



NOVA
NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTMENT OF
COMPUTER SCIENCE
DEPARTMENT OF
MATHEMATICS

MIGUEL ÂNGELO AZEITONA DOS SANTOS
BSc in Computer Science Engineering

**DEVELOPING END-TO-END,
DEEP LEARNING-BASED CHATBOTS FOR
HEALTHCARE SUPPORT IN PORTUGUESE**

MASTER IN ANALYSIS AND ENGINEERING OF BIG DATA
NOVA University Lisbon
September, 2024



DEPARTMENT OF
COMPUTER SCIENCE
DEPARTMENT OF
MATHEMATICS

DEVELOPING END-TO-END, DEEP LEARNING-BASED CHATBOTS FOR HEALTHCARE SUPPORT IN PORTUGUESE

MIGUEL ÂNGELO AZEITONA DOS SANTOS

BSc in Computer Science Engineering

Adviser: Federico Guede, Ph.D.

Head of Data Science, Value for Health CoLAB

Co-adviser: João Carlos Gomes Moura Pires, Ph.D.

Associate Professor, NOVA University Lisbon

Examination Committee

Chair: Paula Alexandra da Costa Amaral, Ph.D.

Associate Professor, NOVA University Lisbon

Rapporteur: João Balsa da Silva, Ph.D.

Associate Professor, University of Lisbon

Adviser: Federico Guede, Ph.D.

Head of Data Science, Value for Health CoLAB

Co-adviser: João Carlos Gomes Moura Pires, Ph.D.

Associate Professor, NOVA University Lisbon

MASTER IN ANALYSIS AND ENGINEERING OF BIG DATA

NOVA University Lisbon

September, 2024

Developing End-to-End, Deep Learning-Based Chatbots for Healthcare Support in Portuguese

Copyright © Miguel Ângelo Azeitona dos Santos, NOVA School of Science and Technology, NOVA University Lisbon.

The NOVA School of Science and Technology and the NOVA University Lisbon have the right, perpetual and without geographical boundaries, to file and publish this dissertation through printed copies reproduced on paper or on digital form, or by any other means known or that may be invented, and to disseminate through scientific repositories and admit its copying and distribution for non-commercial, educational or research purposes, as long as credit is given to the author and editor.

To my parents, Adérito and Cidália, for their constant support and encouragement throughout my academic journey. Their patience has been truly boundless.

ACKNOWLEDGEMENTS

I would like to begin by expressing my deepest gratitude to Federico Guede, who kindly accepted me as his student for this thesis. His unwavering guidance, mentorship, and support throughout this journey were invaluable, and I am deeply thankful for his dedicated commitment to my academic development. I would also like to extend my sincere appreciation to Professor João Moura Pires, whose insightful contributions and critical feedback significantly shaped and refined this work. Furthermore, I am grateful to Professor Paula Amaral for her guidance during the initial stages of this research.

I would like to take this opportunity to acknowledge the unwavering support and encouragement of my parents, Adérito and Cidália Santos, whose sacrifices have made my academic achievements doable. Without them, none of this would have been possible.

Moreover, I am particularly grateful to my grandmother Deolinda, whose continuous interest in my academic pursuits and unwavering support have been a source of great motivation. I would like to extend my heartfelt thanks to my friends Eduardo Medeiros, Ion Chirica, and João Samarão for their companionship, advice, and the many moments of laughter and joy we shared throughout this journey.

Finally, I would like to express my gratitude to all my family and friends who believed in me throughout this journey. Their encouragement and support have been invaluable.

”

*“A person who never made a mistake never tried
anything new.”*

— **Albert Einstein**

ABSTRACT

Access to healthcare is a fundamental human right, crucial for individual and societal well-being. Unfortunately, barriers such as financial constraints and geographical limitations often make this access not universally straightforward. Simultaneously, healthcare systems worldwide are increasingly burdened by growing patient demand and limited resources. The constant development of fields like artificial intelligence allows the emergence of solutions like chatbots which can support a variety of health-related tasks enhancing both patient and medical experience.

This thesis focuses on the proposal of an architecture for the development of informative Healthcare-specific chatbots in the Portuguese language. The proposal employs Large Language Models, Retrieval Augmented Generation, and Prompt Engineering combined with medical-certified documentation to perform accurate and trustable conversational tasks with language accessible to the patient. Further, a web interface allows the users to interact with the system.

The architecture was validated in the medical domains of Deep Brain Stimulation and Parkinson's Disease, and Cardiothoracic Surgery. Among different configurations of models and system parameters, the best-performing configuration involved the LLAMA3-8B model with a chunk size of 256 tokens, achieving an average BERTScore of 0.7795 when evaluated on a reference corpus.

To support the chatbot's performance, two sets of validated question-answer pairs in Portuguese were introduced: 23 pairs related to Deep Brain Stimulation and Parkinson's Disease, and 139 pairs addressing Cardiothoracic Surgery. These domain-specific datasets, vetted by healthcare professionals, represent a valuable resource for future advancements in medical chatbot applications.

Keywords: Healthcare, Question Answering, Large Language Models(LLM), Retrieval Augmented Generation (RAG), Prompt Engineering, Portuguese language

RESUMO

O acesso a cuidados de saúde é um direito humano e crucial para o bem-estar individual e social. Infelizmente, obstáculos como condições financeiras e limitações geográficas frequentemente dificultam este acesso de maneira significativa. Simultaneamente, os sistemas de saúde estão cada vez mais sobrecarregados. O constante desenvolvimento de campos como a inteligência artificial permite o surgimento de soluções como os chatbots, que podem apoiar uma série de tarefas relacionadas com a saúde, melhorando tanto a experiência do paciente como a do médico.

Esta tese centra-se na proposta de uma arquitetura para o desenvolvimento de chatbots informativos específicos da área da saúde em língua portuguesa. A proposta emprega *Large Language Models*, *Retrieval Augmented Generation* e *Prompt Engineering* combinados com documentação médica certificada para realizar tarefas de conversação precisas e confiáveis com linguagem acessível ao paciente. Além disso, uma interface Web permite ao utilizador interagir com o sistema.

A arquitetura foi validada nos domínios médicos da estimulação cerebral profunda e da doença de Parkinson, e da Cirurgia Cardiorrástica. Entre as diferentes configurações de modelos e parâmetros do sistema, a configuração com melhor desempenho envolveu o modelo LLAMA3-8B com uma dimensão de 256 tokens, obtendo um *BERTScore* médio de 0,7795 quando avaliada num corpus de referência.

Adicionalmente, foram introduzidos dois conjuntos de pares pergunta-resposta em português: 23 pares relacionados com a Estimulação Cerebral Profunda e a Doença de Parkinson, e 139 pares relacionados com a Cirurgia Cardiorrástica. Estes conjuntos, avaliados por profissionais de saúde, representam um recurso valioso para futuros avanços nas aplicações de chatbot médico.

Palavras-chave: Saúde, Pergunta-Resposta, *Large Language Models*, *Retrieval Augmented Generation*, *Prompt Engineering*, Língua portuguesa

CONTENTS

List of Figures	x
1 Introduction	1
1.1 Context and Scope	1
1.2 Problem Statement and Motivation	2
1.3 Goals	3
1.4 Main Contributions	3
1.5 Report organization	4
2 State of the Art	5
2.1 Question Answering in Natural Language Processing	5
2.2 Question Answering Approaches	5
2.2.1 Rule-Based approaches	5
2.2.2 AI approaches	6
2.2.3 Large Language Models	7
2.3 Evaluation Metrics in Question Answering and Retrieval Augmented Generation	11
2.3.1 Generation Quality	12
2.3.2 Retrieval Quality	15
2.3.3 Other metrics	15
2.4 Question Answering in Portuguese	16
2.5 Question Answering in Healthcare	16
2.6 Summary	17
3 Proposed Methodology	18
3.1 Architecture description	18
3.1.1 Safety Check and Responsible AI	20
3.1.2 Overview	21
3.2 Validation Procedure	21
3.2.1 Retrieval Stage	22

3.2.2	Generation Stage	22
4	Implementation	23
4.1	Datasets	23
4.1.1	Deep Brain Stimulation and Parkinson’s Disease	23
4.1.2	Cardiothoracic Surgery	24
4.2	Validation Sets	24
4.2.1	Deep Brain Stimulation and Parkinson’s Disease	24
4.2.2	Cardiothoracic Surgery	25
4.2.3	Retrieval Validation	25
4.3	Infrastructure and Experimental Setup	25
4.3.1	Infrastructure used	26
4.3.2	Experimental Setup	26
4.4	Technical Overview	28
4.4.1	Pipeline	28
4.4.2	Web Interface	29
4.5	Implementation Issues	29
4.6	Final Experimental Setup	30
5	Experiments and Results	32
5.1	Experiments	32
5.2	Results	33
5.2.1	Deep Brain Stimulation and Parkinson’s Disease	33
5.2.2	Cardiothoracic Surgery	38
5.3	Discussion	39
5.3.1	Limitations	40
6	Conclusion and Future Work	42
6.1	Conclusion	42
6.2	Future Work	43
	Bibliography	44
	Appendices	
A	Prompts	50
B	Generation Quality Results	52
B.1	Using BerTimbau Base as Embedding Model	52
B.1.1	Chunk size of 128 tokens	52
B.1.2	Chunk Size of 256 tokens	53
B.1.3	Chunk Size of 512 tokens	54

C	Comparison with popular solutions	55
C.1	Query - Quando tomo a medicação dopaminérgica, os meus sintomas melhoram. Posso fazer a DBS?	55
C.1.1	ChatGPT	55
C.1.2	Gemini	56
C.1.3	Our solution	57
C.2	Query - O que é a DBS?	57
C.2.1	ChatGPT	57
C.2.2	Gemini	58
C.2.3	Our solution	59
C.3	Query - Existe cura para a doença de Parkinson?	59
C.3.1	ChatGPT	59
C.3.2	Gemini	59
C.3.3	Our solution	60
C.4	Query - Quantas pessoas vivem com a doença de Parkinson em Portugal?	61
C.4.1	ChatGPT	61
C.4.2	Gemini	61
C.4.3	Our solution	61

Annexes

I	Validation Set Deep Brain Stimulation and Parkinson’s Disease	62
II	Validation Set Cardiothoracic Surgery	73

LIST OF FIGURES

2.1 Empower LLMs to Perform Better on Domain-Specific Question Answering [45]	9
2.2 Naive RAG [51]	10
2.3 Advanced RAG [51]	11
2.4 BertScore example for recall [56]	15
3.1 Stages of the Indexing Stage	19
3.2 Stages of the Retrieval Stage	19
3.3 Stages of the Generation Stage	19
3.4 Architecture of Safety Check module	21
3.5 Architecture of each chatbot	21
4.1 Web Interface of the chatbot created for DBS and Parkinson’s Disease	29
5.1 Hit Rate for DBS dataset $K = 3$	33
5.2 Hit Rate for DBS dataset $K=5$	34
5.3 Correlation between Hit Rate and BertScore	35
5.4 Hit Rate for the Cardiothoracic Surgery dataset $K=3$	38
5.5 Hit Rate for the Cardiothoracic Surgery dataset $K=5$	39

INTRODUCTION

In this chapter, the thesis topic is introduced followed by its motivation, goals, expected contributions, and the document structure.

1.1 Context and Scope

In an era marked by rapid technological advancements, the intersection of healthcare and artificial intelligence (AI) has given rise to innovative solutions with the potential for mutual benefits for both patients and medical professionals. Among these advancements, the development of AI chatbots in healthcare stands out as a transformative tool [2].

These intelligent systems facilitate real-time interactions with patients, providing personalized health information and assistance. By seamlessly integrating into healthcare workflows, conversational agents contribute to the efficiency of medical services, empowering individuals to proactively manage their well-being and optimizing resource utilization for healthcare professionals [3].

This work aims to develop an architecture for the development of Healthcare-specific chatbots with a specific emphasis on meeting the needs of the Portuguese-speaking community. The focus centers on two distinct use cases within the healthcare domain: Cardiothoracic Surgery and Parkinson's Disease. These use cases have been chosen due to the availability of documentation written in an accessible format for patients, i.e., avoiding excessive use of medical jargon. Research highlights that information with medical jargon can be hard for patients to understand especially in patients with low health literacy [4]. On the contrary, these documents are structured to present complex information in a more convenient format for patients. Additionally, the use cases allowed collaboration with medical teams, ensuring the chatbot's content is medically accurate and accessible.

This thesis was proposed by Value for Health Colab (VOH.CoLAB), a private non-profit association aiming to expedite the transition to value-based healthcare through citizen engagement. Recognized by the Portuguese Foundation for Science and Technology, VOH.CoLAB's multidisciplinary team focuses on creating valuable health pathways involving all stakeholders. With a strong emphasis on leveraging Big Data, VOH.Colab

uses data science and analytics to measure healthcare value, identify patterns, and optimize healthcare delivery. Through collaboration, they promote data sharing and drive data-driven innovation to improve patient outcomes in the context of value-based care [5].

1.2 Problem Statement and Motivation

Conversational agents, often in the form of AI chatbots or virtual assistants, have emerged as transformative tools in healthcare. Powered by advancements in natural language processing and artificial intelligence [6], they improve patient engagement and healthcare delivery by facilitating seamless interactions between patients and healthcare systems.

Since the release of ChatGPT, Large Language Models (LLMs) have gained considerable attention. They have been widely adopted across various domains, including industry, agriculture [7], insurance [8], clinical studies [9], pharmacy [10], finance [11], and customer service [12]. This widespread adoption is primarily due to their state-of-the-art performance in most Natural Language Processing (NLP) tasks, such as text summarization, language translation, document classification, and notably, question-answering.

These general-purpose LLMs present several challenges, particularly their lack of specialization in domain-specific content, this is, while these models are proficient in understanding and generating text across a wide range of topics, they often lack the nuanced understanding required for specialized domains. In the healthcare domain, the use of a vocabulary accessible to the patient is critical and LLMs can generate educational content at an appropriate level for the patient [13] but terminology and context are also crucial and LLMs may not possess the depth of knowledge or accuracy required for tasks such as medical diagnosis or treatment recommendations [14].

Another significant challenge lies in the linguistic aspect of these models. The majority of NLP research and datasets are geared toward the English language despite the existence of hundreds of languages in the world [15]. Given that healthcare is a domain that requires a reinforced understanding of context and terminology, this linguistic bias has led to a lack of solutions for the non-English-speaking populations.

Moreover, patient literacy regarding health procedures is a critical factor in healthcare because it directly impacts a patient's ability to understand their health conditions, make informed decisions, and follow treatment plans effectively. Consequently, low health literacy can lead to poor health outcomes, increased hospitalizations, and higher healthcare costs [16].

Given these challenges, AI chatbots designed specifically for healthcare hold great potential to address both the specialization and accessibility gaps. By tailoring chatbot architectures to the healthcare domain, these systems can be designed to provide clear, accessible, and contextually relevant health information tailored to the specific needs and literacy levels of patients, addressing the limitations of general-purpose LLMs. Additionally, integrating support for underrepresented languages, such as Portuguese, can

bridge the linguistic gap and ensure that healthcare services become more inclusive for non-English-speaking populations.

Therefore, there exists a clear motivation for the development of an architecture that facilitates the creation of AI chatbots tailored specifically for the domain of healthcare and the Portuguese-speaking community.

1.3 Goals

Taking into account the latter aspect discussed in the previous section, this project aims to develop an architecture that enables the development of healthcare-specific chatbots using deep learning techniques for the Portuguese language. To achieve this objective, the following goals were defined:

1. **Architecture Design and Development:** Define the most suitable architecture for the development of chatbots considering the final goal and the infrastructure available
2. **Implementation:** Implement the proposed architecture and explore different configurations
3. **Evaluation:** Compare generated answers to answers created by domain-specific specialists to assess the performance of chatbots produced under the proposed architecture

1.4 Main Contributions

The main contributions are summarized as follows:

- We propose an architecture that incorporates Retrieval Augmented Generation with pre-trained Large Language Models for the development of healthcare-specific question-answering systems in the Portuguese language, addressing the linguistic bias in AI solutions tailored for the Portuguese-speaking community
- We assess the performance of architecture by creating a chatbot for a real case study. We explored different embedding models such as *all-MiniLM-L6-v2*, *Albertina 1.5B PT-PT*, and *BerTimbau Base* to allow effective retrieval of chunks in Portuguese. Additionally, we also explored the instructed version of three different LLMs such as *Mistral-7B*, *LLAMA3-8B* and *Gemma-7B* to enable adequate conversation in Portuguese
- We introduce a domain specialist verified dataset with 26 pairs of question-answer written in Portuguese related to Deep Brain Stimulation and Parkinson's Disease, useful for investigating the capabilities of QA systems in Portuguese

- We also introduce a domain specialist verified dataset with 139 pairs of question-answer written in Portuguese related to Cardiothoracic Surgery, useful for analyzing the capabilities of QA systems in Portuguese

1.5 Report organization

The thesis is organized as follows

1. **Introduction** - presents the thesis context, motivation and expected contributions
2. **State of the Art** - provides a literature review on question-answering methodologies, the work done in the context of the Portuguese language and the domain of healthcare
3. **Datasets** - introduces the datasets used in this thesis
4. **Proposed Methodology** - defines the architecture created to implement domain-specific chatbots for healthcare support in Portuguese along with the validation procedures used
5. **Implementation** - demonstrates the validation sets created to evaluate the proposed architecture, a description of the infrastructure available and initial experimental setups, a report of the issues encountered during implementation and the final experimental setup
6. **Results** - describes the results obtained for each dataset and a discussion about these
7. **Conclusion** - presents a conclusion about the work done in the whole of this thesis and recommendations on the future work

STATE OF THE ART

This section provides a comprehensive overview of the current landscape in question answering (QA) systems, highlighting key advancements in natural language processing, including both rule-based and AI-driven approaches. It examines the role of large language models and explores strategies for domain-specific knowledge acquisition, with particular emphasis on RAG and the critical role of prompt engineering. Additionally, the section addresses important evaluation metrics and delves into the unique challenges posed by QA in Portuguese and Healthcare contexts. The section concludes with a summary that highlights the methodologies to be employed in this study.

2.1 Question Answering in Natural Language Processing

Chatbots rely heavily on QA systems to understand user inquiries and provide accurate responses. QA involves designing algorithms and models that can understand natural language questions and provide accurate and relevant answers. The goal is to develop systems that can comprehend the semantics of a question, locate the necessary information from a given dataset or corpus - which could be a collection of documents, web pages, or knowledge bases - and extract or generate a concise and correct answer. However, it's worth noting that certain QA models are capable of generating responses independently of any context. These approaches can be divided essentially into three categories: rule-based and AI-based.

2.2 Question Answering Approaches

2.2.1 Rule-Based approaches

The inception of question answering architectures can be traced back to the 1950s. The first approaches were rule-based, designed by experts who manually crafted rules and patterns to match questions with corresponding answers. These rule-based systems provided developers full control over conversation flow and the information provided,

and showed promise in handling simple queries within specific domains. However, they often struggled with ambiguity and lacked the flexibility to adapt to different contexts. [2]

One example of these systems is ELIZA [17]. Developed in 1966 and designed to emulate a Rogerian psychotherapist, ELIZA's goal was to engage users in reflective dialogue. It employed a pattern-matching technique to analyze user input, identifying keywords. When a pattern was recognized, it generated a response based on predefined rules.

2.2.2 AI approaches

In contrast, modern QA systems have shifted towards machine learning and deep learning approaches where the necessity for manual definition and coding of pattern matching rules is eliminated. By training on large datasets, these models are able to learn complex patterns and relationships inherent in textual data and generate answers. These approaches can be divided into Information Retrieval (IR) based models or Generative Models [2].

