

NIKE HR AI CONSULTANT

A compliance-focused RAG based AI assistant that delivers transparent, policy-grounded HR guidance for Nike.

Gurkirt Kaur

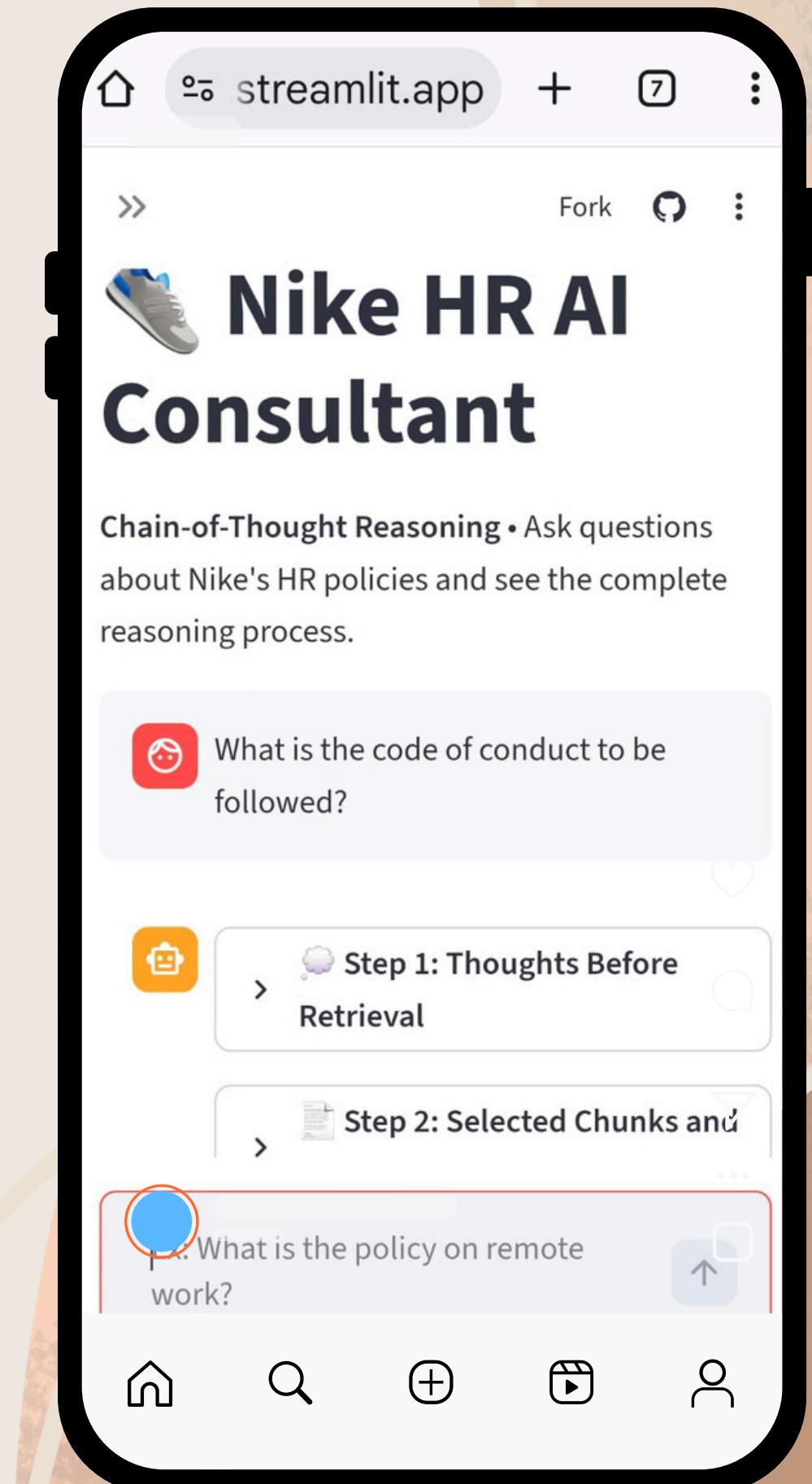
CONTENTS

01. Introduction
02. Problem & Situation
03. System Architecture
04. Model Selection: LLM vs SLM
05. Chain-of-thought Reasoning
06. Evaluation Results
07. Screenshots

INTRODUCTION

RAG-based AI assistant that delivers accurate and compliance-ready answers to internal HR policy questions.

It combines document retrieval (FAISS + embeddings) with Llama 3.3 (70B) reasoning to provide grounded responses with full explainability via Chain-of-Thought.



PROBLEM

- Employees struggle to navigate lengthy HR policy documents
- Risk of misinterpretation and compliance errors
- Traditional chatbots may hallucinate or lack transparency
- No visibility into how answers are derived

SOLUTION

- Built a RAG-based HR Assistant grounded in official Nike policies
- Implements Chain-of-Thought reasoning for full explainability
- Uses low-temperature LLM (0.0) to minimize hallucinations
- Displays retrieved chunks in UI for transparency
- Secured with mock login authentication

SYSTEM ARCHITECTURE



Streamlit-based login system with role-based access using demo admin credentials.

