Take-Home Assignment: Editorial Assistant Chatbot with RAG v1.2

Welcome and Thank You!

Thank you for taking the time to complete this take-home assignment for the Senior AI Engineer role with the Digital Strategy & Products department at CBC.

This exercise is designed to give you the opportunity to showcase your creativity, problem-solving approach, and technical expertise in building AI-powered solutions—specifically in the area of editorial support, which is core to CBC's public service mission.

We appreciate the effort and thought you'll bring to this assignment. Please feel free to explore, have fun with the problem, and show us how you'd approach designing something both practical and innovative.

Best of luck—we're excited to see what you build!

Objective

Design and prototype an AI-powered editorial assistant chatbot that leverages Retrieval-Augmented Generation (RAG) to assist CBC's editorial teams. The chatbot should be able to answer editorial policy FAQs, provide SEO suggestions, and generate summaries or headlines based on internal documentation and article content.

Background

Editorial teams at CBC handle fast-paced content publishing. To support them, we're exploring AI solutions that:

- Answer questions based on internal editorial guidelines.
- Help with tasks like SEO optimization and content summarization.
- Deliver accurate and trustworthy responses with citations from internal documents.

Your Task

Build a **chatbot prototype** that does the following:

RAG System Setup

- Use a **vector database** (e.g., FAISS, Elasticsearch) to index a small collection of news articles (use the provided JSON dataset) and editorial guideline documents (use this link)
- Implement a **retriever** to fetch relevant content chunks from the indexed corpus.
- Create a **generative model** using Hugging Face open-source community platform (Free Plan) to generate answers based on the retrieved context.

Chatbot Features

- Accept queries like:
    - "What's CBC's guideline on citing anonymous sources?"
    - "Suggest an SEO-optimized headline for this article: [Insert article id]"
    - "Summarize this article for a Twitter post."
- The chatbot should:
    - Search across internal content.
    - Answer editorial policy questions.
    - Generate SEO-friendly headlines.
    - Summarize long content into social snippets.

Bonus Features (Optional)

- Handle ambiguous queries gracefully.
- Add **memory** or simple context history.
- Include a **UI** with input fields and display responses in a clean format.
- Return **citations** when applicable (e.g., source doc or news id).

Deliverables

1. In the CBC GitHub repository in which you are a collaborator, please include the following:
    - Source code
    - A README file explaining:
        i. Setup instructions
        ii. Technical choices (e.g., model, vector store, chunking method). Specific to the choice of model, what choices would you make beyond Hugging Face (what this assignment asked for), and why?
        iii. Sample test conversations showing:
            1. A policy Q&A
            2. A headline suggestion
            3. A tweet-style summary of a sample article
2. A link to a short demo video showcasing the system in action **(Optional)**

Evaluation Criteria

- **Relevance** and **accuracy** of RAG answers
- **Generative quality** of summaries/headlines
- **Code quality** and **modularity**
- **Clarity** of documentation
- Thoughtfulness in design decisions and improvements