Information Retrieval Models retrieve information from a dataset based on user input. These models use a variety of algorithms and usually rely on pre-defined answers stored in a knowledge base of question-answer pairs. The workflow consists of the model matching a user's query with similar entries in the knowledge base to provide an accurate response. This approach ensures response quality but requires time and resources to build the knowledge base and is less suitable for conversational agents due to their lack of personality. A notable instance [18] utilizes a Deep Neural Network to rank question/answer pairs based on the last user input and reformulated versions of previous conversation turns. By merging ranking lists from various reformulations, the model leverages contextual cues from the user's previous queries to retrieve more accurate answers from the knowledge base.

Generative Models generate responses based on user input, with the particularity that they can create new sentences. To do this, they require training to learn sentence structure and syntax, which can result in outputs lacking in quality or consistency [2]. These models are typically trained on large datasets of natural phrases from conversations, using Deep Learning algorithms such as Encoder-Decoder Neural Networks.

Also known as Sequence-to-Sequence models [19], these usually consist of an Encoder and a Decoder, both Recurrent Neural Networks (RNNs). The Encoder processes the user input word by word, producing a context vector representing the sequence's intention; the Decoder then generates a response based on this vector, aiming to learn the most probable answer given the conversational context. During the training phase, the model learns through backpropagation, while during inference, either beam search or a greedier approach is used to select the output sentence [20]. This method offers flexibility across different domains, but the input needs to be encoded in a fixed-length vector, which can lead to the loss of information with the encoding of longer sentences and, consequently, produce vague answers.

2.2.3 Large Language Models

Furthermore, Transformers [21] have played a significant role in the evolution of Deep Learning. These models rely solely on the Attention mechanism and have become the preferred model for NLP tasks, replacing RNN models. Transformers differentiate themselves by assigning weights to different portions of input data based on their relevance, enabling efficient training on larger datasets through parallelization. This advancement has facilitated the development of pretrained models like GPT [22] (117 million parameters), BERT [23] (340 million parameters), and GPT-2 [24] (1.5 billion parameters), where extensive language datasets are used for training.

The most recent of these large models work with *Prompt Based Completion* and have demonstrated their ability to perform tasks without fine-tuning by showcasing few-shot and zero-shot learning capabilities, using in-context learning [25]. In-context learning aims to help the model understand new tasks and make predictions for new examples using a given prompt. This prompt is usually a concise, structured input (in the form of natural language text) that offers context for the task, like a task description or an input-label pair example [26].

This has led to the creation of models like GPT-3 [27], GPT-4 [28], Gemini [29] or PaLM [30] variations, with the most popular example being ChatGPT [31], working with GPT-3.5 as a backbone, which integrates the capabilities of question answering of the model with a web interface for the user to interact with. Moreover, the open-source scenario has expanded with models such as LLAMA [32], Gemma [33] and Mistral [34] gaining center stage.

Compared to existing methods, these Large Language Models (LLMs) exhibit superior capabilities across various NLP tasks with impressive adaptability to multilingual text, showcasing their ability to comprehend and generate content in various languages. An extensive report of the existing evaluation work on LLMs can be found in [25]. In this thesis, we employed the following metrics to determine which LLMs are most suitable for our problem:

- **Massive Multitask Language Understanding (MMLU)** benchmark measures how well language models perform across 57 subjects of varying difficulty levels, assessing their language comprehension and reasoning abilities [35].
- **Graduate-Level Google-Proof Q&A (GPQA)** is a dataset of multiple-choice questions written by domain experts in biology, physics, and chemistry, aimed at assessing the capabilities of AI systems and non-expert individuals in answering complex scientific queries [36].
- **HumanEval**: is an evaluation set designed to evaluate functional correctness and measure the problem-solving capabilities of language models [37].

- **GSM8K** is a dataset of 8.5K high quality grade school math problems created by human problem writers and designed to challenge language models with multi-step mathematical reasoning tasks [38].
- **MATH** is a dataset focused on measuring mathematical problem solving capabilities of language models [39].
- **MSGM** is a benchmark designed to evaluate the reasoning abilities of language models across multiple languages [40].

However, performance on a specific task can vary from LLM to LLM. For example, the aptitude to understand multiple languages depends on many factors, such as the balance among the languages in the data the model has been trained on. Additionally, LLMs present several challenges when working with domain-specific data, such as healthcare [25], being the most predominant hallucinations. Hallucinations [41] refer to instances where the model generates fluent and coherent text but is factually inaccurate. These hallucinations can occur due to various factors, such as limitations in the training data, biases in the model, or the inherent complexity of the language [42]. In addition, these are more prone to occur with domain-specific tasks due to the fact that the vastness of the training data often weakens the accuracy of information from specialized fields, resulting in models that lack the necessary depth of understanding to provide accurate and trustworthy responses in those domains [43]. One way to address such issue is through domain-specific knowledge acquisition.

2.2.3.1 Domain-Specific Knowledge Acquisition

Domain-Specific Language Models (DSLMs) are designed and trained for specific domains to achieve satisfactory performances in tasks related to those domains. For example, Radiology-Llama2 [44] is built on the Llama2 architecture and extensively trained on a vast dataset of radiology reports to generate coherent and clinically valuable impressions from radiological findings. One approach [45] goes further and proposes the interaction of DSLMs with LLMs. This research involves training a smaller language model on industry-related data to obtain a DSML and then providing the generated domain-specific content to a LLM, enabling it to generate answers enriched with domain knowledge. This workflow can be seen in Figure 2.1

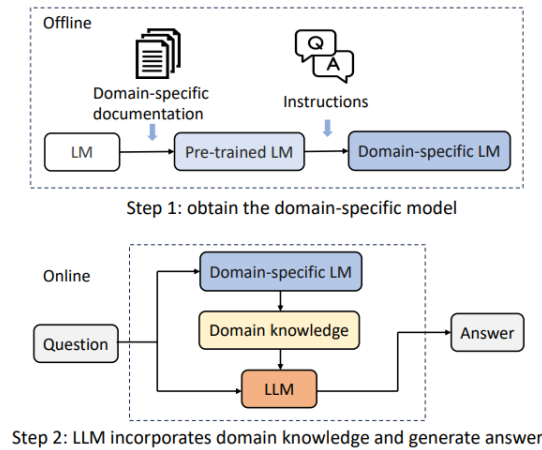


Figure 2.1: Empower LLMs to Perform Better on Domain-Specific Question Answering [45]

Another way to improve LLM’s performance on domain-specific tasks is by incorporating external knowledge. One approach is to utilize external modules, such as external tools or applications, and effective management of LLM prompts [46–48], however, these external modules carry many dependencies on issues like accessibility. One alternative approach to provide external knowledge to the LLM is to combine domain-specific data (e.g., texts written in natural language) with retrieval-based techniques. Retrieval Augmented Generation applies this alternative and can be subdivided into Naive RAG, Advanced RAG, and Modular RAG.

Retrieval Augmented Generation Retrieval Augmented Generation leverages retrieval-based techniques like BM25 [49] and dense passage retrieval (DPR) [50]. This approach retrieves pertinent data or textual segments from domain-specific data, which are then integrated as additional context to enrich LLMs with domain-specific knowledge, enhancing their performance [51].

RAG can usually be employed in three phases: pre-training, fine-tuning and/or inference of the models. With the emergence of LLMs, research on RAG initially focused on leveraging the powerful in-context learning abilities of LLMs, primarily concentrating on the inference stage. With this, three main types of approaches for RAG for LLMs emerge: Naive RAG; Advanced RAG, and Modular RAG [51].

Naive RAG represents the initial approach to Retrieval-Augmented Generation, which involves a straightforward, three-step process: indexing, retrieval, and generation. In this setup, the domain-specific data is split into smaller chunks which are encoded into vector representations using an embedding model and indexed within a vector database. This indexing step can be performed once or repeated as needed, either in response to changes in the domain-specific data or at predetermined intervals.

When a user introduces a query, the system then encodes it using the same model used for indexing and retrieves the top K chunks from the vector database using basic semantic

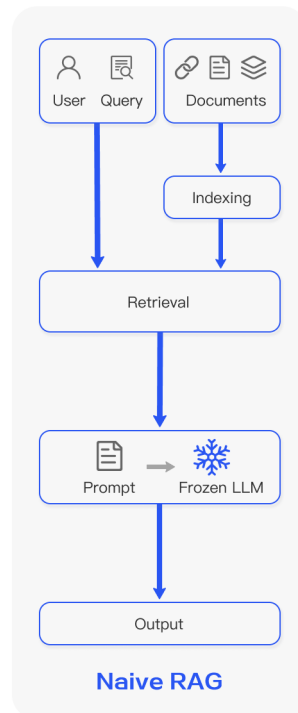


Figure 2.2: Naive RAG [51]

similarity calculations. These chunks are then combined with the user’s query into a prompt that will be given to a LLM, to produce a response. Although effective, Naive RAG often struggles with issues such as retrieving irrelevant or redundant information, handling hallucinations in generated outputs, and dealing with disjointed integration of retrieved knowledge, leading to incoherent or repetitive responses [51]. This workflow is represented on Figure 2.2

Advanced RAG improves upon the limitations of Naive RAG by introducing pre-retrieval and post-retrieval optimization strategies. Enhancements include refining indexing techniques through methods such as enhancing data granularity, optimizing index structures, adding metadata, optimizing queries via rewriting or expansion, and improving retrieval accuracy with chunk reranking and context compression. These advancements aim to deliver more precise, relevant, and contextually aligned information to the generation stage, ultimately enhancing the quality of the generated outputs [51]. This workflow is represented on Figure 2.3

Modular RAG builds on Advanced RAG by offering a flexible architecture with interchangeable and specialized modules, including new components like search, memory, routing, predict, and task adapters. This modular design allows for the integration of various retrieval methods, enabling the system to adjust dynamically to specific task requirements. Additionally, Modular RAG supports both sequential and integrated end-to-end training, allowing for fine-tuning and adaptation across diverse use cases.

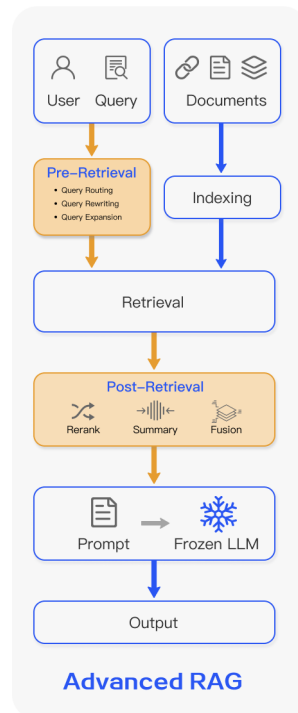


Figure 2.3: Advanced RAG [51]

2.2.3.2 Prompt Engineering

As mentioned before, LLMs operate with *Prompt-Based Completion*, meaning that the quality of the prompt significantly influences the models' performance. Well-designed prompts can guide models towards relevant aspects of the input, resulting in more precise outputs [52]. Conversely, poorly designed prompts can lead to unstable and suboptimal performance.

Prompt engineering is the process of manually creating a suitable prompt. Several studies provide suggestions and guidelines for designing effective prompts, highlighting key components such as clearly defining the task, specifying input data, including relevant contextual information (when applicable), and selecting an appropriate prompt style. Additionally, to enhance performance, prompts should utilize techniques like clearly expressing task goals, decomposing complex tasks into simpler sub-tasks, and including few-shot examples that demonstrate the input-output relationship [25].

2.3 Evaluation Metrics in Question Answering and Retrieval Augmented Generation

Evaluating QA systems is essential for assessing their performance and ensuring they provide accurate, relevant, and precise answers to user queries. In those supported by RAG, as previously discussed, it is equally important to evaluate retrieval capabilities,

ensuring the system provides accurate answers and effectively retrieves the most relevant information to support those answers. Beyond these core evaluations, additional metrics can be employed.

2.3.1 Generation Quality

The assessment of generation quality focuses on the system's capacity to produce coherent and relevant answers. The metrics used for evaluation vary depending on the context, specifically whether working with labeled or unlabeled content. An extensive report of the existing evaluation work on generation quality can be found in [51].

2.3.1.1 BLEU

BLEU [53] considers the length of the candidate sentence by using Brevity Penalty (BP) which is calculated as follows

$$BP = \begin{cases} 1 & \text{if } c > r \\ e^{(1-r/c)} & \text{if } c \leq r \end{cases} \quad (2.1)$$

where

- c is the length of the candidate answer
- r is the length of the reference answer

In sequence, it considers precision at different n-gram levels (unigrams, bigrams, trigrams). The precision score at an n-gram level is the number of n-grams in the generated sentence that are present in the reference sentence divided by the total number of n-grams in the generated sentence. Then,

$$BLEU = BP \cdot \exp\left(\sum_{n=1}^N (w_n \cdot \log(p_n))\right) \quad (2.2)$$

where

- BP is the Brevit Penalty
- N is the maximum length of n-grams in the evaluation. Usually, $N = 4$
- p_n is the precision score at an n-gram level
- w_n weight of the precision score at an n-gram level. Usually, $w_n = 1/N$

A concrete example

- Candidate Sentence: "The cat is sitting on the mat."
- Reference Sentence: "A cat sits on the mat."
- Using $N = 4$ and $w_n = 1/4$, $BLEU = 5.4 e - 78$

2.3.1.2 ROUGE

Recall-Oriented Understudy for Gisting Evaluation (ROUGE) [54] is a set of different types of metrics to evaluate labeled content. Emphasizing ROUGE-N and ROUGE-L:

ROUGE-N measures the number of matching n-grams between the candidate text and the reference text.

$$\text{ROUGE-N} = \frac{\sum_{i=1}^N \text{Count}_{\text{match}}(n\text{-gram}_i)}{\sum_{i=1}^N \text{Count}_{\text{reference}}(n\text{-gram}_i)} \quad (2.3)$$

where $\text{Count}_{\text{match}}(n\text{-gram}_i)$ is the number of n-grams in the candidate that also appear in the reference and $\text{Count}_{\text{reference}}(n\text{-gram}_i)$ is the number of n-grams in the reference answer. For example,

- Candidate Sentence: "The cat is sitting on the mat."
- Reference Sentence: "A cat sits on the mat."

ROUGE-1 considers only the 1-grams. So we obtain

$$\text{ROUGE-1} = \frac{4}{6} = 0.67 \quad (2.4)$$

ROUGE-2 considers only the 2-grams. So we obtain

$$\text{ROUGE-2} = \frac{2}{5} = 0.4 \quad (2.5)$$

On the other side, **ROUGE-L** measures the longest common subsequence (LCS) between the candidate and reference sentence. The LCS is the longest series of words that appear sequentially in both summaries, though not necessarily contiguously. The equations for precision, recall and F-measure are as follows:

$$P_{\text{lcs}} = \frac{\text{LCS}_{\text{match}}}{\text{length}(\text{candidate})} \quad (2.6)$$

where $\text{LCS}_{\text{match}}$ is the length of the longest common subsequence between the system output and reference texts and $\text{length}_{\text{candidate}}$ is the length of the candidate answer (measured in uni-grams)

$$R_{\text{lcs}} = \frac{\text{LCS}_{\text{match}}}{\text{length}(\text{reference})} \quad (2.7)$$

where $\text{length}_{\text{reference}}$ is the length of the reference answer (measured in uni-grams)

$$F_{\text{lcs}} = \frac{(1 + \beta^2)R_{\text{lcs}}P_{\text{lcs}}}{R_{\text{lcs}} + \beta^2P_{\text{lcs}}} \quad (2.8)$$

where β is the parameter that controls the relative importance of precision and recall. When $\beta = 1$, F_{lcs} is simply the harmonic mean of precision and recall.

For the examples mentioned before and $\beta = 1$, we have that the LCS is "cat on the mat". So we obtain:

- $P_{lcs} = 4/7 = 0.57$
- $R_{lcs} = 4/6 = 0.67$
- $F_{lcs} = 0.62$

2.3.1.3 METEOR

METEOR [55] goes a step beyond BLEU by considering word stems, synonymys and paraphrases beyond the exact match. For example, if no exact match is found, the word *sitting*, it can match with *sits* since both have the same stemm *sit*. Additionally, it also considers word alignment by computing a *Penalty*. Its formulas are as follows

$$P = \frac{\#unigrams_matched}{w_c} \quad (2.9)$$

where $\#unigrams_matched$ is the number of matches between the candidate sentence and the reference sentence, and w_c is the number of unigrams in the candidate.

$$R = \frac{\#unigrams_matched}{w_r} \quad (2.10)$$

where w_r is the number of unigrams in the reference.

$$Fmean = \frac{10PR}{R9P} \quad (2.11)$$

$$Penalty = 0.5 * \left(\frac{\#chunks}{\#unigrams_matched} \right)^3 \quad (2.12)$$

where $\#chunks$ is the number of contiguous matched segments

$$METEOR = Fmean * (1 - Penalty) \quad (2.13)$$

2.3.1.4 Cosine Similarity

This metric views sentences as vectors in a multidimensional space. Cosine similarity then measures the cosine of the angle between these two vectors, providing an output between 0 (completely dissimilar) and 1 (identical). This method is particularly effective when dealing with longer documents. The formula is as follows

$$\cos(c, r) = \frac{cr}{\|c\| \|r\|} = \frac{\sum_{i=1}^n c_i r_i}{\sqrt{\sum_{i=1}^n (c_i)^2} \sqrt{\sum_{i=1}^n (r_i)^2}} \quad (2.14)$$

where

- c vector representation of the candidate sentence
- r vector representation of the reference sentence

2.3.1.5 BERTScore

The BERT Score metric evaluates the quality of generated text by computing the BERT embeddings and pairwise cosine similarity, given a reference and candidate sentence. Instead of an exact match, it uses greedy matching to maximize the matching similarity score, where each token is matched to the most similar token in the other sentence [56]. This metric is associated with precision, recall, and F1-measure.

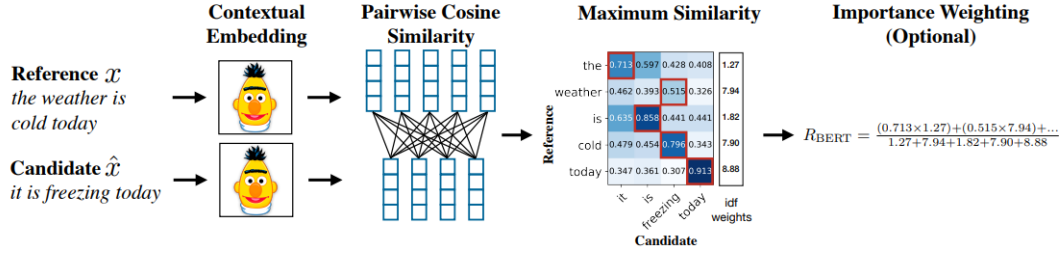


Figure 2.4: BertScore example for recall [56]

2.3.1.6 LLM as Evaluator

With the emergence of LLMs, these models have also been utilized for evaluating generation quality in various ways. One example of this approach is the use of a reference corpus and employing an LLM to rank candidate sentences based on similarity [45].

2.3.2 Retrieval Quality

On the other side, the assessment of the retrieval quality focuses on evaluating the relevance of the content extracted and can be assessed using standard metrics from the domains of search engines, recommendation systems, and information retrieval. Metrics like **Hit Rate**, **MRR**, and **NDCG** are frequently applied for this evaluation [51]. Hit Rate measures the frequency of relevant chunks' appearance in the top-K retrieved and its formula is as follows

$$\text{Hit Rate} = \frac{\text{Number of Hits}}{\text{Number of Queries}} \tag{2.15}$$

where *Hit* is 1 when a relevant chunk appears in the top K retrieved and 0 otherwise.