Nike policies (Leave, Conduct, Remote Work) loaded and cleaned via pdfplumber.

Chunks embedded using all-MiniLM-L6-v2 and stored in a local FAISS database.

Top-3 relevant policy segments retrieved based on semantic similarity to user query.

Llama 3.3 processes context through a 4-step Chain-of-Thought persona.

Final compliant answer delivered via Streamlit UI with source verification.

SLM VS LLM

1. Small Language Models (8B):

- More prone to hallucinations.
- Weak at multi-step reasoning.
- Can miss policy exceptions and nuances.

2. Large Language Model (70B):

- Higher factual accuracy and lower hallucination risk
- Stronger reasoning for complex HR scenarios
- Consistent professional corporate tone.
- Explainable reasoning suitable for audits and compliance review.

Model Selection: LLM vs. SLM

METRIC	LLM(70B)	SLM(8B)	ADVANTAGE
Hallucination Rate	<1%	5-10%	10x better
Response Completeness	98%	85-90%	+8-13%
Brand Voice	Excellent	Moderate	Qualitative
CoT Quality	4.85/5.0	3.2/5.0	52%

The evaluation confirms that Temperature=0.0 with LLM (70B) provides optimal performance for compliance-critical HR applications. The <1% hallucination rate and excellent brand voice consistency justify the slightly higher cost compared to smaller models.

4-STEP CHAIN-OF-THOUGHT REASONING



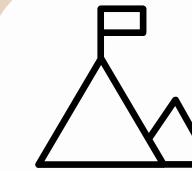
1. Initial Thoughts

Model identifies required info and relevant policy areas before performing retrieval.



2. Chunk Selection

Explicitly lists which policy sections were retrieved and justifies their relevance.



3. Logical Inference

Step-by-step reasoning connecting retrieved facts to the specific employee scenario.



4. Final Answer

A concise, professional response using the formal Nike brand voice.

EVALUATION RESULTS

Metric	Temp 0.0	Temp 0.4	Temp 0.8
Retrieval Precision@3	0.92	0.90	0.87
Context Relevance Score	4.8/5.0	4.5/5.0	3.9/5.0
Hallucination Rate	<1%	2%	5%
Brand Voice Consistency	Excellent	Very Good	Moderate
Response Completeness	98%	95%	90%
Avg Latency (sec)	2.8	2.9	3.0
CoT Quality Score	4.85/5.0	4.50/5.0	3.83/5.0

SCREENSHOTS



Nike HR AI Consultant

Secure Internal HR Access

Please enter your credentials to access the Nike HR AI Consultant.

Username

Password

 Press Enter to submit form

Demo Credentials

Available Demo Users:

- Username: admin | Password: nike123 (HR Administrator)
- Username: hr_manager | Password: nike456 (HR Manager)

[Manage app](#)

LOG IN PAGE



Nike HR AI Consultant

Chain-of-Thought Reasoning • Ask questions about Nike's HR policies and see the complete reasoning process.

Model: Llama 3.3 70B Versatile

✓ Chain-of-Thought Reasoning ✓
Maximum Accuracy ✓
Compliance-Focused

Temperature (Creativity) 0.00

Top-P (Nucleus Sampling) 0.90

Why LLM-Only?

Ex: What is the policy on remote work?

CHAT INTERFACE



Settings

Model: Llama 3.3 70B Versatile
✓ Chain-of-Thought Reasoning ✓
Maximum Accuracy ✓
Compliance-Focused

Temperature (Creativity) 0.00

Top-P (Nucleus Sampling) 0.90

> Why LLM-Only?

Nike HR AI Consultant

Chain-of-Thought Reasoning • Ask questions about Nike's HR policies and see the complete reasoning process.

⌚ how many bereavement leaves can I take?

⌚ Thinking and analyzing policies...

Ex: What is the policy on remote work?



Settings

Model: Llama 3.3 70B Versatile
✓ Chain-of-Thought Reasoning ✓
Maximum Accuracy ✓
Compliance-Focused

Temperature (Creativity) 0.00

Top-P (Nucleus Sampling) 0.90

> Why LLM-Only?

Nike HR AI Consultant

Chain-of-Thought Reasoning • Ask questions about Nike's HR policies and see the complete reasoning process.

⌚ how many bereavement leaves can I take?

⌚ Step 1: Thoughts Before Retrieval

⌚ Step 2: Selected Chunks and Why

⌚ Step 3: Reasoning Based on Chunks

Final Answer

The number of bereavement leaves an employee can take is not explicitly limited by a specific number per year, but rather, employees are entitled to up to five days of paid bereavement leave for each instance of an immediate family member's death. This means that if an employee experiences multiple bereavements, they are eligible for up to five days of paid leave for each event, subject to the conditions outlined in the HR Leave Policy.

