### Developing a Policy QA System Using Retrieval-Augmented Generation (RAG) Framework

#### 一、研究背景及意义

随着信息技术的快速发展，政策信息的获取和理解变得越来越重要。政策问答系统能够帮助用户快速找到相关政策信息，并提供准确的解答。然而，传统的问答系统在处理复杂和动态的政策信息时，常常面临知识更新不足和生成内容不准确的问题。检索增强生成（RAG）框架将外部知识库的检索能力与大语言模型（LLMs）的生成能力相结合，为政策问答系统提供了一种新的解决方案。通过动态获取最新的政策信息，RAG框架能够显著提升政策问答的准确性和实用性。

#### 二、国内外研究现状

近年来，RAG技术在自然语言处理（NLP）领域得到了广泛关注。许多研究集中在RAG的基本框架和应用上，探索如何将其有效应用于问答系统。Gao等[1]对RAG进行了全面的综述，指出其在知识密集型任务中的应用潜力。而Karpukhin等[2]提出的Dense Passage Retrieval方法，进一步增强了检索精度，为政策QA系统提供了有力支持。此外，国内的研究也逐渐增多，探讨了如何结合RAG框架优化政策问答的生成质量[3]。这些研究为RAG在政策领域的应用奠定了基础。

#### 三、文献评述

现有文献对RAG框架的研究主要集中在其结构和应用效能上。早期的简单RAG方法虽然在某些任务上表现良好，但在检索精度和生成质量上存在不足[1]。为解决这些问题，研究者们提出了先进RAG和模块化RAG，通过引入预检索和后检索优化策略，显著提升了信息检索的效率。然而，尽管RAG在技术上取得了进展，如何在政策问答系统中有效整合这些技术仍然是一个值得深入探讨的问题。

在实际应用中，政策QA系统需要处理大量的政策文本和法律条款，这对生成模型的理解和推理能力提出了更高的要求[3]。此外，文献中普遍存在对评估标准和方法缺乏统一的认识，这使得不同研究间的比较变得困难。因此，未来的研究需要在评估方法上达成共识，以便更好地衡量RAG在政策问答中的有效性。

1. **参考文献**

（引用格式不对，后期再改）

|  |
| --- |
| [1]<https://doi.org/10.48550/arXiv.2312.10997> |

[2]

|  |
| --- |
| <https://doi.org/10.48550/arXiv.2004.04906> |

[3]Yang, H.; Li, S.; Gonçalves, T. Enhancing Biomedical Question Answering with Large Language Models. *Information* **2024**, *15*, 494. <https://doi.org/10.3390/info15080494>