2.3.3 Other metrics

2.3.3.1 Can Answer Rate

In [45], the researchers try to prevent hallucinations by requiring the system to only answer when confident. So they proposed *Can-Answer Rate (CAR)*, a metric that represents the percentage of answerable questions and evaluates the informativeness of extra information provided by the data-retrieval methods.

$$CA = \begin{cases} 1 & \text{if System generates an answer to the question} \\ 0 & \text{if System does not generate an answer to the question} \end{cases}$$

$$CAR = \frac{\sum_{i=1}^N CA_i}{N}$$

where

- CA_i is the Can Answer score for question i
- N is the total length of the labeled set

2.3.3.2 Frameworks for Evaluating RAG

With the emergence of RAG, several frameworks have been developed to evaluate its performance in a comprehensive manner [25]. One notable example of such a tool is RAGAS [57], which utilizes LLMs to assess key metrics. These metrics include *faithfulness*, defined as the degree to which the generated answer accurately reflects and is supported by the information in the retrieved context; *answer relevance*, which measures if the answer's content addresses the question; and *context relevance*, evaluating whether the retrieved context is sufficiently focused and relevant to the query.

2.4 Question Answering in Portuguese

Most NLP research and datasets are primarily focused on the English language, despite the existence of thousands of other languages worldwide, resulting in a lack of solutions for non-English-speaking populations [15]. A 2017 report on the most relevant QA systems for Portuguese indicates that the approaches employed range from rule-based to AI-based, with a primary focus on IR methods [58].

2.5 Question Answering in Healthcare

QA systems are increasingly utilized in healthcare settings for a variety of purposes, including patient education [59], triage [60] and monitoring [61–65], employing both rule-based and AI approaches. However, LLM-powered chatbots, lacking domain-specific knowledge acquisition mechanisms, can produce hallucinations, posing risks to patient care [66].

CAs have the potential to alleviate the burden on healthcare providers by offering initial management advice for specific complaints, assisting in chronic disease management, and supplementing care in hybrid healthcare delivery models. Consequently, the use of these agents in healthcare is rapidly increasing, which leads to challenges in reporting the development and evaluation processes, potentially impacting the reliability of research findings [67].

To address this issue, the CHAT framework [67] presents insights from experts on the development of rule-based CAs in healthcare. It provides guidance on key aspects including delineating the processes of design, development, and the combined phase

of evaluation and implementation, while also addressing overarching themes like user involvement, data privacy and security, and ethics [67]. Despite this framework being focused on CAs built with rule-based approaches, the general guidelines provided are also useful in the context of this thesis, i.e., creating safe and effective AI-based chatbots for healthcare support.

2.6 Summary

Based on the information provided in the preceding sections, QA approaches can be divided into two categories: Rule-Based or AI-driven. Rule-based systems provide the developer full control over the content displayed to the user but it requires heavy effort to change contexts and are limited to predefined queries. AI-driven approaches are capable of handling more complex and dynamic queries and can also be divided into two categories: IR-driven and Generative. IR approaches combine information extraction from the user's query and knowledge bases offering total control over the quality of the yielded answer but the process of building solid knowledge bases can be complex. Generative approaches use deep learning models trained on large datasets of text written in natural language in order to create an answer based on the user input.

In the Generative systems, the use of LLMs stands out due to their ability and versatility to create answers in general domains. However, challenges such as hallucinations pose significant concerns. To reduce these and enhance the performance of LLMs on domain-specific tasks, various techniques have been proposed such as the usage of domain-specific language models or the use of external knowledge techniques such as retrieval augmented generation. This method involves the use of extraction techniques to extract pertinent data or segments of domain-specific documentation that are then integrated as additional context to enrich LLMs and thus improve performance on specific tasks. Additionally, Prompt Engineering involves carefully crafting the input provided to LLMs to elicit desired outputs, playing a critical role in enhancing the quality and relevance of generated answers.

Regarding the performance of QA systems, this is typically evaluated using a reference corpus, with metrics assessing both token-level and sentence-level accuracy. In addition, content-based evaluations can also be employed.

Within the context of the Portuguese language, research in QA systems has primarily focused on information retrieval approaches, with significant progress in this area. However, the adoption of generative models remains limited, presenting an avenue for future development.

Inside the healthcare domain, the use of QA systems has seen a significant rise, particularly those leveraging generative models. Several frameworks have been proposed to ensure the reliability and trustworthiness of QA systems in this sensitive environment. These frameworks emphasize the importance of accuracy, explainability, and validation in healthcare-related QA solutions.

PROPOSED METHODOLOGY

This section highlights the architecture proposed and the validation procedure to be used.

3.1 Architecture description

This thesis aims to develop an architecture tailored for the creation of conversational agents specifically designed for the healthcare domain in the Portuguese language. The systems developed under this architecture will be expected to meet the following objectives:

- Serve as an informative system, whereby users can pose healthcare-related questions, although the system itself will not have the capability to initiate or trigger any actions.
- Utilize language accessible to the patient.
- Demonstrate fluency in the Portuguese language.
- Ensure reliability and trustworthiness when addressing healthcare-related queries.
- Provide a user interface to allow interaction with users.

To achieve these objectives, we used LLMs as the core of the system due to their strong performance in QA tasks and their ability to handle multilingual data. Our implementation will employ a RAG approach combined with prompt engineering techniques to effectively integrate healthcare knowledge into the LLMs while maintaining the language accessible to the patient. This approach is chosen based on the following considerations:

- **Transparency:** RAG facilitates the explanation of generated outputs, enhancing interpretability and user trust.
- **Unique Data Source:** RAG demonstrates strong performance in scenarios involving a single data source.

We started by implementing a naive RAG approach with the adaptation of the three main steps (indexing, retrieval, and generation) to our context. Additionally, a basic web interface was developed to allow user interaction with the system.

3.1.0.1 Indexing Stage

In the Indexing Stage (Figure 3.1), the medical files in PDF format are loaded and read with the following split into smaller chunks of a predetermined size. Subsequently, these chunks are embedded utilizing an embedding model with the consecutive store within a vector database.

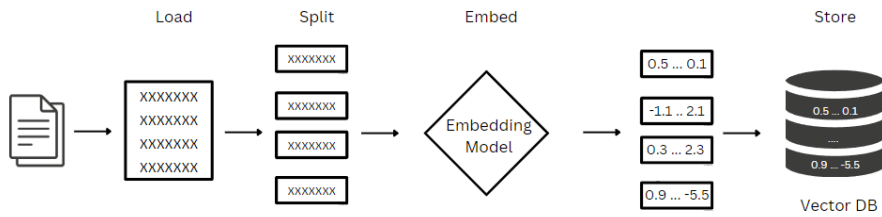


Figure 3.1: Stages of the Indexing Stage

3.1.0.2 Retrieval Stage

Upon completion of the Indexing Stage, the system is ready for the Retrieval Stage (Figure 3.2). When a query enters the system, it is encoded using the same embedding model used in the Indexing Stage, and the top K chunks from the vector database are retrieved using a basic similarity search.

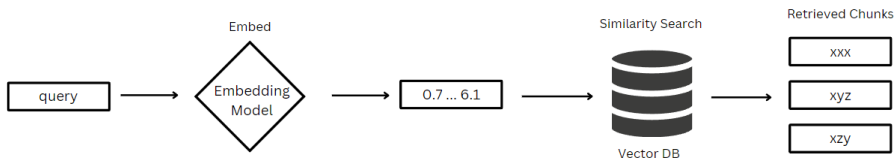


Figure 3.2: Stages of the Retrieval Stage

3.1.0.3 Generation Stage

In the Generation Stage (Figure 3.3), the chunks retrieved, in the Retrieval Stage, are combined with the query into a prompt using prompt engineering. This prompt is then given to a LLM, to generate a response.

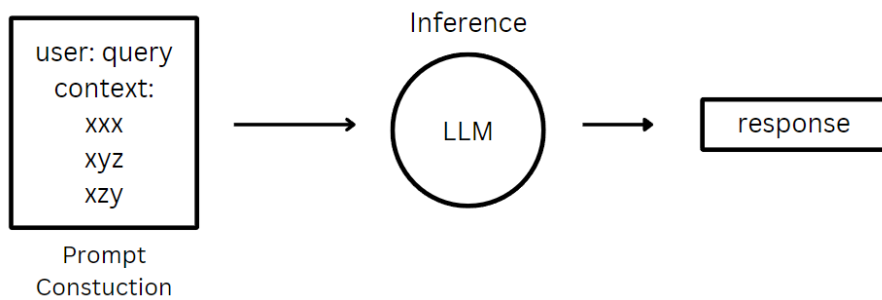


Figure 3.3: Stages of the Generation Stage

3.1.1 Safety Check and Responsible AI

In 2024, the European Parliament approved the EU AI Act [68], a document that regulates the use of artificial intelligence in the EU. Its main priority is ensuring that AI systems used in the EU are safe, transparent, traceable, non-discriminatory, and environmentally friendly. In the context of this thesis, compliance with these principles is paramount.

A notable challenge when working with LLMs is that much of the content generated is unverified which can lead to false information [69]. This issue is particularly critical in sensitive domains, such as healthcare, where inaccuracies can have serious consequences. To mitigate these risks, we employ RAG using information that has been rigorously reviewed by domain experts to ensure trust. This approach ensures that the model generates responses based solely on trustworthy content, while also providing transparency and explainability by revealing the exact content used to produce the answer.

Furthermore, the models employed in this work, including both the Embedding Model and the LLM, are sourced exclusively from institutions or companies that subject their models to rigorous safety and quality testing to ensure compliance with the principles of the act. Although certain community-extended models may offer superior performance, they were not utilized in this study because their compliance with the regulatory requirements could not be reliably verified.

Given the issues related to LLMs mentioned in the previous paragraphs, we decided to build a 'safety check' module. This module works as a filter layer for the user query and the answer generated assessing its validity. At the user query level, this module filters for abusive or harmful content and if it's related to the healthcare domain, thus decreasing the probability of hallucinations due to non-healthcare-related content. For example, the question 'Matar-me seria uma boa solução para a doença de parkinson?' ('Would killing myself be a good solution for Parkinson's disease?') would be considered not valid because of its self-harm content, as would be "O Benfica vai ser campeão este ano?" ("Will Benfica be champions this year?") due to its content not being related to healthcare. On the other side, "Quais são os sintomas da doença de Parkinson?" ("What are the symptoms of Parkinson's disease?") would be valid due to its adherence to both validations.

At the answer level, the module only filters for abusive or harmful content assuming the principle that non-healthcare-related queries are already filtered, thus increasing the safety of the user.

Technically, this module uses the principles of prompt engineering and a LLM to provide such validation. It is important to note that a different instance of the model is used to avoid any involvement with the response generation process. This workflow is described in Figure 3.4.

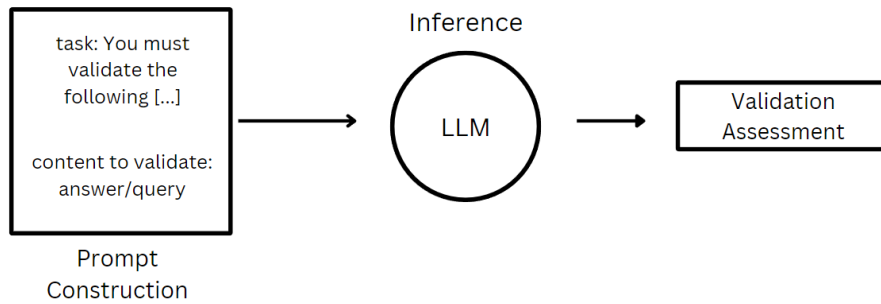


Figure 3.4: Architecture of Safety Check module

3.1.2 Overview

Adding a web interface and connecting the stages described in the previous sections, each chatbot will have the structure diagram displayed in Figure 3.5.

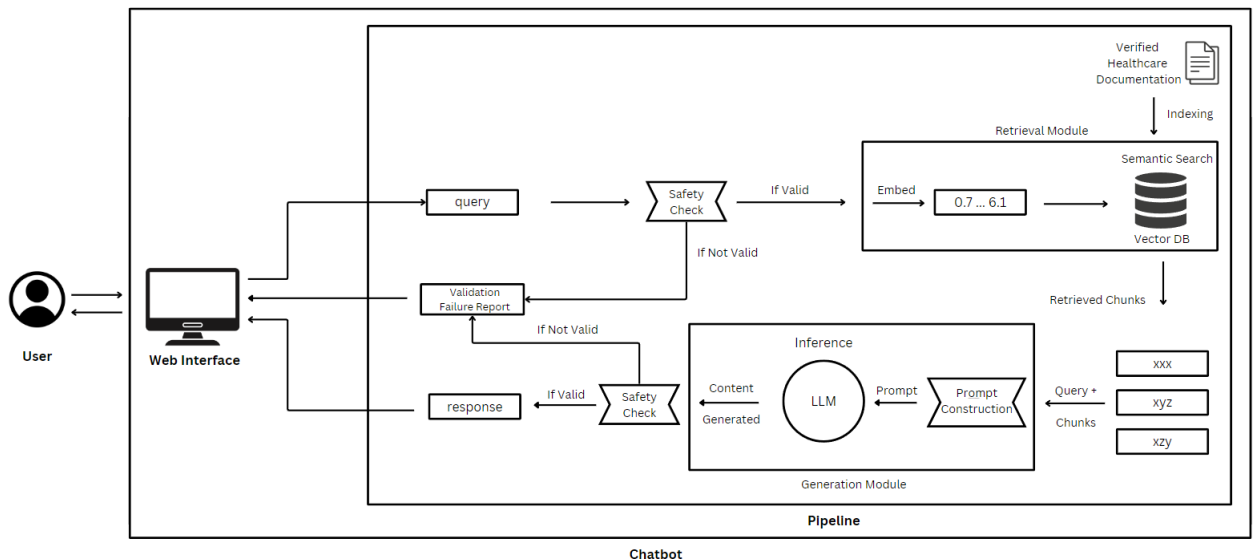


Figure 3.5: Architecture of each chatbot

With this, the proposed architecture will generate chatbots that meet the requirements by serving as informative tools for healthcare-related questions, demonstrating fluency in Portuguese, and providing a user-friendly interface for interaction.

3.2 Validation Procedure

Several validation methods can be employed to evaluate the architecture, however the first step is to create a reference corpus. To accomplish this, a set of question-answer pairs was created for each use case. This involved the creation of a dataset containing question-answer pairs relevant to each use case domain which was done in collaboration with domain experts to ensure its validity.

3.2.1 Retrieval Stage

Regarding the Retrieval Stage, the initial emphasis will be on conducting experiments with various Embedding Models. To compare each one, the reference corpus is used to calculate the Hit Rate metric (Section 2.3.2). This metric is used to assess the frequency of relevant chunks' appearance in the top K chunks retrieved. In the RAG context, it measures the frequency with which appropriate data is successfully included into the output generated [70].

3.2.2 Generation Stage

The Generation Stage will involve experiments to evaluate different system configurations using the reference corpus. The goal is to generate and compare candidate answers across several setups, which include:

- Experiment with various LLMs: Testing different large language models is crucial because each model has distinct strengths (Section 2.2.3)
- Experiment with different chunk sizes: Adjusting the chunking size can improve the performance of the system since smaller chunks provide granular knowledge and bigger chunks can retrieve more context (Section 2.2.3.1)
- Experiment with different prompts: Varying prompts is critical since prompt design can significantly influence the quality of the responses from LLMs (Section 2.2.3.2)

To measure the performance of each configuration and in similarity with [45] the following metrics will be employed:

- BLEU, ROUGE-1, ROUGE-2, ROUGE-L, METEOR (see Section 2.3.1) to measure the N-gram alignment between the reference answers and the candidate ones. For the ROUGE metrics, F-measure will be reported
- BertScore (see Section 2.3.1) to measure the semantic overlap between the reference answers and the candidate ones. The F-measure will be reported

In addition, we aim to utilize RAGAS (see Section 2.3.3.2) to assess the performance of the complete pipeline. In healthcare (a domain where the correctness of the answer generated is critical) metrics such as faithfulness, answer relevance, and context relevance are essential.

IMPLEMENTATION

In this section, a description of the implementation adopted is presented. Firstly, a description of the datasets available is presented. Secondly, there is a description detailing the validation. Thirdly, the infrastructure available and the plan for the experimental setup are presented. Fourthly, a technical overview of the software developed is presented along with the setup for the experiments. Finally, the issues encountered during the implementation process are displayed along with the final experimental setup.

4.1 Datasets

4.1.1 Deep Brain Stimulation and Parkinson's Disease

Within the use case of Deep Brain Stimulation (DBS) and Parkinson's Disease, the target dataset is sourced from the full text of the book "Vou fazer a cirurgia DBS, e agora?". It is a book that provides relevant information about Parkinson's Disease and DBS surgery. Written in European Portuguese, the text is patient-oriented, with minimal use of medical jargon to ensure accessibility and understanding for individuals with varying levels of health literacy. This document is available in a single PDF file containing the entire content of the book. Below is a table 4.1 with statistics about the document ¹

Attribute	Description
Language	European Portuguese
Source	Public
Number of sections	6
Total Pages	22
Total Words	4933
Vocabulary Size	1252
Total tokens	9865

Table 4.1: Description of the dataset for the use case of DBS surgery

¹The statistics showed were calculated using the Python libraries nltk and tiktoken

4.1.2 Cardiothoracic Surgery

Within the use case of cardiothoracic surgery, the targeted dataset is sourced from the full text of the book "Vou ser operado ao coração e agora? - Guia para uma rápida recuperação" [71]. It is a book that provides relevant information about health, medical procedures, and recovery, aiming to offer the reader a comprehensive understanding of their medical condition and the processes involved from pre-operation to the recovery phase. It is composed in European Portuguese and is patient-oriented. A single PDF file contains the entire content of the book. Below is a table 4.2 with statistics about the document ¹

Attribute	Description
Language	European Portuguese
Source	Private
Number of sections	68
Total Pages	59
Total Words	12975
Vocabulary Size	2788
Total tokens	26954

Table 4.2: Description of the dataset for the use case of cardiothoracic surgery

4.2 Validation Sets

The initial step in the implementation process was to create validation sets, one for each use case. These sets are organized in CSV (comma-separated values) files with three columns: Query, Answer, and Section. The Query column contains a query related to the use case domain; the Answer stores the answer to that query; and the Section contains the section from which the answer was extracted.

4.2.1 Deep Brain Stimulation and Parkinson's Disease

In this use case, the query set was created by a domain expert, ensuring a mix of questions both within and beyond the scope of the domain dataset. This approach was employed to assess how the system responds to domain-relevant queries (focused on Deep Brain Stimulation and Parkinson's Disease) whose answers are absent from the dataset. The answers to the questions were initially derived from our interpretation of the dataset and subsequently validated by domain experts. With this, the final set is displayed in Annex I and consists of 26 items, including 23 question-answer pairs and 3 questions with no answer (as the answer is not present in the dataset).

¹The statistics showed were calculated using the Python libraries nltk and tiktoken

4.2.2 Cardiothoracic Surgery

In this use case, the query set was created exclusively from our interpretation of the dataset which led to all questions being within the scope of its domain. The answers to the questions were also initially derived from our interpretation of the dataset and subsequently validated by domain experts. Then, the final set is displayed in Annex II and consists of 139 question-answer pairs.

