CAPSTONE PROJECT SUBMISSION FORM INSTRUCTIONS

Form Field Descriptions

FORM FIELD	DESCRIPTION OF REQUIRED INFORMATION
Project Title	Provide a project title to reflect the problem or solution space.
Organization	Provide the name of the partner organization. If the partner is a specific department within a large organization provide that information (e.g., Company XYZ Department of IoT Analytics).
Project POC(s)	Provide the name, title, and email address contact information for all the partner POCs who will be working with the project team throughout the semester. The first individual listed is considered the organization's primary POC for the project.
Knowledge Domain(s)	Click on the check box to select all knowledge domains which apply to this project.
Specialized Skills	Identify any specialized skills you believe might be useful for the project team members to know for the project.
Max Number of Project Teams	Click on the check box to select the number of project teams you can support with this project. This is useful when capstone partners want to compare approaches by multiple project teams.
New/Follow-on Project	Click on the check box to indicate if this is a new project or a follow-on project from a previous semester. Be sure to specify the semester and year of the earlier project (e.g., Fall/Spring/Summer 202x).
U.S. Citizenship Requirement	Click on the check box to indicate whether U.S. Citizenship is a requirement to work on this project.
Problem Description	Provide a brief description of the project problem which also incorporates sufficient background domain knowledge to provide context to understand the problem and potential solutions. We recommend one paragraph based on 5-6 sentences.
Project Goals	Provide a brief description of the project goals. Be as specific as possible to assist in defining the scope of the project.
Data	Provide a brief description of any partner-provided or open-source data to be used by the project.
Partner Intellectual Property	Identify any partner Intellectual Property (IP) (e.g., proprietary data, process, ML model, source code, etc.) which is to be part of the project solution. Identifying partner intellectual property up front at the beginning of the project will ensure it is not part of any graduate student intellectual property that is part of the project solution (typically research work associated with and generated by PhD students working on their PhD dissertation).
References	Indicate any reference materials you feel the project team should read in preparation for this project.
Project Deliverables	All DAEN 690 capstone projects generate a **Final Project Report**, a public **Final Public Project Presentation**, and a **Data and Code Repository**. Click on the check box for any additional project deliverables the team must provide as part of this project along with a brief description of the additional project deliverable.

Form Submission Instructions

Please email the completed capstone project submission form to Bernard Schmidt, Assistant Director, GMU Data Analytics Engineering (DAEN) Program at bschmid5@gmu.edu.

IMPORTANT Notice Regarding Non-Disclosure Agreements (NDAs)

The Data Analytics Engineering (DAEN) program is unable to support Non-Disclosure Agreements (NDAs) for capstone projects for two very simple reasons: a) the DAEN 690 capstone course is an *academic course* that students take for *academic credit* towards their graduate degree requirements and is not *paid sponsored research* where the faculty researcher and graduate research assistants *are university employees*, and b) since students in an academic course *are not university employees* they therefore can't be required to sign and be legally bound by a non-disclosure agreement that applies to the university and its employees.

If your project requires a Non-Disclosure Agreement (NDA) we will regrettably be unable to accept the project.

Project Title	Proposal Helper: Intelligent Proposal Generator
Organization	*Please provide the name of the partner organization for this project:* Allwyn Corporation
Project POC(s)	*Please provide the name, title, and email contact information of the partner POC(s) supporting this project:* Madhu Garlanka, CEO, Allwyn Corporation, madhu.garlanka@allwyncorp.com Swathi Young, CTO, Allwyn Corporation, swathi.young@allwyncorp.com Aditya Srinivas, Intern, Allwyn Corporation, aditya.muralidharan@allwyncorp.com
Knowledge Domain(s)	*Please select all knowledge domains which apply to this project:* ☒ Systems Engineering ☒ Data Engineering ☒ Data Mining ☐ Data Analytics ☐ Data Modeling/Simulation ☐ Data Visualization ☐ Computer Vision ☐ Natural Language Processing (NLP) ☒ AI/ML ☒ Generative AI ☐ DevSecOps ☐ MLOps ☐ Data Management ☐ Programming and Scripting ☐ Governance and Compliance
Specialized Skills	*Please indicate any specialized skills required to work on this project:* Experience with Large Language Models and ChatGPT APIs, web development (frontend/backend), authentication systems implementation.
Max Number of Project Teams	*Please indicate the maximum number of project teams which can work on this project during the semester:* ☒ 1 ☐ 2 ☐ 3 ☐ 4
New/Follow-on Project	*Please indicate whether this is a <u>new project</u> or a <u>follow-on project</u> from a previous semester:* ☒ New project ☐ Follow-on project from a previous semester (Semester Year): Fall/Spring/Summer 202x
U.S. Citizenship Requirement	*Please indicate whether U.S. citizenship is a requirement to work on this project:* ☐ Yes - U.S. citizenship required ☒ No - U.S. citizenship not required

