**Prompt Engineering Strategy and Rationale**

**1. Page Summarization**

**Prompt:**

Summarize the following medical content from page {page\_number} in 100-150 words:

{page\_content}

Focus on key cardiac physiology concepts, conduction system details, and any important diagrams or charts mentioned.

**Rationale:**

* Specifies the desired length (100-150 words) to ensure concise summaries.
* Focuses on key cardiac physiology concepts and conduction system details to capture the most relevant information.
* Mentions diagrams and charts to ensure visual elements are not overlooked.

**Potential Improvements:**

* Could add a request for bullet points or a structured format to make summaries more scannable.
* Might include a request to highlight any new terms or concepts introduced on the page.

**2. Flashcard Generation**

**Prompt:**

Create 5 flashcards (question-answer pairs) based on the following medical content from page {page\_number}:

{page\_content}

Ensure the flashcards cover key cardiac physiology concepts, conduction system details, and any important diagrams or charts mentioned.

Format your response as a JSON array of objects, each with 'question' and 'answer' keys.

**Rationale:**

* Specifies the number of flashcards (5) to ensure consistency across pages.
* Focuses on key concepts, details, and visual elements to capture the most important information.
* Requests a specific JSON format for easy parsing and structuring of the output.

**Potential Improvements:**

* Could specify a maximum length for questions and answers to ensure conciseness.
* Might request a mix of different question types (e.g., definitions, processes, comparisons) for variety.

**3. Search Query Generation**

**Prompt:**

Generate a short, general search query (5-7 words) to find medical images related to the cardiac physiology content on page {page\_number}:

{page\_content}

The query should be related to cardiac anatomy, physiology, or the heart's conduction system mentioned in the content.

Provide only the search query, without any additional text or quotation marks.

**Rationale:**

* Specifies the desired length (5-7 words) to create focused yet not overly specific queries.
* Directs the focus to cardiac anatomy, physiology, and the conduction system to ensure relevance.
* Requests only the query without additional text to simplify parsing.

**Potential Improvements:**

* Could request multiple queries per page to increase the chances of finding relevant images.
* Might include a request to prioritize visual elements mentioned in the text (e.g., specific diagrams or structures).

**Iterations and Improvements**

During development, several iterations were made to refine the prompts:

1. Added specific word counts and format instructions to improve consistency of outputs.
2. Included mentions of diagrams and charts in summarization and flashcard prompts to ensure visual information was captured.
3. Refined the search query prompt to focus on generating queries that would yield relevant medical images.
4. Adjusted the flashcard prompt to request JSON format output for easier processing.