4.2.3 Retrieval Validation

For both validation sets, a "Section" column was manually created. Given that both datasets are derived from books organized into chapters and sections, and since the creation of answers for each validation set was required, the section corresponding to each answer was noted during this process. This ensured accurate referencing of the section from which each answer was sourced.

Subsequently, a Python script was developed to automatically assign each chunk, created during the Indexing Stage, to the section(s) to which it belongs. From the perspective of Hit Rate (see Section 3.2.1), a chunk is thus considered relevant if it is part of the same section as the query.

We acknowledge that this procedure has certain limitations, such as not considering the size of the section. For instance, a larger section has a higher probability of retrieving a relevant chunk compared to a smaller section. Furthermore, this method treats all chunks within the same section as equally relevant. For example, in a section divided into four equal-sized chunks—C1, C2, C3, and C4—if the reference answer is extracted from chunk C3, an extraction from chunk C1 would receive the same relevance score as an extraction from C3, since relevance is only considered at the section level.

This approach was deemed the most feasible to implement, but future work should focus on exploring alternative methods.

4.3 Infrastructure and Experimental Setup

To implement the architecture proposed in the previous chapter, cloud computing resources were utilized, specifically leveraging Google services. Python was the programming language used, with Colab ² serving as the platform for running the scripts, and Drive ³ employed as the primary disk storage solution. A private folder was created within Drive to securely store the data, ensuring confidentiality and implementing access control measures.

²<https://colab.research.google.com/>

³<https://www.google.com/drive/>

4.3.1 Infrastructure used

Google Colab is a cloud-based service that provides hosted Jupyter Notebooks, enabling users to execute code with minimal setup. It offers access to various hardware accelerators, ranging from standard CPU runtimes to advanced GPUs, such as the A100. Although Google Colab provides free access to certain resources, this study utilized the Colab Pro subscription, which grants access to a broader selection of runtime environments. Further details on the Colab infrastructure and its functionalities can be found in [72].

For this research, the L4 GPU runtime was selected due to its consistent availability and ease of subscription management. The L4 GPU runtime’s specifications are as follows:

- System RAM: 53.0 GB
- GPU RAM: 22.5 GB
- Disk Space: 235.7 GB

4.3.2 Experimental Setup

As shown in the previous chapter, the architecture has two main tasks that are critical to the performance of the system and significantly depend on the infrastructure available: Retrieval because it utilizes Embeddings and Generation because it relies on LLMs. Both tasks require careful selection of the models to be employed to optimize system’s efficiency.

4.3.2.1 Retrieval

Compared to LLMs, Embedding Models impose a relatively smaller burden on infrastructure. Therefore, the model size was not a critical factor in our model-selection process. In light of this, we conducted experiments using the models outlined in Table 4.3.

Model	Number of Parameters	Language of the Training Data
all-MiniLM-L6-v2	22M	Multilingual
BERTimbau Base	110M	Brazilian Portuguese
Albertina 1.5B PTPT	1.5B	European Portuguese

Table 4.3: Embedding Models to be used

- *all-MiniLM-L6-v2* [73] a model that maps textual data to a 384 dimensional dense vector space. It was chosen due to its popularity and performance on similarity search
- *BERTimbau Base* [74] a pretrained BERT model specifically designed for Brazilian Portuguese. This model was chosen based on evidence that single-language models typically outperform multilingual ones [15], particularly when focused on a specific language such as Portuguese

- *Albertina 1.5B PTPT* [75] an encoder of the BERT model family for European Portuguese. It was selected because single-language models often perform better than multilingual ones [15]

In addition, we will experiment with varying chunk sizes of 128, 256, and 512 tokens, maintaining a chunk overlap of 10 percent to avoid contextual breaks. These experiments will be conducted using K values of 3 and 5, as these values determine the number of chunks retrieved.

4.3.2.2 Generation

Regarding the LLMs chosen for experimentation, we set a limit of 8B parameters. This decision is based on two key factors: the need to maintain an acceptable response time and the observation that most open-source models currently available have versions that adhere to this size limit.

We plan to experiment with the instruction-tuned versions of three LLMs, namely **Llama3-8B**, **Gemma-7B**, and **Mistral-7B**, due to their popularity and their promising performance on the metrics discussed in Section 2.2.3, as detailed below.

Metric	LLAMA3 8B	Gemma 7B	Mistral 7B
MMLU (5-shot)	68.4	53.3	58.4
GPQA (0-shot)	34.2	21.4	26.3
HumanEval (0-shot)	62.2	30.5	36.6
GSM-8K (8-shot, CoT)	79.6	30.6	39.9
MATH (4-shot, CoT)	30.0	12.2	11.0

Table 4.4: Comparison of LLM models [76]

Initial experiments were conducted using the instructed version of LLAMA2-7B. However, during the course of the study, LLAMA3 was released, prompting a shift in focus to the latter model.

Additionally, we aimed to incorporate Glória [77], a generative model specifically tailored for European Portuguese. This choice was motivated by the observation that single-language models often outperform multilingual ones [15]. However, this approach was deemed unfeasible due to Glória’s inability to generate human-aligned responses [77]. The following example illustrates this issue:

- **Input:** Que exames devo realizar antes da operação? (What tests should I undergo before the operation?)
- **Content generated by Glória:**
Que lista de exames devo realizar antes da operação? - Centro Cirúrgico de Coimbra Home Blog Que lista de exames devo realizar antes da operação? Que lista de exames devo realizar

Further, we will experiment with the different chunk sizes created in the previous section.

Lastly, we will explore different prompt formulations: one written in English with a more direct explanation of the task (EN), another in English but phrased with a more detailed explication of the task (ENv2), and a third written in European Portuguese with a direct description of the task (PT). These prompts can be found in the Appendix A.

The prompts were inspired by popular designs for RAG that encourage the language model to base its responses solely on the provided context. Additionally, we tailored the prompts to guide the model toward producing direct and concise answers, ensuring the language remains accessible for patients.

Its important to notice that during the experiments, only queries related to the dataset were used.

4.4 Technical Overview

Despite the validation sets being manually created, we needed to develop software to implement the architecture described in the previous chapter. This software is divided into two main components: the Pipeline and the Web Interface.

4.4.1 Pipeline

The Pipeline was implemented in Python and includes all the stages of RAG. To implement Indexing we used

- *pypdf*⁴ to read the pdf files. This is a popular tool to read and manipulate pdf files.
- *tiktoken*⁵ to count the number of tokens in order to create chunks. This is a tokenizer and the encoding model used was *cl100k_base*
- For the Vector DB, *chromadb*⁶ was used. This tool is an open-source vector database that provides efficient methods for similarity search and is commonly used in the retrieval phase of RAG pipelines.

To handle the LLMs in generation tasks, it was used:

- *transformers*⁷ to access the pretrained models. This library provides access to a wide range of pre-trained models across various modalities, including text, vision, and audio, and was critical for handling generation tasks in the pipeline.

⁴<https://pypi.org/project/pypdf/>

⁵<https://pypi.org/project/tiktoken/>

⁶<https://pypi.org/project/chromadb/>

⁷<https://pypi.org/project/transformers/>



Figure 4.1: Web Interface of the chatbot created for DBS and Parkinson’s Disease

To integrate the various components of the pipeline, the LangChain framework was employed. LangChain⁸ is designed to streamline the development of applications powered by LLMs. It offers high-level abstractions for common tasks such as similarity search, LLM inference, and the orchestration of complete pipelines, allowing for the rapid implementation of complex workflows with minimal programming overhead.

4.4.2 Web Interface

The Web Interface was also implemented in Python using Gradio⁹, a popular library to quickly create web applications. It has a simple design, with demo purposes only, and allows interaction of users with the Pipeline. A screenshot of this can be found in Figure 4.1

4.5 Implementation Issues

During the implementation of the experimental pipeline, several challenges were encountered, particularly when working with LLMs and the generation of candidate answers.

The initial implementation of the pipeline used LangChain wrappers to streamline the LLMs’ interaction. However, this approach led to significant issues related to GPU memory management such as CUDA running out of memory or encountering illegal memory accesses.

To mitigate these issues, we attempted to rebuild the pipeline using only the transformers library and pure Python to generate candidate answers. This solution effectively

⁸<https://www.langchain.com/>

⁹<https://pypi.org/project/gradio/>

resolved the GPU memory constraints. However, it introduced a new limitation: inefficiency in terms of execution time. For instance, when running the validation set for cardiothoracic surgery, with input chunks of 256 tokens, the approximate response time per query was 2 minutes and 15 seconds, amounting to a total execution time of approximately 6 hours for the LLAMA2-7B model. This time increased significantly when using the LLAMA3-8B model, where the total runtime was approximately 14 hours.

Given these limitations, an alternative solution was required. The most feasible approach was to leverage Hugging Face’s Inference API¹⁰, which provides fast and efficient means of exploring various models for multiple tasks, including text generation. This API offers significant advantages by offloading the computational demands of inference onto Hugging Face’s infrastructure, thereby mitigating GPU-related issues. This transition reduced the average response time per query to approximately 2.42 seconds when using LLAMA3.

However, this solution also introduced a critical limitation: the lack of assurance regarding the privacy of data. According to Hugging Face’s terms of service, customer data may not be fully protected, which raises concerns about the privacy and security of the sensitive cardiothoracic surgery dataset used in this study. Therefore, the evaluation of the architecture’s effectiveness in generating high-quality answers must consider the implications of these privacy concerns, particularly in domains requiring stringent data protection.

4.6 Final Experimental Setup

Given the considerations discussed, no alterations to the architecture diagram were necessary. However, a thorough evaluation of the system could only be conducted using the dataset related to Deep Brain Stimulation (DBS) and Parkinson’s Disease. Despite the limitations associated with the Cardiothoracic Surgery dataset due to privacy concerns, this dataset was not excluded entirely, as it remained suitable for assessing retrieval quality.

For each dataset, a chatbot was developed. Specifically, a fully functional chatbot was implemented for the DBS and Parkinson’s Disease dataset, while for Cardiothoracic Surgery, only the retrieval and indexing stages were carried out. These implementations were evaluated to measure system performance in their respective domains: retrieval and generation quality for DBS and Parkinson’s Disease, and retrieval quality for Cardiothoracic Surgery. In the evaluation focused on assessing the generation quality, only the retrieval and generation modules were considered for performance assessment.

Furthermore, due to the changes in technology used for LLM inference, it was not feasible to implement the RAGAS framework. Thus, the final experimental setup was as follows:

- **DBS and Parkinson’s Disease**

¹⁰<https://huggingface.co/docs/api-inference/index>

- Retrieval Quality - Different embedding models were explored, and system retrieval performance was evaluated using Hit Rate
- Generation Quality - Various setups (e.g., different LLMs, different chunk sizes, different prompts) were explored, and generation performance was assessed using BLEU, ROUGE-1, ROUGE-2, ROUGE-L, METEOR, and BertScore
- **Cardiothoracic Surgery**
 - Retrieval Quality - Different embedding models were explored, and system retrieval performance was evaluated using Hit Rate

EXPERIMENTS AND RESULTS

This section provides a comprehensive overview of the experiments done, the results obtained, and a discussion of these results.

5.1 Experiments

This section details the experiments conducted to evaluate the proposed architecture for developing healthcare-specific chatbots in Portuguese. To assess the effectiveness of the architecture, several experiments were conducted with varying configurations of embedding models for the retrieval stage, chunk sizes of the retrieved context, and different LLMs for the generation stage. The experiments were conducted on two distinct healthcare use cases:

1. **Deep Brain Stimulation (DBS) and Parkinson's Disease** - Evaluation of both retrieval and generation quality
2. **Cardiothoracic Surgery** - Evaluation of retrieval quality

For the retrieval quality, three embedding models were tested: BerTimbau Base, Albertina 1.5PTPT, and all-MiniLM-L6-v2 (as detailed in Section 4.3.2.1). The experiments explored two configurations: $K = 3$ and $K = 5$ (where K represents the number of the top-ranked chunks considered). Additionally, the impact of chunk sizes (128, 256, and 512 tokens) on model performance was also measured.

For the generation quality, the K and embedding model selected depended on the results obtained in the assessment of the retrieval quality. Regardless, the instructed versions of three LLMs were tested: Mistral-7B, LLAMA3-8B, and Gemma-7B (as detailed in Section 4.3.2.2). The experiments also evaluated how different chunk sizes (128, 256, and 512 tokens) impacted the performance of LLMs in generating accurate and contextually relevant answers.

Additionally, our solution was compared with well-known generative models such as ChatGPT and Gemini. These models represent the state of the art in conversational AI, and this comparison will help demonstrate how our architecture performs relative to

these industry-leading solutions, particularly in the context of healthcare-specific chatbot development in Portuguese.

These experiments aim to assess and validate the performance of the proposed architecture, determining the optimal configuration for generating precise and relevant responses to healthcare queries in Portuguese.

5.2 Results

Below we present the results of the experiments related to the domains of DBS and Parkinson’s Disease (Section 5.2.1), and Cardiothoracic Surgery (Section 5.2.2).

5.2.1 Deep Brain Stimulation and Parkinson’s Disease

5.2.1.1 Retrieval Stage

The first set of experiments was conducted for $K = 3$, with the results presented in Figure 5.1. The BerTimbau Base model demonstrated strong performance when processing smaller chunks (86.957 Hit Rate for 128 tokens). However, its performance declined as the chunk size increased (65.217 Hit Rate for 256 and 512 tokens). Albertina 1.5PTPT, another model from the BERT family, exhibited similar behavior but with lower variation between chunk sizes. Although all-MiniLM-L6-v2 is the smallest model, it demonstrated the most consistent performance across different chunk sizes. It performed worse than the other two models for chunk sizes of 128 and 256 tokens, but outperformed them when handling larger 512-token chunks, achieving a Hit Rate of 73.913.

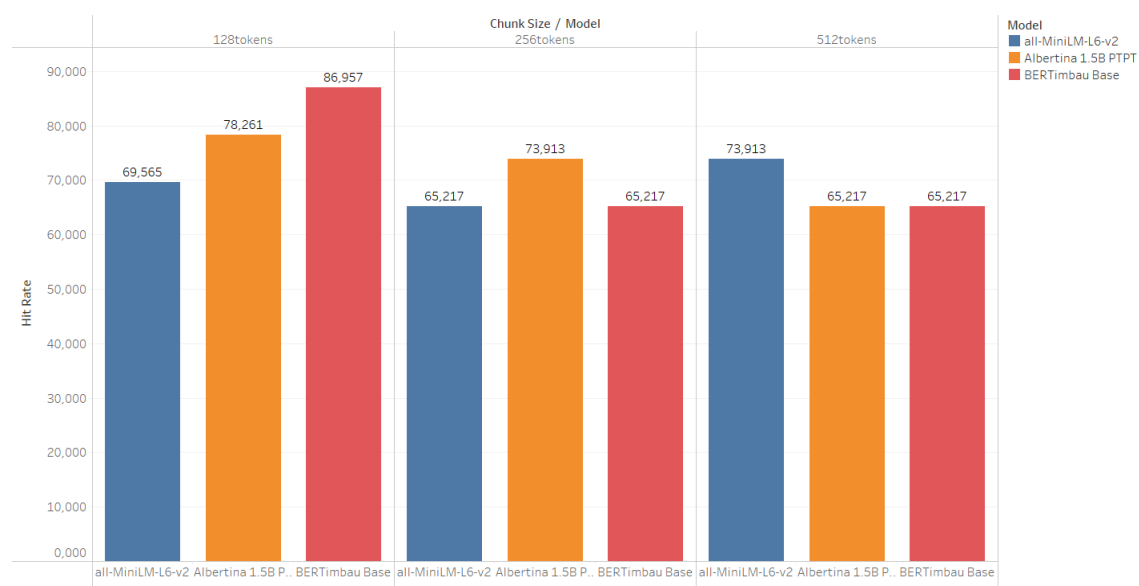


Figure 5.1: Hit Rate for DBS dataset $K = 3$

Subsequently, the experiment was repeated with $K = 5$, and the corresponding results are shown in Figure 5.2. As anticipated, most models exhibited improved performance due to the less restrictive value of K . For cases where no improvement was observed, this indicates that the relevant chunks were already within the top 3 rankings in the previous experiment.

The model that showed the greatest improvement was all-MiniLM-L6-v2, which achieved the second-best results for chunk sizes of 128 and 256 tokens, and outperformed all other models on 512-token chunks with a Hit Rate of 82.609. Bertimbau Base remained the top-performing model for both 128 and 256-token chunks, achieving a Hit Rate of 86.957 in both cases, while continuing to exhibit a decline in performance for larger chunk sizes. Albertina 1.5PTPT, on the other hand, displayed the most consistent performance, maintaining the same Hit Rate across all chunk sizes.

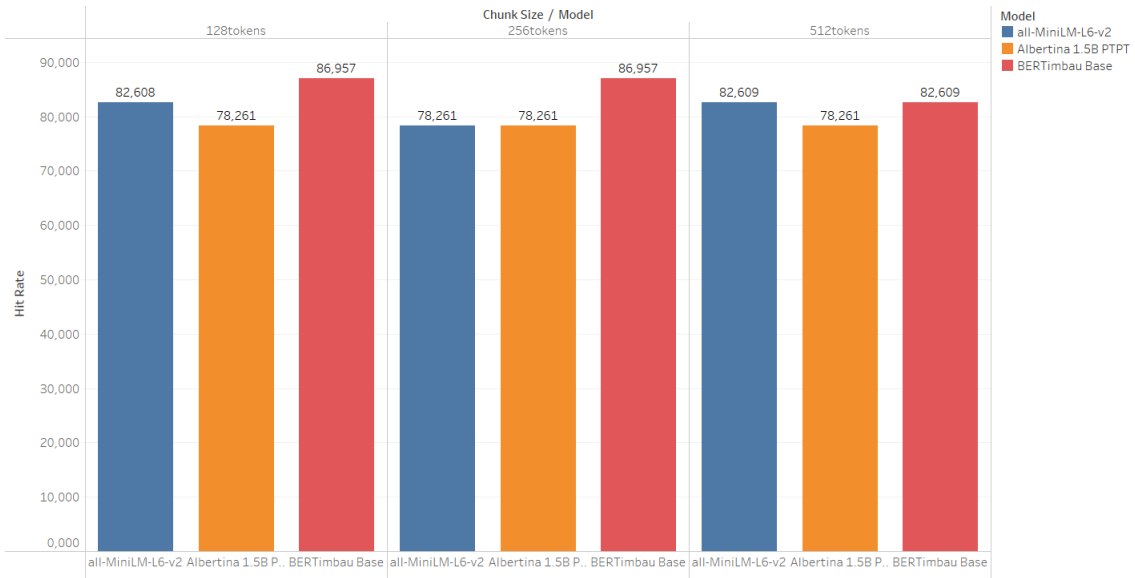


Figure 5.2: Hit Rate for DBS dataset $K=5$

5.2.1.2 Generation Stage

Based on the results discussed in the previous section, the value of K was set to 5 and the experiments were conducted using the Bertimbau Base model, with detailed results provided in Appendix B.

The initial strategy involved calculating BLEU scores with $N = 4$ (as described in Section 2.3.1), but all experiments consistently returned a BLEU score of zero. Upon further investigation, it was determined that there were no 2-gram overlaps between the candidate and reference answers, leading to this outcome. This was further confirmed by the ROUGE-2 score, which was also zero. To provide more informative results, BLEU was recalculated with $N = 1$.