PROBLEM DESCRIPTION

Allwyn Corp. face significant challenges in efficiently creating and managing proposals in an expedited manner, often requiring extensive manual effort to compile past performance data and case studies. The current process is time-consuming and prone to inconsistencies. There is a need for an automated Generative AI Retrieval-Augmented Generation (RAG) chatbot solution that can intelligently assist in proposal generation while ensuring that our past performance is fully utilized.

PROJECT GOALS

The project will be based on the Amazon Web Services (AWS) QnABot Solution, an open-source framework, designed to help organizations create conversational interfaces, such as chatbots, using AWS resources. It leverages services like Amazon Lex, AWS Lambda, Amazon Elasticsearch, and Amazon DynamoDB to provide a scalable and serverless architecture. Users can configure and manage the bot through a user-friendly Content Designer UI, allowing them to customize responses, integrate external APIs, and support dynamic, personalized interactions without requiring advanced technical skills. The AWS QnABot Solution offers cost-effectiveness by employing a serverless model where users pay only for the resources consumed. Organizations can deploy the solution quickly using AWS CloudFormation templates, enabling them to create engaging and efficient conversational experiences tailored to their specific needs. The chatbot is built following a series of steps specified in the implementation guide, but summarized below.

STEP 1: AWS SERVICES

The chatbot utilizes AWS serverless services which must first be built out using the QnABot CloudFoundation template.

STEP 2: RETRIEVAL-AUGMENTED GENERATION (RAG) INTELLIGENT SEARCH

The Retrieval-Augmented Generation (RAG) intelligent search step is incorporated to dynamically fetch relevant information from external sources, such as knowledge bases, documents, or APIs, to provide context-aware and accurate answers. This approach combines retrieval-based and generative methods to improve the bot's ability to handle complex or unstructured queries. Indexing and Storage of Knowledge Sources is done at this step. The project team will:

- Implement functionality to automatically extract and format past performance data.
- Design and implement secure data handling protocols.

STEP 3: CHATBOT WEB INTERFACE DESIGN

The chatbot web interface design step is crucial for creating a user-friendly and intuitive platform where users can interact with the chatbot. In the AWS QnABot Solution, designing the web interface involves configuring and customizing the chat widget or embedding the chatbot within a web application. The project team will develop an intuitive user interface for proposal generation.

STEP 4: PROMPT ENGINEERING DESIGN

The prompt engineering design step is an essential part of chatbot development, particularly when implementing advanced capabilities like dynamic responses or retrieval-augmented generation (RAG). In the context of the AWS QnABot Solution, prompt engineering focuses on crafting effective input prompts to guide the chatbot's behavior and ensure accurate, contextually relevant responses. For the purposes of this project the project team will utilize a prompt engineering design used with the Anthropic Claude foundation model. The Anthropic Prompt Engineering Overview offers a structured approach to crafting effective prompts, covering essential techniques such as clarity, directness, and the use of examples. Additionally, the guide discusses advanced methods like role prompting, where assigning Claude specific roles can enhance response accuracy and relevance.

STEP 5: ROLE PROMPTING

Role prompting is a technique in prompt engineering where a language model is assigned a specific role, persona, or context to guide its behavior and improve the relevance and tone of its responses. By explicitly defining the role—such as a historian, customer service representative, or software engineer—the model focuses on providing answers aligned with the assigned context. This method also allows users to control the response style, such as making it formal, concise, or creative. For example, assigning the model the role of a kindergarten teacher ensures its explanations are simple and engaging for young learners.