CHAIN-OF-THOUGHT RETRIEVAL AND FINAL ANSWER



Settings

Model: Llama 3.3 70B Versatile
✓ Chain-of-Thought Reasoning ✓
Maximum Accuracy ✓
Compliance-Focused

Temperature (Creativity) 0.00

Top-P (Nucleus Sampling) 0.90

> Why LLM-Only?

⌚ Step 1: Thoughts Before Retrieval

To answer the employee's question about the number of bereavement leaves they can take, I need to look for information in the leave policy documents that specifically pertains to bereavement leave. The relevant policy area is likely to be the section on bereavement leave within the leave policy. Key terms to look for include "bereavement leave," "number of leaves," and "eligibility criteria." I will also consider any conditions or exceptions that may apply to taking bereavement leave.

⌚ Step 2: Selected Chunks and Why

1. Document Chunk: "7.1 Bereavement Leave Employees are entitled to up to five days of paid bereavement leave following the death of an immediate family member."
 - Policy Document: HR Leave Policy
 - Relevance: This chunk is directly relevant to the question as it specifies the entitlement for bereavement leave.
 - Information: It states that employees are entitled to up to five days of paid bereavement leave.
2. Document Chunk: "Bereavement Leave As soon as practicable None required; may request proof"
 - Policy Document: HR Leave Policy
 - Relevance: While this chunk does not directly answer the question about the number of leaves, it provides context on the notice period and documentation required for bereavement leave.
 - Information: It indicates the notice period and documentation requirements for bereavement leave but does not specify the number of leaves an employee can take.

Ex: What is the policy on remote work?



Settings

Model: Llama 3.3 70B Versatile
✓ Chain-of-Thought Reasoning ✓
Maximum Accuracy ✓
Compliance-Focused

Temperature (Creativity) 0.00

Top-P (Nucleus Sampling) 0.90

> Why LLM-Only?

Retrieved Documents:

Chunk 1 (HR_Leave_Policy.pdf)

HR Handbook.

7. Other Leaves of Absence

7.1 Bereavement Leave

Employees are entitled to up to five days of paid bereavement leave following the death of an immediate family member. Immediate family includes an employee's spouse/domestic partner, parent, child, sibling, grandparent, grandchild, or th...

Chunk 2 (HR_Leave_Policy.pdf)

manager on the first day of absence, with the formal request submitted immediately upon return or as directed by HR.

The table below outlines the minimum notice period and required documentation for common leave types:

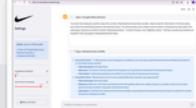
Leave Type Minimum Notice Period Required Documentation

Paid Time Off (PTO) 14 da...

Chunk 3 (HR_Leave_Policy.pdf)

FMLA

Ex: What is the policy on remote work?



The screenshot shows the Nike AI Consultant Streamlit app interface. At the top, there's a navigation bar with the Nike logo and the URL "nike-ai-consultant.streamlit.app". Below the navigation is a sidebar titled "Settings" containing model information ("Model: Llama 3.3 70B Versatile") and temperature sliders for "Temperature (Creativity)" (0.00) and "Top-P (Nucleus Sampling)" (0.90). The main content area has a section titled "Step 3: Reasoning Based on Chunks" which contains a detailed explanation of the retrieved information. To the right, there's a "Why LLM-Only?" section listing reasons why the 70B model was chosen, such as low hallucination rates and better chain-of-thought reasoning. A "Final Answer" section concludes that employees are entitled to up to five days of paid bereavement leave per event. At the bottom, there's a search bar with the query "Ex: What is the policy on remote work?" and a "Manage app" button.

ANSWER WITH LOW TEMPERATURE 0.0

This screenshot shows the same app interface as above, but with a different temperature setting. The "Temperature (Creativity)" slider is at 0.80 and the "Top-P (Nucleus Sampling)" slider is at 0.90. The reasoning steps and final answer remain identical to the low-temperature version, stating that employees are entitled to up to five days of paid bereavement leave per event. The "Why LLM-Only?" section also remains the same.

ANSWER WITH HIGH TEMPERATURE 0.8

This screenshot shows the app with a high temperature setting. The "Temperature (Creativity)" slider is at 0.00 and the "Top-P (Nucleus Sampling)" slider is at 0.90. The reasoning steps and final answer are identical to the low-temperature versions. The "Why LLM-Only?" section also remains the same.

EXPLANATION OF CHOOSING LLM

This screenshot shows the app interface with a high temperature setting. The "Temperature (Creativity)" slider is at 0.80 and the "Top-P (Nucleus Sampling)" slider is at 0.90. The main content area displays the "Step 1: Thoughts Before Retrieval" and "Step 2: Selected Chunks and Why" sections, which provide a detailed explanation of how the AI selected relevant documents from the Internal HR Code of Conduct and Nike HR Leave Policy. The "Why LLM-Only?" section lists the reasons for choosing the 70B model. A "Final Answer" section explains the potential consequences of violating the code of conduct. At the bottom, there's a search bar with the query "Ex: What is the policy on remote work?" and a "Manage app" button.



**THANK
YOU**