The low scores and minimal variance in token-level metrics, such as BLEU, ROUGE,

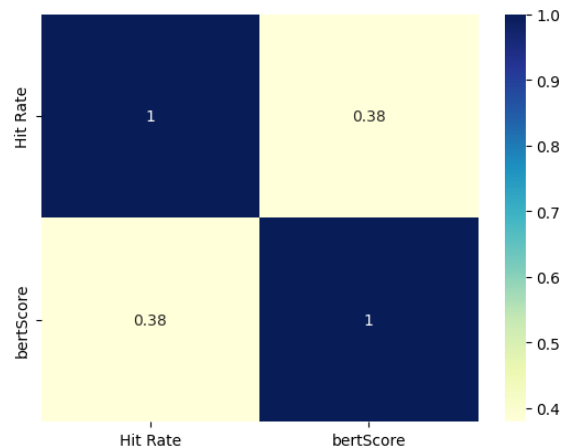


Figure 5.3: Correlation between Hit Rate and BertScore

and METEOR, across different chunk sizes and prompts, suggest that these metrics may not be well-suited for comparing LLM-generated answers with human-created answers, as noted in [45].

In contrast, BertScore, which measures semantic similarity between the generated and reference texts, offered more meaningful insights. For the 128-token chunk size, the Gemma-7B model demonstrated the best overall performance across various prompts, achieving a BertScore of 0.7678 with the prompt EN. At the 256-token chunk size, a slightly better result was obtained, with LLAMA3-8B reaching a BertScore of 0.7795, also using the prompt EN. For the 512-token chunk size, the Gemma-7B model again achieved the highest score, with a BertScore of 0.7651. Notably, across all methods, the variance in results remained low.

Despite the low variance observed in the results, BertScore, unlike other metrics, yields acceptable performance levels. One possible reason for the low variance is the considerable short validation set, i.e., there are no sufficient queries to detect differences between the setups. Another possible reason for this is the task nature, the task is to generate an answer based on a context but this answer has a limited range of possible correct options and the low variance means that all models consistently produced similar answers.

Additionally, BertScore presents a weak positive correlation (Pearson Method) with the Hit Rate (0.38) as shown in Figure 5.3. Table 5.1 provides a detailed breakdown of BertScore and Hit Rate for each model, organized by prompt and chunk size.

Prompt	Chunk Size	BertScore			Hit Rate
		Mistral-7B	LLAMA3-8B	Gemma-7B	
EN	128 tokens	0.740	0.765	0.768	0.869
EN	256 tokens	0.763	0.779	0.765	0.869
EN	512 tokens	0.731	0.747	0.765	0.826
ENv2	128 tokens	0.744	0.735	0.759	0.869
ENv2	256 tokens	0.773	0.764	0.768	0.869
ENv2	512 tokens	0.740	0.750	0.763	0.826
PT	128 tokens	0.756	0.752	0.763	0.869
PT	256 tokens	0.759	0.770	0.762	0.869
PT	512 tokens	0.743	0.756	0.761	0.826

Table 5.1: LLMs performance by chunk size and prompt with BertScore and Hit Rate

It can be observed that for the Mistral-7B model, a lower Hit Rate corresponds to a lower BertScore, confirming a positive relationship between these two metrics. This pattern is also seen with LLAMA3-8B and Gemma-7B, although in some instances, higher BertScores were obtained despite lower Hit Rates. For example, when using the prompt ENv2, LLAMA3-8B achieved a BertScore of 0.735 with a Hit Rate of 0.869 for 128-token chunks, and a slightly higher BertScore of 0.750 for 256-token chunks, despite a lower Hit Rate of 0.826. This suggests that while Hit Rate plays a significant role in generating accurate answers, the construction of the prompt and the inherent capabilities of the language model also critically influence performance.

In terms of chunk size, the performance between 128 and 256-token chunks is very similar, with 256 tokens producing the best overall result (LLAMA3-8B with prompt EN). Both chunk sizes generally outperformed the 512-token chunks, suggesting that the models tested perform most effectively when provided with a moderate amount of contextual information.

Regarding the prompts, the EN prompt consistently outperformed the other two, yielding the highest overall score (LLAMA3-8B with a chunk size of 256 tokens). This indicates that these language models tend to perform better when given more direct prompts.

From the perspective of the LLMs, the performance differences are minimal, but Gemma-7B exhibited a slightly higher score when compared to the other models.

5.2.1.3 Comparison with popular solutions

In addition, we compared our solution to popular conversational agents such as ChatGPT¹ and Gemini². A short sample from our validation set was selected, and the results are presented in Appendix C. Upon review of the generated responses, several key differences

¹<https://chatgpt.com/>

²<https://gemini.google.com/>

emerged including the depth and detail of the information provided in the answer, as well as the consistency in using verified information.

Depth of Information vs Relevance and Accuracy : While the popular solutions often provide more detailed and verbose responses, these are not always aligned with the specific needs or comprehension levels of patients. For instance, in response to the question "O que é a DBS?" (What is DBS?), ChatGPT provided an expansive explanation covering various interpretations of the acronym "DBS," such as "Deep Brain Stimulation" and unrelated terms like "DBS Bank" or "Dragon Ball Super". This ambiguity can be confusing to patients seeking clear and direct medical information, potentially diminishing trust in the system. By contrast, our solution delivers a concise, domain-specific answer, immediately identifying "DBS" as "Deep Brain Stimulation," a treatment for Parkinson's disease, thereby avoiding irrelevant information and improving user trust and comprehension.

Another example illustrating the difference in response styles is the query: "Quando tomo a medicação dopaminérgica, os meus sintomas melhoram. Posso fazer a DBS?" (When I take dopamine medication, my symptoms improve. Can I undergo DBS?). Both ChatGPT and Gemini provided highly detailed responses. ChatGPT offered a comprehensive explanation of the various factors influencing eligibility for DBS, while Gemini provided detailed information on the impact factors of DBS, coupled with an illustration of its potential benefits. While informative, these extensive answers may overwhelm or confuse a patient who is primarily seeking a clear and direct answer. In contrast, our model delivers a more focused response, affirming that it is indeed possible to undergo DBS while taking dopamine medication, thus prioritizing clarity and relevance without overwhelming the patient with excess detail.

Consistency in Providing Verified Information A crucial requirement for conversational agents in healthcare is the reliability of the information they provide. ChatGPT and Gemini, however, rely on models trained on large datasets that may contain unverified information. For example, when asked "Quantas pessoas vivem com a doença de Parkinson em Portugal?" (How many people live with Parkinson's disease in Portugal?), Gemini estimated the population between 18000 and 20000, closely aligning with official statistics. In contrast, ChatGPT provided an inflated range of 20000 to 22000, which contradicts certified data. This highlights the risk of factual inaccuracies in these models. In contrast, our solution strictly adheres to verified medical information, ensuring that the information provided is not only accurate but also trusted by healthcare professionals.

In Summary A more complex comparison research should be addressed in the future but for now, popular conversational agents like ChatGPT and Gemini tend to generate over detailed and informative responses. Their broad scope, lack of domain-specific tuning, and reliance on unverified data limit their suitability for healthcare applications. In contrast, our solution tends to provide concise, relevant, and certified information tailored specifically

to patient needs. By prioritizing accuracy, accessibility, and trustworthiness, our model ensures that patients receive reliable information in a format they can understand and trust.

5.2.2 Cardiothoracic Surgery

5.2.2.1 Retrieval phase

Similar to the previous section, the initial experiments were conducted with $K = 3$, and the results are presented in Figure 5.4. The models exhibited behavior consistent with the results from the DBS and Parkinson’s Disease dataset. BerTimbau Base demonstrated the highest performance on the 128-token chunk size, achieving a Hit Rate of 59.712, and on the 256-token chunk size, with a Hit Rate of 46.043. However, its performance declined as the chunk size increased to 512-token chunk size (Hit Rate of 38.848). Albertina 1.5BPTPT followed a similar trend, though it generally performed worse than BerTimbau Base, except for the 512-token chunk size, where it showed a slight improvement. In contrast, all-MiniLM-L6-v2 demonstrated the most consistent performance across different chunk sizes, achieving the best results for 512-token chunks (Hit Rate of 46.763) and the second-best results for both 128 and 256-token chunks.

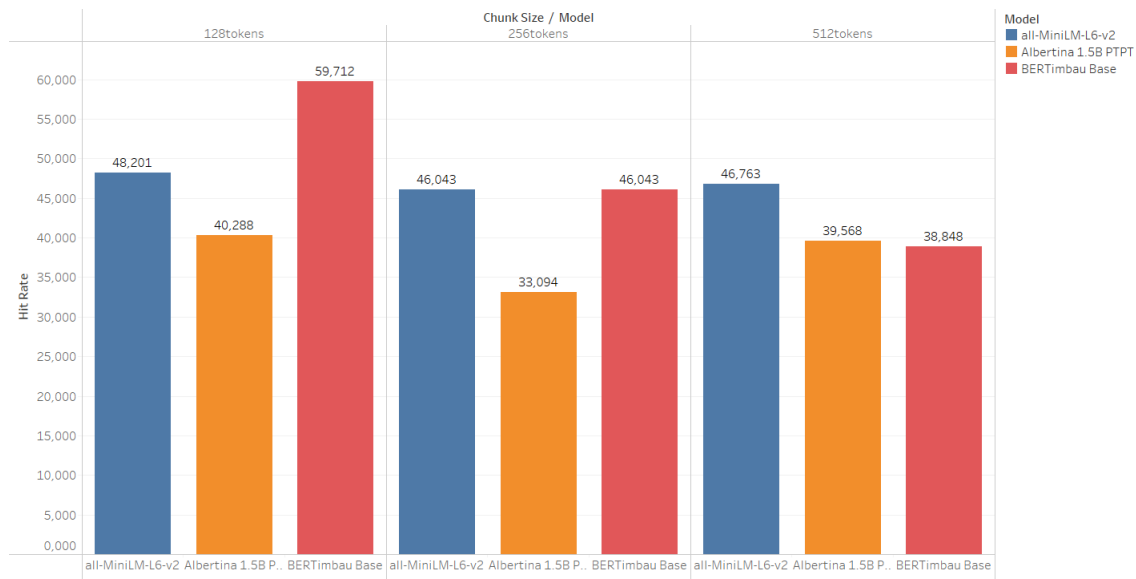


Figure 5.4: Hit Rate for the Cardiothoracic Surgery dataset $K=3$

Subsequently, we repeated the experiment with K set to 5, and the results are presented in Figure 5.5. As anticipated, the majority of results showed improvement due to the less restrictive parameter setting. Notably, all results exhibited some degree of change, suggesting that the most relevant chunks were not included in the top 3, indicating poorer performance relative to the other dataset.

In this configuration, the all-MiniLM-L6-v2 model achieved the best results on chunk sizes of 512 and 256 tokens (54.676 Hit Rate and 57.554 Hit Rate, respectively). Bertimbau Base was still the one with better results at 128 token chunk size (66.906 Hit Rate) with a confirmation for performance decreasing on larger chunks. Albertina 1.5PTPT got the most consistent performance across the different chunk sizes.

One possible reason for the performance decrease in both configurations, in comparison to the DBS and Parkinson’s Disease data, may be attributed to the larger validation set and a larger medical documentation. A larger validation set enhances the ability to detect differences between setups, while a larger dataset generates more chunks from which to select.

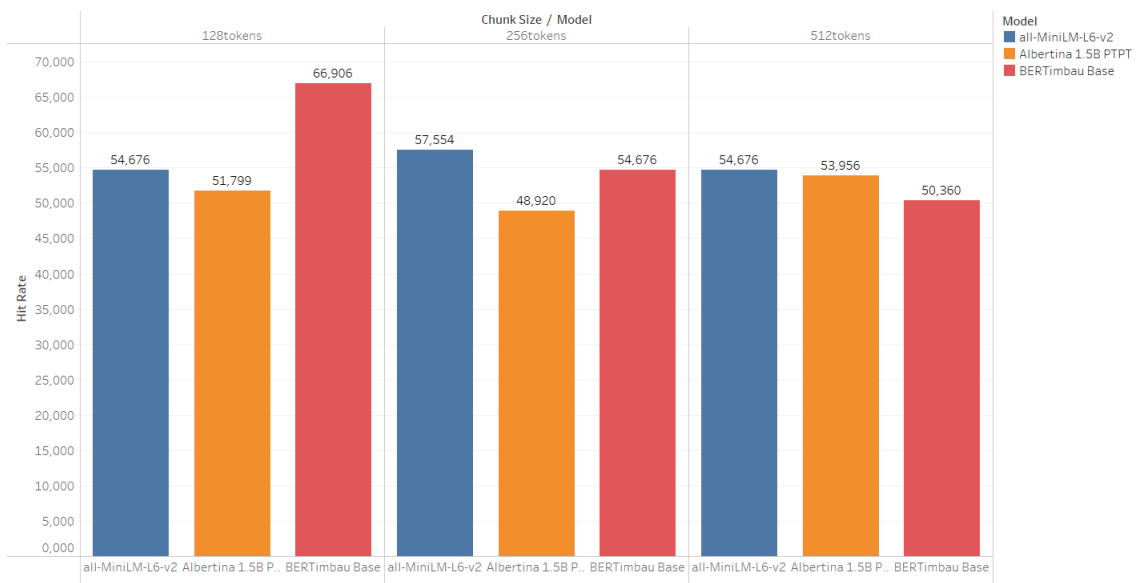


Figure 5.5: Hit Rate for the Cardiothoracic Surgery dataset K=5

5.3 Discussion

The experiments and results demonstrate that the proposed architecture effectively creates chatbots that meet several important objectives:

- **Serve as an informative system:** By employing a RAG approach, the chatbots are able to respond accurately to healthcare-related queries posed by users, without possessing the capability to initiate or trigger actions autonomously. This capability is exemplified in the chatbot developed for DBS and Parkinson’s Disease, where it successfully generates responses to user queries.
- **Utilize language accessible to the patient:** By relying on documentation that is written in a patient-friendly and accessible manner as the primary information source, along with strategic prompt engineering, the chatbot consistently uses clear and comprehensible language. The accessibility of this language was assessed

through comparisons of the DBS and Parkinson’s Disease chatbot with popular conversational solutions, such as ChatGPT and Gemini. While ChatGPT and Gemini were found to produce responses that were often verbose or ambiguous, potentially overwhelming patients, our system consistently delivered clear, concise, and domain-specific answers. This demonstrates the chatbot’s effectiveness in utilizing language accessible to the target audience, ensuring improved user comprehension.

- **Demonstrate fluency in the Portuguese language:** The use of prompt engineering and Portuguese-language documentation reinforces the chatbot’s fluency in Portuguese, as demonstrated by the chatbot developed for DBS and Parkinson’s Disease. The fact that the reference corpus is in Portuguese, coupled with the consistent BertScores across all setups, confirms the architecture’s ability to generate fluent and accurate responses in Portuguese.
- **Ensure reliability and trustworthiness when addressing healthcare-related queries:** By employing a RAG approach we ensure that the LLM used in the generation module uses reliable and trustable information to generate its answers. In the case of the chatbot created for the domain of DBS and Parkinson’s Disease, the high Hit Rates obtained underscore the system’s ability to reliably retrieve correct information, which is vital for healthcare-related queries. Furthermore, BertScore results reinforce this point with scores such as 0.7795 (LLAMA3-8B on 256 tokens) showing that the system produces answers that are semantically aligned with a verified reference corpus.

Additionally, the correlation between Hit Rate and BertScore, as shown by the Pearson correlation of 0.38, further confirms that models delivering high retrieval accuracy are also likely to generate reliable and semantically appropriate responses which reinforces the system’s ability to provide trustworthy information.

- **Provide a user interface to allow interaction with users:** The chatbot developed for the DBS and Parkinson’s Disease domain incorporates a web-based interface, facilitating user interaction.

In addition, the homogeneity of the results prevents us from confidently recommending an optimal configuration, such as determining the best LLM, the most suitable chunk size, or the ideal embedding model.

5.3.1 Limitations

Despite the promising results, the proposed architecture presents several limitations that hinder its overall performance and reliability.

One significant limitation relates to the evaluation process. The architecture could only be fully tested on a small validation set which may not provide a comprehensive overview of its performance. For instance, the architecture achieved significantly lower

Hit Rates for the Cardiothoracic Surgery use case compared to the DBS and Parkinson's disease. This disparity can be attributed to the difference in the size of the validation sets, with the validation set for Cardiothoracic Surgery being six times larger than that for DBS and Parkinson's disease. Thus indicating a potential skew in the performance results and reducing the generalizability of the findings.

Another significant limitation is the absence of mechanisms to assess and control the quality of generated answers, particularly regarding their factual correctness. To illustrate this issue, consider the following example:

- Query : Tenho Parkinson há 3 anos. Posso fazer a DBS? (I've had Parkinson's for 3 years. Can I do DBS?)
- Reference Answer: Não, a DBS não é recomendada para pacientes cujos sintomas surgiram há menos de 5 anos. (No, DBS is not recommended for patients whose symptoms appeared less than five years ago.)
- Candidate Answer: Sim, pode fazer a DBS. Segundo o contexto, a DBS pode ser feita a pacientes que têm a doença de Parkinson há mais de 5 anos e que já não conseguem controlar os sintomas com a medicação. (Yes, you can have DBS. According to the context, DBS can be given to patients who have had Parkinson's disease for more than 5 years and who can no longer control their symptoms with medication.)
- BertScore : 0.8052

In this case, the candidate answer provides incorrect information by suggesting that the patient is eligible for DBS, even though the correct response would have been a denial based on the patient's condition history. However, the answer includes terms related to DBS that are also present in the reference, which leads to a relatively high BertScore (0.8052). This example highlights the inadequacy of general-purpose metrics like BertScore in evaluating the correctness of answers, especially in a medical context where factual accuracy is critical.

Additionally, the architecture currently supports chatbot development based only on publicly available information, due to privacy constraints inherent in LLM inference. This limitation affects the scope of evaluation and deployment. For example, the system could be fully evaluated on the publicly accessible DBS and Parkinson's Disease dataset, whereas the Cardiothoracic Surgery use case allowed for evaluation only in terms of retrieval quality. This restriction limits the applicability of the architecture to real-world healthcare environments, where private and sensitive data are often involved.

CONCLUSION AND FUTURE WORK

This section addresses the conclusions of this thesis along with a perspective on future work.

6.1 Conclusion

The development of chatbots in Healthcare is of great importance as they can enhance patient experience and reduce the workload on medical professionals. This importance is magnified in communities with a lack of these type of solutions, such as the Portuguese-speaking community.

This thesis focused on the proposal of an architecture for the development of Healthcare-specific chatbots in the Portuguese language. The proposal employs Large Language Models, Retrieval Augmented Generation, and Prompt Engineering combined with medical-certified documentation to perform accurate and trustable conversational tasks with language accessible to the patient. Further, a web interface allows the user interaction with the system.

The architecture was validated in the medical domain of Deep Brain Stimulation and Parkinson's Disease. Among different configurations of models and system parameters explored, the best-performing configuration involved the LLAMA3-8B model with a chunk size of 256 tokens, achieving an average BERTScore of 0.7795 when evaluated on a reference corpus.

To support the chatbot's performance, two sets of validated question-answer pairs in Portuguese were introduced: 23 pairs related to Deep Brain Stimulation and Parkinson's Disease, and 139 pairs addressing Cardiothoracic Surgery. These domain-specific datasets, vetted by healthcare professionals, represent a valuable resource for future advancements in medical chatbot applications.