The benefits of role prompting include improved response relevance, consistency in tone, and adaptability to various scenarios. It can be used across applications like customer support, education, creative writing, and technical assistance. Best practices include clearly defining the role and context, specifying constraints like tone or format, and providing examples of desired outputs. By iterating and refining the prompt design, users can tailor the model's behavior to meet specific needs, making role prompting an essential strategy for leveraging AI effectively. The project team will create intelligent role case study generation capabilities.

STEP 6: TESTING AND EVALUATION

The testing and evaluation step is critical in constructing a chatbot to ensure it meets user expectations, operates reliably, and delivers a high-quality experience. This phase involves assessing the chatbot's functionality, accuracy, and usability through systematic testing and iterative refinement. For the purposes of this project, the chatbot testing and evaluation questions will be provided by Allwyn Corporation.

DATA SOURCES AND DATA

The following is an initial, but not comprehensive, list of data sources and data.

1. Allwyn's past performance database on SharePoint.
2. Historical proposal documents.
3. Case study repository.
4. Authentication system APIs.
5. Additional data sources to be provided by Allwyn Corporation.

PARTNER INTELLECTUAL PROPERTY

The following items shared by the partner with the project team are considered the Intellectual Property of Allwyn Corporation and are to be acknowledged as such by the project team who will not disclose any of the IP to the general public nor after the semester capstone ends.

1. Allwyn authentication system.
2. Past performance database.
3. Proprietary proposal templates and formats.
4. Internal case study database.

REFERENCES

1. **AWS Solutions Library / AWS Solution / QnABot on AWS** https://aws.amazon.com/solutions/implementations/qnabot-on-aws/
2. **QnABot Implementation Guide** https://docs.aws.amazon.com/solutions/latest/qnabot-on-aws/solution-overview.html
3. **Anthropic Prompt Engineering Overview** https://docs.anthropic.com/en/docs/build-with-claude/prompt-engineering/overview.
4. **Anthropic Interactive Prompt Engineering Tutorial** https://github.com/anthropics/prompt-eng-interactive-tutorial.
5. Dhareddy, P., Nayudu, S., Panjala, R., Tanniru, M., Vattipally, P., *"Generative AI Retrieval-Augmented Generation (RAG) Chatbot Prototype for City Governments,"* presented at the Fall 2024 DAEN 690 Capstone Presentation Showcase, Fairfax, Virginia, December 10, 2024.

PROJECT DEVELOPMENT ENVIRONMENT

Unless otherwise specified by the capstone project partner or mandated by the project specifics, project teams assigned to this project will be required to make use George Mason University (GMU) and College of Engineering and Computing (CEC) free compute resources to include:

- College of Engineering and Computing VSE Lab Computing Resources – including OpenStack (https://labs.vse.gmu.edu/)
- GMU Office of Research Computing (ORC) High-Performance Computing (HPC) Hopper and Argo clusters (https://orc.gmu.edu/resources/computing-systems/computing-resources/)
- DAEN 690 AWS Academy Learner Lab environment ($50 individual student accounts only)

For team-based commercial cloud service environments students can request access to the following CEC commercial cloud environments.

- Amazon Web Services
- Microsoft Azure

These commercial cloud-based environments are _not free_ in that they incur a charge to the College of Engineering and Computing for the use of their services. Project teams _must_ request access through their DAEN 690 instructor by providing _by the end of Sprint 1_

a _detailed budget for cloud services_ to be used by their respective projects throughout the semester. Access will be granted on a case-by-case basis once the project budget has been approved.

Note: Capstone project partners may be assessed a fee by the College of Engineering and Computing to cover the cost of commercial cloud service environment resources for their respective projects.

PROJECT OPEN SOURCE LICENSING

All code and project deliverables generated by the project team will be published under the **Apache License 2.0** permissive open-source software license.

WEEKLY PARTNER MEETING

The project team assigned to this project will be required to meet weekly with the capstone project partner via Zoom meeting on Thursdays from 2:00 PM EST to 3:00 PM EST starting on Thursday, January 30, 2025 and ending on Thursday, May 1, 2025.

PROJECT DELIVERABLES

- ☒ Capstone Showcase Presentation & PowerPoint Slides
- ☒ Final Project Report
- ☒ Repository for Data and Code Artifacts
- ☐ Machine Learning Model
- ☐ Working Prototype
- ☐ Other (please specify below)
    ⇒