations. This oversight undermines the security of the platform.





Solution: To address this challenge, the team designed and developed a proof of concept (PoC) / pilot to improve methods for data ingestion and data sharing on the Analytical Platform. The team used various methods, including connections to databases via AWS DMS; direct connection to HMCTS Azure blob storage; manual uploads to S3; and tools such as register-my-data and the data-uploader. The team crafted a solution "share-my-data", where users could more securely and automatically share data via email, while automatically checking with a certified 'data owner' and with the team receiving the data before sharing.

Business Outcome: This approach can reduce duplication of work, driving a more efficient, cost-effective and secure mechanism of data sharing. Track the path of data share and enable users to utilise more programmatic methods to share data in future, promoting security and audit-ability.

3. Design and implement a self-service pipeline architecture for training Large Language Models.

Challenge: This use case aimed to enhance the efficiency of data scientists by enabling them to concentrate on model development, in place of grappling with the intricacies of infrastructure management. As an example: Inconsistency in document uploads, with each user uploading in different formats, increases the complexity, resulting in significant efforts invested in verifying unstructured data.

Solution: To address this challenge, the team designed and implemented a Proof of Concept self-service pipeline architecture that could train Large Language Models. It specified that the architecture would enable users to initiate, monitor, and manage their training jobs seamlessly and explored the use of retrieval augmented generation (RAG) to improve quality of LLM outputs with accurate data.

Business outcomes: The autonomy and flexibility offered by the self-service architecture can expedite model insights and improvements. Use of LLMs on local infrastructure can enable its use in sensitive MoJ scenarios. Team created a POC of resilient and secure AWS infrastructure that provides an end-to-end solution, including both frontend and backend components integrated with the LLM solution. This significantly enhances efficiency by automating manual efforts and reduced risks. This cost-effective framework template for similar requirements across MoJ.

In conclusion, the MoJ Innovation Hackathon demonstrated the strength of the teams' combined abilities and collaboration in a cohesive environment. The use of AI, Big Data, and Large Language Models proved that they could help to automate the interpretation of legacy code, generate comprehensive documentation, identify code dependencies, and build a self-service data infrastructure whilst delivering greater understanding about the organisation and its wider employee ecosystem of Developers, Data Engineers, and Data Scientists. Technology partners, Cognizant and AWS in collaboration with MoJ defined the problem statements and validated the solution and its impact. Designing the solutioning by broadening the horizon of team to explore the art of possible solutions and operationalising them.

On the other hand, participants also gained hands-on experience of using innovative AI tools, enhancing their capabilities in code analysis, documentation, and migration strategies. Overall, the Hackathon instilled best practices of applying and adhering to standards of design, development, testing, and documentation, helping to build resilient solutions going forward.



About Cognizant

Cognizant (Nasdaq-100: CTSH) engineers modern businesses. We help our clients modernize technology, reimagine processes and transform experiences so they can stay ahead in our fast changing world. Together, we're improving everyday life. See how at www.cognizant.com or @Cognizant.

UK Headquarters

Cognizant Worldwide Limited 280 Bishopsgate, London EC2M 4AG

P.Kulandaivel@cognizant.com +447768526349

Copyright 2024, Cognizant. All rights reserved. No part of this document ay be reproduced, stored in a retrieval system, transmitted in any onic, mechanical, photocopying, out the express written permission means, el herwise , wi nation contained All other trademark οс here

C

fo