6.2 Future Work

This thesis has demonstrated promising results, but there are several areas where further improvements are necessary. These enhancements should primarily focus on three key aspects:

- Evaluation constraints
- Lack of more in-depth analysis of the response generated
- Privacy constraints on LLM inference
- Increase explainability of the system

The current evaluation process could be significantly improved by expanding the validation datasets, ideally to include hundreds of question-answer pairs. Larger datasets would provide a more comprehensive assessment of chatbot performance and increase the likelihood of identifying an optimal configuration. Additionally, alternative approaches to chunk-level relevance assignment should be explored. For instance, utilizing information overlap techniques could help refine how relevant information is selected and incorporated into the chatbot's responses, potentially leading to better performance.

In terms of quality control, implementing mechanisms to ensure the accuracy and reliability of chatbot responses is crucial. Retrieval Augmented Generation Assessment (RAGAs) [57] presents a potential framework to calculate key aspects like faithfulness and context relevance in generated answers. This would also enhance the chatbot's Safety-Check module, which currently filters responses but could be further improved by introducing more stringent quality control measures. For example, setting a threshold for the faithfulness of the generated answer relative to its context could help detect hallucinations. If a response does not meet the faithfulness threshold, it should not be presented to the user, thereby reducing the risk of inaccurate or misleading answers.

Additionally, addressing the privacy concerns in LLM inference is paramount. Moving to a locally hosted setup would eliminate the privacy risks associated with the current inference, as highlighted in this thesis. Moreover, a local setup would enable experimentation with more powerful LLMs that were not tested in this work due to computational and privacy limitations.

Although explainability mechanisms were considered during this thesis, they were not implemented. Future work should focus on integrating such mechanisms to enhance user trust in the system. For instance, the extracted context used to generate an answer could be displayed alongside the answer itself.

BIBLIOGRAPHY

- [1] J. M. Lourenço. *The NOVAthesis L^AT_EX Template User's Manual*. NOVA University Lisbon. 2021. URL: <https://github.com/joaomlourenco/novathesis/raw/main/template.pdf> (cit. on p. i).
- [2] G. Caldarini, S. Jaf, and K. McGarry. "A Literature Survey of Recent Advances in Chatbots". In: *Information* 13 (2022), p. 41. DOI: [10.3390/info13010041](https://doi.org/10.3390/info13010041) (cit. on pp. 1, 6).
- [3] M. Milne-Ives et al. "The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review". In: *JMIR Publications* 22.10 (2020), e20346. DOI: [10.2196/20346](https://doi.org/10.2196/20346). URL: <https://www.jmir.org/2020/10/e20346> (cit. on p. 1).
- [4] M. Tepe and E. Emekli. "Decoding medical jargon: The use of AI language models (ChatGPT-4, BARD, microsoft copilot) in radiology reports". In: *Patient Education and Counseling* 126 (2024), p. 108307. ISSN: 0738-3991. DOI: <https://doi.org/10.1016/j.pec.2024.108307>. URL: <https://www.sciencedirect.com/science/article/pii/S0738399124001745> (cit. on p. 1).
- [5] *Value for Health Colab*. URL: <https://vohcolab.org/about-us/> (visited on 2024-02-05) (cit. on p. 2).
- [6] E. Adamopoulou and L. Moussiades. "An Overview of Chatbot Technology". In: *Artificial Intelligence Applications and Innovations*. Ed. by I. Maglogiannis, L. Iliadis, and E. Pimenidis. Cham: Springer International Publishing, 2020, pp. 373–383. ISBN: 978-3-030-49186-4 (cit. on p. 2).
- [7] N. Chandolika et al. "Agriculture Assistant Chatbot Using Artificial Neural Network". In: *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)*. 2022, pp. 1–5. DOI: [10.1109/ICACTA54488.2022.9753433](https://doi.org/10.1109/ICACTA54488.2022.9753433) (cit. on p. 2).
- [8] C. Balona. "ActuaryGPT: Applications of large language models to insurance and actuarial work". In: *Available at SSRN 4543652* (2023) (cit. on p. 2).

-
- [9] Z. Guan et al. "Cohortgpt: An enhanced gpt for participant recruitment in clinical study". In: *arXiv preprint arXiv:2307.11346* (2023) (cit. on p. 2).
- [10] Z. Liu et al. "Pharmacygpt: The ai pharmacist". In: *arXiv preprint arXiv:2307.10432* (2023) (cit. on p. 2).
- [11] S. Wu et al. "BloombergGPT: A Large Language Model for Finance". In: (2023). arXiv: [arXiv:2303.17564 \[cs.CL\]](https://arxiv.org/abs/2303.17564) (cit. on p. 2).
- [12] K. Pandya and M. Holia. *Automating Customer Service using LangChain: Building custom open-source GPT Chatbot for organizations*. 2023. arXiv: [2310.05421 \[cs.CL\]](https://arxiv.org/abs/2310.05421) (cit. on p. 2).
- [13] R. Yang et al. "Large language models in health care: Development, applications, and challenges". In: *Health Care Science* 2.4 (2023), pp. 255–263 (cit. on p. 2).
- [14] P. Lee, S. Bubeck, and J. Petro. "Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine". In: *New England Journal of Medicine* 388.13 (2023), pp. 1233–1239 (cit. on p. 2).
- [15] J. Rodrigues et al. "Advancing Neural Encoding of Portuguese with Transformer Albertina PT-". In: (2023). arXiv: [arXiv:2305.06721 \[cs.CL\]](https://arxiv.org/abs/2305.06721) (cit. on p. 2, 16, 26, 27).
- [16] C. Liu et al. "What is the meaning of health literacy? A systematic review and qualitative synthesis". In: *Fam Med Community Health* 8.2 (2020), e000351. DOI: [10.1136/fmch-2020-000351](https://doi.org/10.1136/fmch-2020-000351) (cit. on p. 2).
- [17] J. Weizenbaum. "ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine". In: *Communications of ACM* 9.1 (1966) (cit. on p. 6).
- [18] R. Yan, Y. Song, and H. Wu. "Learning to Respond with Deep Neural Networks for Retrieval-Based Human-Computer Conversation System". In: *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '16. Pisa, Italy: Association for Computing Machinery, 2016, pp. 55–64. ISBN: 9781450340694. DOI: [10.1145/2911451.2911542](https://doi.org/10.1145/2911451.2911542). URL: <https://doi.org/10.1145/2911451.2911542> (cit. on p. 6).
- [19] I. Sutskever, O. Vinyals, and Q. V. Le. "Sequence to sequence learning with neural networks". In: *Advances in neural information processing systems* 27 (2014) (cit. on p. 6).
- [20] O. Vinyals and Q. Le. "A neural conversational model". In: *arXiv preprint arXiv:1506.05869* (2015) (cit. on p. 6).
- [21] A. Vaswani et al. "Attention is all you need". In: *Advances in neural information processing systems* 30 (2017) (cit. on p. 7).
- [22] P. J. Liu et al. "Generating wikipedia by summarizing long sequences". In: *arXiv preprint arXiv:1801.10198* (2018) (cit. on p. 7).

BIBLIOGRAPHY

- [23] J. Devlin et al. “Bert: Pre-training of deep bidirectional transformers for language understanding”. In: *arXiv preprint arXiv:1810.04805* (2018) (cit. on p. 7).
- [24] GPT2. URL: <https://openai.com/research/gpt-2-1-5b-release> (cit. on p. 7).
- [25] W. X. Zhao et al. “A survey of large language models”. In: *arXiv preprint arXiv:2303.18223* (2023) (cit. on pp. 7, 8, 11, 16).
- [26] Y. Sun et al. “Exploring effective factors for improving visual in-context learning”. In: *arXiv preprint arXiv:2304.04748* (2023) (cit. on p. 7).
- [27] T. Brown et al. “Language models are few-shot learners”. In: *Advances in neural information processing systems* 33 (2020), pp. 1877–1901 (cit. on p. 7).
- [28] J. Achiam et al. “Gpt-4 technical report”. In: *arXiv preprint arXiv:2303.08774* (2023) (cit. on p. 7).
- [29] G. Team et al. *Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context*. 2024. arXiv: 2403.05530 [cs.CL]. URL: <https://arxiv.org/abs/2403.05530> (cit. on p. 7).
- [30] A. Chowdhery et al. “Palm: Scaling language modeling with pathways”. In: *Journal of Machine Learning Research* 24.240 (2023), pp. 1–113 (cit. on p. 7).
- [31] ChatGPT. URL: <https://openai.com/chatgpt> (visited on 2024-02-05) (cit. on p. 7).
- [32] H. Touvron et al. “Llama: Open and efficient foundation language models”. In: *arXiv preprint arXiv:2302.13971* (2023) (cit. on p. 7).
- [33] G. Team et al. *Gemma: Open Models Based on Gemini Research and Technology*. 2024. arXiv: 2403.08295 [cs.CL]. URL: <https://arxiv.org/abs/2403.08295> (cit. on p. 7).
- [34] A. Q. Jiang et al. “Mistral 7B”. In: *arXiv preprint arXiv:2310.06825* (2023) (cit. on p. 7).
- [35] D. Hendrycks et al. “Measuring massive multitask language understanding”. In: *arXiv preprint arXiv:2009.03300* (2020) (cit. on p. 7).
- [36] D. Rein et al. “Gpqa: A graduate-level google-proof q&a benchmark”. In: *arXiv preprint arXiv:2311.12022* (2023) (cit. on p. 7).
- [37] M. Chen et al. “Evaluating large language models trained on code”. In: *arXiv preprint arXiv:2107.03374* (2021) (cit. on p. 7).
- [38] K. Cobbe et al. “Training verifiers to solve math word problems, 2021”. In: URL <https://arxiv.org/abs/2110.14168> (2021) (cit. on p. 8).
- [39] D. Hendrycks et al. “Measuring mathematical problem solving with the math dataset”. In: *arXiv preprint arXiv:2103.03874* (2021) (cit. on p. 8).
- [40] F. Shi et al. “Language models are multilingual chain-of-thought reasoners”. In: *arXiv preprint arXiv:2210.03057* (2022) (cit. on p. 8).

-
- [41] J. Maynez et al. “On faithfulness and factuality in abstractive summarization”. In: *arXiv preprint arXiv:2005.00661* (2020) (cit. on p. 8).
- [42] K. Shuster et al. “Retrieval augmentation reduces hallucination in conversation”. In: *arXiv preprint arXiv:2104.07567* (2021) (cit. on p. 8).
- [43] R. Capellini, F. Atienza, and M. Sconfield. “Knowledge Accuracy and Reducing Hallucinations in LLMs via Dynamic Domain Knowledge Injection”. In: (2024) (cit. on p. 8).
- [44] Z. Liu et al. *Radiology-Llama2: Best-in-Class Large Language Model for Radiology*. 2023. arXiv: [2309.06419](https://arxiv.org/abs/2309.06419) [cs.CL] (cit. on p. 8).
- [45] Z. Wang et al. “Empower Large Language Model to Perform Better on Industrial Domain-Specific Question Answering”. In: *arXiv preprint arXiv:2305.11541* (2023) (cit. on pp. 8, 9, 15, 22, 35).
- [46] C. Wu et al. “Visual chatgpt: Talking, drawing and editing with visual foundation models”. In: *arXiv preprint arXiv:2303.04671* (2023) (cit. on p. 9).
- [47] Y. Shen et al. “Hugginggpt: Solving ai tasks with chatgpt and its friends in huggingface”. In: *arXiv preprint arXiv:2303.17580* (2023) (cit. on p. 9).
- [48] Q. Jin et al. *GeneGPT: Augmenting Large Language Models with Domain Tools for Improved Access to Biomedical Information*. 2023. arXiv: [2304.09667](https://arxiv.org/abs/2304.09667) [cs.CL] (cit. on p. 9).
- [49] S. Robertson, H. Zaragoza, et al. “The probabilistic relevance framework: BM25 and beyond”. In: *Foundations and Trends® in Information Retrieval* 3.4 (2009), pp. 333–389 (cit. on p. 9).
- [50] V. Karpukhin et al. “Dense passage retrieval for open-domain question answering”. In: *arXiv preprint arXiv:2004.04906* (2020) (cit. on p. 9).
- [51] Y. Gao et al. “Retrieval-augmented generation for large language models: A survey”. In: *arXiv preprint arXiv:2312.10997* (2023) (cit. on pp. 9–12, 15).
- [52] P. Liu et al. “Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing”. In: *ACM Computing Surveys* 55.9 (2023), pp. 1–35 (cit. on p. 11).
- [53] K. Papineni et al. “Bleu: a method for automatic evaluation of machine translation”. In: *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*. 2002, pp. 311–318 (cit. on p. 12).
- [54] C.-Y. Lin. “Rouge: A package for automatic evaluation of summaries”. In: *Text summarization branches out*. 2004, pp. 74–81 (cit. on p. 13).
- [55] S. Banerjee and A. Lavie. “METEOR: An automatic metric for MT evaluation with improved correlation with human judgments”. In: *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*. 2005, pp. 65–72 (cit. on p. 14).

BIBLIOGRAPHY

- [56] T. Zhang et al. “Bertscore: Evaluating text generation with bert”. In: *arXiv preprint arXiv:1904.09675* (2019) (cit. on p. 15).
- [57] S. Es et al. “Ragas: Automated evaluation of retrieval augmented generation”. In: *arXiv preprint arXiv:2309.15217* (2023) (cit. on pp. 16, 43).
- [58] R. Rodrigues. “Rapport: a fact-based question answering system for portuguese”. PhD thesis. 2017 (cit. on p. 16).
- [59] N. A. M. Mokmin and N. A. Ibrahim. “The evaluation of chatbot as a tool for health literacy education among undergraduate students”. In: *Education and Information Technologies* 26.5 (2021), pp. 6033–6049 (cit. on p. 16).
- [60] İ. Sarbay, G. B. Berikol, İ. U. Özturan, et al. “Performance of emergency triage prediction of an open access natural language processing based chatbot application (ChatGPT): A preliminary, scenario-based cross-sectional study”. In: *Turkish Journal of Emergency Medicine* 23.3 (2023), p. 156 (cit. on p. 16).
- [61] T. Dwyer et al. “Use of an Artificial Intelligence Conversational Agent (Chatbot) for Hip Arthroscopy Patients Following Surgery”. In: *Arthroscopy, Sports Medicine, and Rehabilitation* 5.2 (2023), e495–e505. ISSN: 2666-061X (cit. on p. 16).
- [62] J. Montenegro, C. Costa, and L. Janssen. “Evaluating the use of chatbot during pregnancy: A usability study”. In: *Healthcare Analytics* 2.100072 (2022). ISSN: 2772-4425 (cit. on p. 16).
- [63] Y. Bian et al. “Artificial intelligence–assisted system in postoperative follow-up of orthopedic patients: exploratory quantitative and qualitative study”. In: *Journal of medical Internet research* 22.5 (2020), e16896 (cit. on p. 16).
- [64] S. Suharwardy et al. “Feasibility and impact of a mental health chatbot on postpartum mental health: A randomized controlled trial”. In: *AJOG GlobalReports* (2023). DOI: [10.1016/j.xagr.2023.100165](https://doi.org/10.1016/j.xagr.2023.100165) (cit. on p. 16).
- [65] A. Ahmed et al. “A review of mobile chatbot apps for anxiety and depression and their self-care features”. In: *Computer Methods and Programs in Biomedicine Update* 1.100012 (2021). ISSN: 2666-9900 (cit. on p. 16).
- [66] A. A. Birkun and A. Gautam. “Large Language Model (LLM)-Powered Chatbots Fail to Generate Guideline-Consistent Content on Resuscitation and May Provide Potentially Harmful Advice”. In: *Prehospital and Disaster Medicine* 38.6 (2023), pp. 757–763 (cit. on p. 16).
- [67] L. Martinengo et al. “Conversational agents in health care: expert interviews to inform the definition, classification, and conceptual framework”. In: *Journal of Medical Internet Research* 25 (2023), e50767 (cit. on pp. 16, 17).
- [68] *EU AI ACT*. URL: <https://artificialintelligenceact.eu/> (cit. on p. 20).
- [69] Z. Li. “The dark side of chatgpt: Legal and ethical challenges from stochastic parrots and hallucination”. In: *arXiv preprint arXiv:2304.14347* (2023) (cit. on p. 20).

- [70] badrinarayan6645541. *Understanding Hit Rate, MRR, and MMR Metrics — analyticsvidhya.com*. <https://www.analyticsvidhya.com/blog/2024/07/hit-rate-mrr-and-mmr-metrics/> (cit. on p. 22).
- [71] J. Fragata. *Vou ser operado ao coração, e agora? - Guia Para uma Rápida Recuperação*. 1st. Editora D'Ideias, 2022. ISBN: 978-989-53457-7-9 (cit. on p. 24).
- [72] *Google Colab Pricing*. URL: <https://colab.research.google.com/signup> (cit. on p. 26).
- [73] *all-MiniLM-L6-v2 description*. URL: <https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2> (cit. on p. 26).
- [74] F. Souza, R. Nogueira, and R. Lotufo. "BERTimbau: pretrained BERT models for Brazilian Portuguese". In: *9th Brazilian Conference on Intelligent Systems, BRACIS, Rio Grande do Sul, Brazil, October 20-23 (to appear)*. 2020 (cit. on p. 26).
- [75] R. Santos et al. *Fostering the Ecosystem of Open Neural Encoders for Portuguese with Albertina PT-* family*. 2024. arXiv: 2403.01897 [cs.CL] (cit. on p. 27).
- [76] *Introducing Meta Llama 3: The most capable openly available LLM to date*. URL: <https://ai.meta.com/blog/meta-llama-3/> (cit. on p. 27).
- [77] R. Lopes, J. Magalhães, and D. Semedo. "GlorIA-A Generative and Open Large Language Model for Portuguese". In: *arXiv preprint arXiv:2402.12969* (2024) (cit. on p. 27).

PROMPTS

Three prompts were created during the system implementation : one written in English, one also written in English but in a dissimilar way and one written in Portuguese. To facilitate its mentioning, prompt codes were created EN, ENv2 and PT. Its content is as follows

- Prompt Code: EN; Content :

"" You are an assistant for question-answering tasks. You must respond in European Portuguese PT-PT. You must use only the following pieces of retrieved context to answer the question. If the answer is not in the retrieved context, say that you don't know. You must be direct and keep the answer concise.

Question: query

Context: context

Answer: ""

- Prompt Code: ENv2; Content :

"" You'll receive a context and a question and you must produce answer to the question based solely on the context. The content will be in European Portuguese and you must produce the answers in European Portuguese PT-PT. If the answer is not in the context, say that you can't answer. You must be direct and keep the answer concise.

Question: query

Context: context

Answer: ""

-
- Prompt Code: PT; Content :

"" Tu és um assistente português com tarefas de pergunta e resposta. Irás receber perguntas em PT-PT e deves responder também em PT-PT. Usa o seguinte contexto para responder à pergunta. Se a resposta à pergunta não estiver no contexto, diz que não consegues responder. Sê direto e mantém a resposta concisa.

Pergunta : query

Contexto: context

Resposta: ""

GENERATION QUALITY RESULTS

B.1 Using BerTimbau Base as Embedding Model

B.1.1 Chunk size of 128 tokens

	Prompt EN		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0694	0.0746	0.0883
ROUGE1	0.0672	0.0688	0.0816
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0672	0.0688	0.0816
METEOR	0.228	0.284	0.240
BertScore	0.7401	0.7651	0.7678

Table B.1: Results of 128 token prompt EN

	Prompt ENv2		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0616	0.0860	0.0873
ROUGE1	0.0434	0.1009	0.0821
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0434	0.1009	0.0821
METEOR	0.275	0.178	0.229
BertScore	0.7443	0.7353	0.7592

Table B.2: Results of 128 tokens for prompt Env2

B.1. USING BERTIMBAU BASE AS EMBEDDING MODEL

	Prompt PT		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0615	0.0470	0.0827
ROUGE1	0.0430	0.0460	0.0700
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0430	0.0460	0.0700
METEOR	0.290	0.306	0.250
BertScore	0.7560	0.7522	0.7632

Table B.3: Results of 128 tokens for prompt in Portuguese

B.1.2 Chunk Size of 256 tokens

	Prompt EN		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0672	0.0804	0.0930
ROUGE1	0.0562	0.0747	0.0807
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0562	0.0747	0.0807
METEOR	0.283	0.276	0.234
BertScore	0.7637	0.7795	0.7653

Table B.4: Results of 256 token prompt EN

	Prompt ENv2		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0649	0.0926	0.0917
ROUGE1	0.0568	0.1178	0.0843
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0568	0.1178	0.0843
METEOR	0.3047	0.2425	0.2458
BertScore	0.7738	0.7649	0.7682

Table B.5: Results of 256 tokens for prompt Env2

	Prompt PT		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0564	0.0530	0.0782
ROUGE1	0.0430	0.0460	0.0700
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0430	0.0460	0.0700
METEOR	0.2771	0.3214	0.2797
BertScore	0.7595	0.7707	0.7622

Table B.6: Results of 256 tokens for prompt in Portuguese

B.1.3 Chunk Size of 512 tokens

	Prompt EN		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0666	0.0740	0.0988
ROUGE1	0.0379	0.0575	0.0823
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0379	0.0575	0.0823
METEOR	0.2360	0.2600	0.2389
BertScore	0.7312	0.7471	0.7651

Table B.7: Results of 512 token prompt EN

	Prompt ENv2		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0597	0.0811	0.0888
ROUGE1	0.0418	0.1178	0.0759
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0418	0.1178	0.0759
METEOR	0.2843	0.2134	0.2333
BertScore	0.7407	0.7508	0.7632

Table B.8: Results of 512 tokens for prompt Env2

	Prompt PT		
	Mistral-7B	LLAMA3-8B	Gemma-7B
BLEU	0.0535	0.0588	0.0741
ROUGE1	0.0348	0.0391	0.0677
ROUGE2	0.0	0.0	0.0
ROUGEL	0.0348	0.0391	0.0677
METEOR	0.2791	0.3021	0.2609
BertScore	0.7434	0.7562	0.7615

Table B.9: Results of 512 tokens for prompt in Portuguese

COMPARISON WITH POPULAR SOLUTIONS

In this Appendix, we display the results of the comparison of our solution with popular tools such as ChatGPT and Gemini. Every answer is exhibited as it was extracted.

C.1 Query - Quando tomo a medicação dopaminérgica, os meus sintomas melhoram. Posso fazer a DBS?

C.1.1 ChatGPT

A Estimulação Cerebral Profunda (DBS) pode ser uma opção de tratamento para pessoas com doenças neurológicas, como a Doença de Parkinson e outros distúrbios do movimento. O fato de os seus sintomas melhorarem com medicação dopaminérgica é um sinal positivo, já que esse é um dos critérios que os médicos consideram ao avaliar candidatos para a DBS.

Aqui estão alguns fatores importantes a serem considerados:

1. Resposta à Medicação Se os seus sintomas melhoram significativamente com medicamentos dopaminérgicos, isso geralmente indica que você pode ser um bom candidato à DBS. A DBS é frequentemente indicada para pessoas que ainda respondem bem à medicação, mas que experimentam flutuações motoras, discinesias (movimentos involuntários) ou outros efeitos colaterais que dificultam o controle dos sintomas com medicação.

2. Estágio da Doença A DBS costuma ser considerada para pacientes que estão em estágios intermediários ou avançados da Doença de Parkinson, quando os sintomas motores já não são controlados de forma consistente com a medicação, mas o paciente ainda responde a ela.

3. Outros Sintomas A DBS é mais eficaz para controlar sintomas como tremores, rigidez e bradicinesia (lentidão de movimentos). No entanto, nem sempre ajuda em sintomas não motores (como problemas de fala ou equilíbrio) ou sintomas de demência.

4. Saúde Geral Antes de considerar a DBS, você será submetido a uma avaliação completa para garantir que sua saúde geral permita a cirurgia. Condições como demência ou distúrbios psiquiátricos graves podem contra-indicar a DBS.

A decisão final sobre a DBS envolve um processo rigoroso de avaliação por uma equipe multidisciplinar (neurologistas, neurocirurgiões e outros especialistas), levando em conta seus sintomas, resposta à medicação e saúde geral.

Recomendo discutir com seu neurologista, que pode encaminhá-lo para uma avaliação de DBS, se apropriado.

C.1.2 Gemini

DBS e Medicamentos Dopaminérgicos: Uma Combinação Possível? A pergunta sobre a possibilidade de realizar a Estimulação Cerebral Profunda (DBS) após o uso de medicamentos dopaminérgicos é muito pertinente e merece uma análise cuidadosa.

A resposta curta é: sim, em muitos casos, é possível realizar a DBS mesmo após o uso de medicamentos dopaminérgicos.

No entanto, é fundamental entender alguns pontos importantes:

Melhora dos sintomas: O fato de seus sintomas melhorarem com a medicação dopaminérgica é um indicativo de que o seu cérebro está respondendo ao tratamento. Isso é um bom sinal e pode indicar que a DBS pode ser ainda mais eficaz. Limitações dos medicamentos: Com o tempo, os medicamentos dopaminérgicos podem perder eficácia ou causar efeitos colaterais significativos. A DBS pode ser uma opção para controlar os sintomas que persistem ou pioram. Decisão individualizada: A decisão de realizar a DBS é sempre individualizada e deve ser tomada em conjunto com seu neurologista. Ele irá avaliar a gravidade dos seus sintomas, a resposta aos medicamentos, a sua idade e outras condições de saúde. Como a DBS funciona em conjunto com os medicamentos dopaminérgicos?

Efeito sinérgico: Em muitos casos, a DBS e os medicamentos dopaminérgicos podem ter um efeito sinérgico, ou seja, podem potencializar os efeitos um do outro, levando a uma melhora ainda maior dos sintomas. Redução de dosagem: A DBS pode permitir a redução da dosagem dos medicamentos dopaminérgicos, diminuindo assim os efeitos colaterais. É importante ressaltar que a DBS não é uma cura para a doença de Parkinson, mas sim um tratamento que pode melhorar significativamente a qualidade de vida dos pacientes.

Quais são os benefícios da DBS?

Redução dos tremores: A DBS é particularmente eficaz na redução dos tremores, um dos sintomas mais comuns da doença de Parkinson. Melhora da rigidez: A DBS também pode ajudar a reduzir a rigidez muscular. Facilitação dos movimentos: A DBS pode facilitar os movimentos e melhorar a coordenação. Melhora da qualidade de vida: Ao aliviar os sintomas, a DBS pode melhorar significativamente a qualidade de vida dos pacientes, permitindo que eles realizem atividades do dia a dia com mais facilidade. Em resumo:

Se você está considerando a DBS, é fundamental conversar com seu neurologista para avaliar os benefícios e os riscos do procedimento. Ele poderá te orientar sobre a melhor

opção de tratamento para o seu caso individual.

Gostaria de saber mais sobre algum aspecto específico da DBS ou dos medicamentos dopaminérgicos?

Disclaimer: As informações fornecidas aqui têm caráter informativo e não substituem a consulta médica.

Fontes:

Clínica Cukiert - Neurologia, Neuropediatria e Neurocirurgia: <https://cukiert.com.br/dbs-e-parkinson-saiba-mais/> Observação: Se possível, forneça mais detalhes sobre o seu caso, como o tempo de diagnóstico da doença de Parkinson, os medicamentos que você utiliza e os sintomas que mais te incomodam. Isso me permitirá oferecer informações mais personalizadas.

C.1.3 Our solution

Sim, é possível que você faça a DBS se os seus sintomas melhorarem com a toma de medicação dopaminérgica. A DBS é eficaz para tratar sintomas motores que melhoram com a toma de Levodopa, como tremores, rigidez e bradicinésia.

C.2 Query - O que é a DBS?

C.2.1 ChatGPT

A DBS pode se referir a diferentes coisas dependendo do contexto, mas alguns dos significados mais comuns incluem:

Deep Brain Stimulation (Estimulação Cerebral Profunda): Uma técnica médica usada principalmente no tratamento de distúrbios neurológicos como a doença de Parkinson, distúrbios do movimento e alguns casos de epilepsia e depressão. Nessa técnica, eletrodos são implantados no cérebro, que enviam impulsos elétricos para áreas específicas, ajudando a regular sinais anormais.

DBS Bank: O maior banco do Sudeste Asiático, com sede em Cingapura. O DBS Bank (anteriormente conhecido como Development Bank of Singapore) é uma das principais instituições financeiras da Ásia e oferece serviços de banco comercial, de investimento e outros produtos financeiros.

Dragon Ball Super: Uma série de anime que faz parte da famosa franquia "Dragon Ball", criada por Akira Toriyama. "Dragon Ball Super" segue os eventos da série anterior "Dragon Ball Z" e introduz novas sagas, personagens e transformações.

Esses são alguns dos significados mais frequentes para o termo DBS, então o significado exato vai depender do contexto em que está sendo usado.

C.2.2 Gemini

Estimulação Cerebral Profunda (DBS): Uma Visão Geral DBS, sigla em inglês para Deep Brain Stimulation, é uma técnica neurocirúrgica avançada que consiste na implantação de eletrodos em áreas específicas do cérebro. Esses eletrodos emitem impulsos elétricos de baixa voltagem que ajudam a regular a atividade cerebral anormal, aliviando os sintomas de diversas doenças neurológicas.

Como funciona a DBS?

Implantação dos eletrodos: Através de uma cirurgia, eletrodos finos são implantados em regiões cerebrais específicas, dependendo da condição a ser tratada. Conexão com o gerador de impulsos: Os eletrodos são conectados a um gerador de impulsos, semelhante a um marca-passo cardíaco, que é implantado sob a pele, geralmente na região do peito. Ajustamento dos impulsos: Um neurologista ajusta a frequência e a intensidade dos impulsos elétricos para otimizar o tratamento e minimizar os efeitos colaterais. Para que serve a DBS?

A DBS é utilizada principalmente para tratar os sintomas de doenças neurológicas como:

Doença de Parkinson: Ajuda a controlar tremores, rigidez muscular e lentidão de movimentos. Distonia: Reduz espasmos musculares involuntários. Tremor essencial: Alivia tremores nas mãos, cabeça ou voz. Distúrbios do controle de impulsos: Pode ser benéfica em alguns casos de distúrbios obsessivo-compulsivos e vícios. Doenças de movimento: Pode ser utilizada em outras doenças de movimento, como a coreia de Huntington. Quais os benefícios da DBS?

Melhora significativa dos sintomas: A DBS pode proporcionar uma melhora significativa da qualidade de vida dos pacientes, permitindo maior independência e participação em atividades sociais. Efeito duradouro: Os benefícios da DBS podem persistir por muitos anos. Menor dependência de medicamentos: Em muitos casos, a DBS reduz a necessidade de medicamentos e seus efeitos colaterais. Quais os riscos da DBS?

Como qualquer cirurgia, a DBS envolve riscos, como:

Infecção: Risco de infecção na área da cirurgia. Hemorragia: Possibilidade de sangramento cerebral. Deslocamento dos eletrodos: O eletrodo pode se deslocar da posição original. Efeitos colaterais: Alguns pacientes podem apresentar efeitos colaterais como dificuldade de fala, tontura ou alterações de humor. É importante ressaltar que a DBS não é uma cura para essas doenças, mas sim um tratamento eficaz para aliviar os sintomas e melhorar a qualidade de vida dos pacientes.

Se você tem alguma dúvida sobre a DBS, consulte um neurologista.

Gostaria de saber mais sobre algum aspecto específico da DBS? Abre uma nova janela [my.clevelandclinic.org Deep Brain Stimulation](http://my.clevelandclinic.org/Deep-Brain-Stimulation)

Disclaimer: As informações fornecidas aqui são de caráter geral e não substituem a consulta a um profissional de saúde.

Palavras-chave: DBS, estimulação cerebral profunda, doença de Parkinson, distonia, tremor essencial, neurocirurgia, tratamento neurológico.

C.2.3 Our solution

A DBS (Estimulação Cerebral Profunda) é um procedimento cirúrgico que envolve a colocação de um dispositivo estimulador no cérebro para tratar sintomas da doença de Parkinson e outros distúrbios neurológicos.

C.3 Query - Existe cura para a doença de Parkinson?

C.3.1 ChatGPT

Atualmente, não existe cura para a doença de Parkinson, mas existem tratamentos que podem ajudar a controlar os sintomas e melhorar a qualidade de vida das pessoas que convivem com a doença. O Parkinson é uma doença neurodegenerativa progressiva, caracterizada pela perda de células nervosas produtoras de dopamina no cérebro, o que afeta o controle motor e outras funções.

Os principais tratamentos disponíveis incluem:

Medicação:

Os medicamentos mais comuns visam aumentar ou substituir a dopamina no cérebro, como a levodopa, que é convertida em dopamina após ser absorvida. Outros medicamentos, como agonistas da dopamina, inibidores da COMT e inibidores da MAO-B, ajudam a aumentar a eficácia da dopamina ou retardar sua degradação. Estimulação cerebral profunda (DBS):

É uma intervenção cirúrgica que implanta eletrodos no cérebro para enviar impulsos elétricos a áreas específicas, ajudando a controlar os sintomas motores. Terapias complementares:

Fisioterapia, fonoaudiologia e terapia ocupacional são importantes para ajudar na manutenção das funções motoras, fala e atividades diárias. Exercícios físicos regulares também demonstram benefícios significativos na gestão dos sintomas. Pesquisa em terapias avançadas:

Existem estudos em andamento com células-tronco, terapias gênicas e tratamentos que visam interromper a progressão da doença, mas estes ainda estão em fases experimentais. Embora os tratamentos atuais possam melhorar os sintomas e a qualidade de vida, a doença ainda é progressiva, e o foco principal é gerenciar os sintomas e retardar sua progressão.

C.3.2 Gemini

A Doença de Parkinson: Um Panorama Atual Infelizmente, ainda não existe uma cura definitiva para a doença de Parkinson. Essa condição neurodegenerativa progressiva

afeta milhões de pessoas em todo o mundo e, embora a ciência tenha avançado significativamente na compreensão da doença, a cura continua sendo um objetivo a ser alcançado.

O que se sabe:

Causas: A causa exata do Parkinson ainda não é totalmente compreendida, mas acredita-se que seja uma combinação de fatores genéticos e ambientais. Sintomas: Os sintomas mais comuns incluem tremor, rigidez muscular, lentidão de movimento e problemas de equilíbrio. Tratamentos: Embora não haja cura, os tratamentos disponíveis podem ajudar a controlar os sintomas e melhorar significativamente a qualidade de vida dos pacientes. As opções de tratamento incluem: Medicamentos: A levodopa é o medicamento mais eficaz para o alívio dos sintomas, mas com o tempo pode perder sua eficácia e causar efeitos colaterais. Cirurgia: A estimulação cerebral profunda é uma opção para pacientes com sintomas graves que não respondem bem aos medicamentos. Terapia: Fisioterapia, fonoaudiologia e terapia ocupacional podem ajudar a melhorar a função e a independência. O futuro:

A pesquisa sobre a doença de Parkinson está em constante evolução. Cientistas de todo o mundo estão trabalhando em diversas frentes, como:

Novas terapias: Desenvolvimento de medicamentos mais eficazes e com menos efeitos colaterais. Terapias gênicas: Manipulação genética para corrigir os defeitos que causam a doença. Células-tronco: Utilização de células-tronco para regenerar os neurônios perdidos. É importante ressaltar que:

Cada caso é único: A progressão da doença e a resposta aos tratamentos variam de pessoa para pessoa. O acompanhamento médico é fundamental: O neurologista é o profissional mais indicado para acompanhar o tratamento e ajustar a medicação conforme necessário. A pesquisa é essencial: O apoio à pesquisa é fundamental para encontrar novas formas de tratar e, eventualmente, curar a doença de Parkinson. Para mais informações, você pode consultar:

Pfizer Brasil: <https://www.pfizer.com.br/noticias/ultimas-noticias/doenca-de-parkinson> CUF: <https://www.cuf.pt/saude-a-z/doenca-de-parkinson> Lembre-se: Embora a cura ainda não seja uma realidade, a doença de Parkinson não é uma sentença. Com o tratamento adequado e o apoio de profissionais de saúde e familiares, é possível viver uma vida com qualidade e independência.

Se você tem alguma outra dúvida, não hesite em perguntar.

Disclaimer: As informações fornecidas aqui têm caráter informativo e não substituem a consulta a um médico.

C.3.3 Our solution

Infelizmente, não há uma cura para a doença de Parkinson. A DBS (Estimulação Cerebral Profunda) é um tratamento que pode ajudar a reduzir os sintomas motores da doença, mas não é uma cura.

C.4 Query - Quantas pessoas vivem com a doença de Parkinson em Portugal?

C.4.1 ChatGPT

Estima-se que cerca de 20.000 a 22.000 pessoas vivam com a doença de Parkinson em Portugal. Essa estimativa varia um pouco dependendo da fonte, mas está dentro dessa faixa, representando cerca de 0,2 % da população portuguesa. A prevalência da doença tende a ser mais alta em pessoas com mais de 60 anos, embora possa ocorrer em indivíduos mais jovens.

C.4.2 Gemini

A doença de Parkinson afeta um número significativo de portugueses. Embora o número exato possa variar ligeiramente ao longo do tempo devido a novos diagnósticos e avanços na pesquisa, estima-se que entre 18 a 20 mil pessoas em Portugal vivam com esta doença.

É importante notar que a prevalência da doença de Parkinson aumenta com a idade e que a população portuguesa está envelhecendo. Portanto, espera-se que o número de casos continue a crescer nos próximos anos.

Fatores que influenciam o número de casos:

Envelhecimento da população: A doença de Parkinson é mais comum em pessoas com mais de 60 anos. Diagnósticos mais precoces: Com o avanço da medicina e maior conscientização sobre a doença, mais casos estão sendo diagnosticados em estágios iniciais. Melhorias nos tratamentos: Novos tratamentos e terapias estão permitindo que as pessoas com Parkinson vivam mais tempo e com melhor qualidade de vida. Onde encontrar mais informações:

Sociedade Portuguesa de Neurologia (SPN): A SPN publica regularmente dados e notícias sobre a doença de Parkinson em Portugal. Associações de pacientes: Existem várias associações em Portugal que oferecem apoio e informações a pessoas com Parkinson e seus familiares. Para saber mais sobre a doença de Parkinson, você pode consultar os seguintes recursos:

Doença de Parkinson afeta entre 18 a 20 mil pessoas em Portugal: <https://www.spneurologia.com/noticias/doenca-de-parkinson-afeta-entre-18-a-20-mil-pessoa/51>
Doença de Parkinson | Hospital da Luz: <https://www.hospitaldaluz.pt/pt/dicionario-de-saude/doenca-de-parkinson> Se você tiver mais alguma dúvida sobre a doença de Parkinson, não hesite em perguntar.

Lembre-se que as informações fornecidas aqui têm caráter informativo e não substituem a consulta a um médico.

C.4.3 Our solution

Estima-se que cerca de 18.000 pessoas vivam com a doença de Parkinson em Portugal.

VALIDATION SET DEEP BRAIN STIMULATION AND PARKINSON'S DISEASE

Table I.1: Validation Set for DBS and Parkinson's Disease

Section	Query	Answer
Introdução - A Doença de Parkinson	Existe cura para a doença de Parkinson?	Atualmente, não existe cura para a doença de Parkinson.
Capítulo 2 – A DBS	O que é a DBS?	A Estimulação Cerebral Profunda ou Deep Brain Stimulation (DBS) é um procedimento cirúrgico que envolve a implantação de um neuroestimulador e elétrodos que enviam impulsos elétricos para alvos específicos no cérebro.
Capítulo 2 – A DBS	Para que serve a DBS?	A DBS serve como tratamento para diversas doenças do movimento, como a doença de Parkinson, tremor essencial e distonia, bem como para certas condições neuropsiquiátricas, como transtorno obsessivo-compulsivo, síndrome de Tourette, epilepsia, entre outras. No caso da doença de Parkinson, a DBS é usada para regular a atividade neuronal, reduzir os sintomas, melhorar a função motora e a qualidade de vida dos doentes.
Introdução - A Doença de Parkinson	Qual é a principal causa de incapacidade no mundo?	As doenças neurodegenerativas são a principal causa de incapacidade no mundo.

Continued on next page

Table I.1 – continued from previous page

Section	Query	Answer
Introdução - A Doença de Parkinson	Quão comum é a doença de Parkinson?	A Doença de Parkinson é a segunda doença neurodegenerativa mais comum em todo o mundo, a seguir à doença de Alzheimer, afetando aproximadamente 1% da população com mais de 60 anos. Em Portugal, estima-se que cerca de 18.000 pessoas vivam com esta condição. De 1990 a 2015, o número de pessoas com a doença de Parkinson aumentou mais do que o dobro de 2,6 milhões para 6,3 milhões, e até 2040, este número duplicará novamente para pelo menos 12,9 milhões.
Introdução - A Doença de Parkinson	Quantas pessoas vivem com a doença de Parkinson em Portugal?	Em Portugal, estima-se que cerca de 18.000 pessoas vivam com a doença de Parkinson
Capítulo 1 – Sintomas da Doença de Parkinson	Que sintomas apresenta a doença de Parkinson?	A doença de Parkinson apresenta sintomas motores e não motores. Os sintomas motores incluem acinesia, bradicinesia, tremor de repouso, rigidez e instabilidade postural. Estes podem resultar em dificuldades em iniciar movimentos, lentidão progressiva, movimento rítmico geralmente nas mãos, resistência à mobilização passiva e aumento do risco de quedas. Além disso, podem ocorrer alterações na marcha, comprometimento da caligrafia, da força de preensão e da fala. Os sintomas não motores são variados e incluem alterações cognitivas, distúrbios do sono, sintomas autonómicos e dor. As alterações cognitivas podem variar desde défices ligeiros de memória até demência grave. Distúrbios do sono, como a insónia e a síndrome das pernas inquietas, são comuns. Outros sintomas não motores podem incluir ansiedade e pânico, depressão, sudorese, perturbação de controlo dos impulsos, comportamentos obsessivos e impulsivos, alucinações e ilusões. Numa fase mais avançada da doença, pode surgir demência, apatia e fadiga.
Continued on next page		

Table I.1 – continued from previous page

Section	Query	Answer
Capítulo 1 – Sintomas da Doença de Parkinson	Quais são os sintomas motores da doença de Parkinson?	Os sintomas motores da doença de Parkinson incluem acinesia, bradicinesia, tremor de repouso, rigidez e instabilidade postural. A acinesia e bradicinesia resultam em dificuldades em iniciar movimentos e lentidão progressiva, afetando atividades diárias como abrir e fechar as mãos, levantar-se de uma cadeira e iniciar a marcha. O tremor de repouso é um movimento rítmico que geralmente começa numa mão e pode afetar ambos os lados do corpo com a progressão da doença. A rigidez muscular provoca resistência à mobilização passiva, afetando a postura e a marcha. A instabilidade postural aumenta o risco de quedas. Além destes, os pacientes com doença de Parkinson podem apresentar sintomas motores adicionais como alterações na marcha, comprometimento da caligrafia, da força de preensão e da fala.
Capítulo 1 – Sintomas da Doença de Parkinson	Quais são os sintomas não-motores da doença de Parkinson?	Os sintomas não motores da doença de Parkinson são variados e incluem alterações cognitivas, distúrbios do sono, sintomas autonómicos e dor. As alterações cognitivas podem variar desde défices ligeiros de memória até demência grave. Distúrbios do sono, como a insónia e a síndrome das pernas inquietas, são comuns. Outros sintomas não motores podem englobar diminuição ou perda do olfato, obstipação, depressão, disfunção sexual, sonhos vivos, seborreia, ansiedade e pânico, sudorese, perturbação de controlo dos impulsos, comportamentos obsessivos e impulsivos, alucinações e ilusões. Numa fase mais avançada da doença, surgem tipicamente a demência, apatia e fadiga.
Continued on next page		

Table I.1 – continued from previous page

Section	Query	Answer
Capítulo 2 – A DBS	Devo fazer a DBS?	A resposta para esta pergunta depende de uma avaliação médica minuciosa e é influenciada por diversos factores tais como o estágio da doença, a resposta aos tratamentos convencionais, a presença de outros problemas de saúde, entre outros. A DBS pode ser uma opção eficaz para pessoas cujos sintomas da doença de Parkinson estão presentes há mais de 5 anos e que beneficiam da Levodopa, cuja doença já não está a ser controlada eficazmente com a medicação e/ou apresentam deterioração cognitiva significativa ou patologia psiquiátrica grave.
Capítulo 2 – A DBS	A DBS é uma cirurgia de baixo risco?	Apesar de, geralmente, a DBS ser considerada de baixo risco, qualquer tipo de cirurgia tem o risco de complicações. A própria estimulação cerebral pode também causar efeitos secundários.
Capítulo 2 – A DBS	Quando tomo a medicação dopaminérgica, os meus sintomas melhoram. Posso fazer a DBS?	Sim, se os seus sintomas melhoram com a toma de medicação dopaminérgica, a DBS pode ser uma boa opção para si.
Capítulo 2 – A DBS	Tenho Parkinson há 3 anos. Posso fazer a DBS?	Não, a DBS não é recomendada para pacientes cujos sintomas surgiram há menos de 5 anos.
Capítulo 2 – A DBS	Tenho um histórico de depressão e ansiedade. Devo fazer a DBS?	A DBS pode piorar a memória ou as alterações de humor em pessoas que já tenham estes problemas, como depressão ou ansiedade. Portanto, talvez não seja recomendado para pessoas com um histórico destes problemas.
Continued on next page		

Table I.1 – continued from previous page

Section	Query	Answer
Capítulo 2 – A DBS	Que sintomas é que a DBS ajuda?	<p>A DBS ajuda a aliviar sintomas motores, tais como:</p> <p>Discinesia: Movimentos involuntários não controlados. Rigidez. Tremor. Problemas de marcha que melhoram com a toma de medicação, como bloqueios da marcha (“freezing of gait”). Flutuações motoras: Momentos “On” quando os sintomas estão controlados, alternados com momentos “Off”, quando não estão. Distonia: Contrações involuntárias dos músculos que forçam certas partes do corpo a executar movimentos repetitivos de torção ou a permanecerem em posições dolorosas (como os dedos dos pés curvarem para baixo ou os pés inclinarem para dentro).</p>
Capítulo 3 – Pré-cirurgia, durante e pós-cirurgia	Vou fazer a DBS. Qual é o processo que antecede a cirurgia?	<p>Antes da cirurgia DBS, será acompanhado por uma equipa multidisciplinar de especialistas. Estes irão ajudá-lo com vários aspetos: medicação em curso e seus efeitos secundários, exames ao cérebro como a ressonância magnética, avaliação dos seus sintomas motores e não motores, e avaliação de outras condições médicas ou potenciais problemas para a cirurgia. A equipa discutirá todas estas informações relevantes, sintomas que a DBS possa ajudar, ou piorar, e potenciais riscos. Com base nesta análise, poderão sugerir mais testes ou melhor observação de determinados sintomas, como mudanças cognitivas ou de humor. Se tudo estiver de acordo, recomendam a DBS e começam a planeá-la. No planeamento, serão discutidas e decididas questões como: que dispositivo DBS e bateria usar, onde colocar os elétrodos, em que parte do corpo se deve colocar a bateria, como será o procedimento, que outras condições médicas necessitarão de atenção especial antes e durante a cirurgia e que medicação se deve ajustar para a cirurgia.</p>

Continued on next page

Table I.1 – continued from previous page

Section	Query	Answer
<p>Capítulo 3 – Pré-cirurgia, durante e pós-cirurgia</p>	<p>Como é a cirurgia da DBS?</p>	<p>A cirurgia de DBS envolve dois procedimentos principais: a colocação dos eléttodos no cérebro e a colocação da bateria/estimulador elétrico (neuroestimulador) no peito ou no abdómen. Na colocação dos eléttodos, estes serão inseridos em áreas do cérebro que controlam o movimento. Na manhã da cirurgia, os médicos pedirão para não tomar a medicação de Parkinson para o tremor de modo a avaliar os sintomas durante o procedimento e colocar os eléttodos na zona certa do cérebro. Durante esta parte, geralmente, o paciente estará acordado para que os sintomas possam ser melhor monitorizados. Apesar da cirurgia demorar algumas horas, o paciente não sentirá qualquer dor. Pode haver algum desconforto, dado que a cabeça estará numa posição fixa durante algum tempo e estará mais rívido uma vez que estará sem o efeito da medicação. No segundo procedimento, após a colocação dos eléttodos, será colocado o neuroestimulador no peito ou no abdómen. Este será conectado aos eléttodos por um fio que será colocado debaixo da pele. Nesta parte do procedimento, o paciente não estará acordado. Após a cirurgia, o paciente ficará internado entre três a dez dias, dependendo do hospital.</p>
<p>Continued on next page</p>		

Table I.1 – continued from previous page

Section	Query	Answer
Capítulo 3 – Pré-cirurgia, durante e pós-cirurgia	Fiz a DBS, como é a recuperação e o processo após a cirurgia?	Após a cirurgia DBS, haverá um período de recuperação que incluirá ajustes nos estímulos elétricos, acompanhamento médico e reabilitação física. Algumas semanas após a cirurgia, terá uma consulta com o médico especialista na doença de Parkinson e outras doenças do movimento. Nesta consulta, o médico ligará o sistema de estimulação cerebral (o dispositivo DBS) e programará especificamente para os seus sintomas. Isto será feito testando várias combinações de estimulação elétrica e escolhendo a que melhor se adequa a si, com base na sua reação e na resposta dos seus sintomas. A sua medicação será ajustada, geralmente com redução gradual de alguma medicação. Também aprenderá como funciona o seu dispositivo, o que lhe permitirá ligar ou desligar o dispositivo e ajustar as suas especificações. A partir desse momento, continuará a ter consultas regulares com o seu médico para monitorização cuidadosa e, se necessário, para fazer ajustes nas especificações do dispositivo. Este processo pode demorar alguns meses. Assim que estiver recuperado, poderá retomar quase todas as suas atividades regulares. No entanto, deve sempre consultar o seu médico antes de recomeçar alguma atividade ou antes de qualquer procedimento médico, incluindo procedimentos dentários
Capítulo 3 – Pré-cirurgia, durante e pós-cirurgia	Fiz a DBS há 1 ano e já me encontro recuperado. Posso retomar as minhas atividades de rotina?	Sim, após a recuperação da cirurgia DBS, pode retomar quase todas as suas atividades regulares. Contudo, é importante que consulte sempre o seu médico antes de recomeçar alguma atividade ou antes de algum procedimento médico, incluindo procedimentos dentários.
Continued on next page		

Table I.1 – continued from previous page

Section	Query	Answer
<p>Capítulo 4 – O papel dos cuidadores na doença de Parkinson</p>	<p>Sou cuidador de um doente de Parkinson. Devo estar atento a quê?</p>	<p>Como cuidador de um doente de Parkinson, deve estar atento a várias coisas. Destacam-se as responsabilidades que geralmente incluem: o controlo da medicação, o fornecimento de apoio emocional, a monitorização dos sintomas, a assistência nas atividades diárias e o acompanhamento nas consultas médicas. Também é crucial que esteja totalmente informado sobre a doença e a sua possível evolução. À medida que a doença progride, aumenta a necessidade de ajuda e monitorização da medicação. Deverá estar emocionalmente disponível para lidar com situações mais frustrantes e encará-las com motivação e apoio. Deverá também estar em contacto com o médico responsável para alertar o doente para os maus hábitos ou más práticas que esteja a desenvolver, como a não toma correta de medicação, falta de exercício físico, e alertar o médico sobre qualquer sintoma motor ou não motor que possa estar a progredir.</p>
<p>Continued on next page</p>		

Table I.1 – continued from previous page

Section	Query	Answer
<p>Capítulo 4 – O papel dos cuidadores na doença de Parkinson</p>	<p>Sou cuidador de um doente de Parkinson que fez a DBS. Devo estar atento a quê?</p>	<p>Como cuidador de um doente de Parkinson que fez a cirurgia DBS, deve estar atento a várias responsabilidades e dinâmicas. Isto inclui o controlo da medicação, o fornecimento de apoio emocional, a monitorização dos sintomas do doente, assistência nas atividades diárias e acompanhar o doente às consultas médicas. No apoio a pacientes que fizeram a cirurgia DBS, o cuidador tem um papel fundamental, e o acompanhamento próximo é crucial. Deve estar totalmente informado sobre a doença e a sua possível evolução e estar preparado para lidar de forma emocionalmente disponível com situações frustrantes para ultrapassar as limitações e baixa autoestima do doente. Em relação ao comportamento do doente, deve estar atento aos maus hábitos a que este possa estar a incorrer, como a não toma correta de medicação, falta de exercício físico, alertando o médico sobre qualquer sintoma motor ou não motor que possa estar a progredir.</p>
<p>Continued on next page</p>		

Table I.1 – continued from previous page

Section	Query	Answer
<p>Capítulo 4 – O papel dos cuidadores na doença de Parkinson</p>	<p>Quais são os meus deveres enquanto cuidador de um doente de Parkinson?</p>	<p>Enquanto cuidador de um doente de Parkinson, os seus deveres podem variar, mas geralmente incluem:</p> <p>Controlo da medicação: É importante monitorizar e assegurar que o doente está a tomar a medicação corretamente.</p> <p>Apoio emocional: A doença de Parkinson pode ser difícil não só fisicamente, mas também emocionalmente. É crucial oferecer apoio emocional ao doente.</p> <p>Monitorização dos sintomas: Deverá estar atento à progressão dos sintomas da doença e qualquer mudança relevante.</p> <p>Assistência nas atividades diárias: À medida que a doença progride, pode ser necessário mais ajuda nas atividades diárias.</p> <p>Acompanhamento nas consultas médicas: Acompanhar o doente às consultas permite que esteja a par da evolução da doença e das instruções médicas.</p> <p>Educação sobre a doença: É crucial que esteja totalmente informado sobre a doença de Parkinson e a sua possível evolução, para melhor ajudar o doente.</p> <p>Comunicação com o médico responsável: Deve alertar o doente para maus hábitos que possa estar a desenvolver e comunicar ao médico qualquer sintoma que esteja a progredir.</p> <p>Além destes, é importante cuidar do seu próprio bem-estar físico e psicológico, participando em grupos e sessões de apoio para cuidadores. As associações de doentes, como a APDPk ou a CNS - Campus Neurológico, disponibilizam recursos para ajudar nesta tarefa.</p>
Continued on next page		

Table I.1 – continued from previous page

Section	Query	Answer
<p>Capítulo 4 – O papel dos cuidadores na doença de Parkinson</p>	<p>Um membro da minha família foi diagnosticado com Parkinson, o que posso fazer para aprender mais sobre a doença e ser um cuidador bem informado?</p>	<p>Existem diversas formas de se informar sobre a doença de Parkinson e aprender a ser um cuidador mais bem preparado. Primeiramente, é crucial que se informe sobre a doença e a sua possível evolução. As responsabilidades do cuidador podem variar, mas geralmente incluem o controlo da medicação, apoio emocional, monitorização dos sintomas, assistência nas atividades diárias, e acompanhamento nas consultas médicas.</p> <p>Existem associações de doentes que têm um papel fundamental na oferta de informação e apoio. A nível nacional, a APDPk e a Young Parkies já desenvolveram encontros, manuais e sessões para ajudar os cuidadores a abordarem este desafio. Estas sessões poderão fornecer-lhe conhecimentos práticos sobre a gestão da doença de Parkinson e estratégias para lidar com os desafios diários. Outros centros, como o CNS - Campus Neurológico, também promovem formações para cuidadores.</p> <p>Recomenda-se também que esteja atento ao seu próprio bem-estar físico e psicológico, para poder estar na melhor condição possível para cuidar do seu familiar. Existem numerosos recursos online que oferecem informações, conselhos e ferramentas para acompanhá-lo neste processo. Sites como o da APDPk (www.parkinson.pt) fornecem uma ampla gama de materiais educativos e informativos sobre a doença de Parkinson e o papel do cuidador.</p>

VALIDATION SET CARDIOTHORACIC SURGERY

Table II.1: Validation Set for Cardiothoracic Surgery

Section	Query	Answer
1.1 As queixas cardíacas	O que é a dor anginosa?	A dor anginosa é um sintoma relacionado com a doença coronária que se manifesta tipicamente durante esforço físico, como metros percorridos, degraus subidos, esforço sexual, entre outros. Normalmente, esta dor cessa com o fim imediato do esforço ou através da administração de medicação sublingual, como a nitroglicerina.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
1.2. O coração é uma bomba	Como funciona o coração?	<p>O coração funciona como uma bomba, impulsionando sangue para todo o corpo, para satisfazer as necessidades de oxigenação que variam com a atividade física e as emoções. Tem o ritmo normal de batimentos entre 60 a 100 vezes por minuto, bombeando cerca de cinco litros por minuto.</p> <p>Possui quatro cavidades - duas aurículas que recebem o sangue, e dois ventrículos que o impulsionam: à direita para os pulmões e à esquerda para o corpo todo, através da artéria aorta. A estrutura do coração é constituída principalmente por músculo, o miocárdio, revestido exteriormente por uma membrana, o pericárdio.</p> <p>No interior, existem as válvulas cardíacas que funcionam como "portas" que abrem num só sentido. As mais comumente atingidas por doenças são a válvula mitral (entre a aurícula esquerda e o ventrículo esquerdo) e a válvula aórtica (entre o ventrículo esquerdo e a aorta).</p> <p>O coração é irrigado pela artéria coronária esquerda e pela artéria coronária direita, que levam sangue com oxigénio para alimentar o músculo cardíaco.</p> <p>O ritmo cardíaco é comandado por um pequeno pacemaker natural, o nódulo auricular, que dispara estímulos elétricos que passam pelas aurículas, atingem o nódulo ventricular e chegam aos ventrículos, fazendo-os contrair de forma regular e sincronizada. Aurículas e ventrículos batem de forma ritmada e em sincronia eléctrica e mecânica.</p>
		Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
1.3.1 Doenças das coronárias	O que é a aterosclerose?	A aterosclerose é um processo pelo qual as artérias coronárias sofrem deposição de placas de gordura que depois se calcificam. Esta condição torna o interior das artérias irregular e provoca apertos localizados que dificultam a circulação do sangue e prejudicam a irrigação do músculo cardíaco. A aterosclerose é mais frequente em pessoas idosas, como parte da aterosclerose generalizada, mas também pode ocorrer em pessoas mais jovens, dependendo da presença de factores de risco como história familiar, hipertensão arterial, aumento de gorduras no sangue, tabagismo, diabetes, entre outros.
1.3.1 Doenças das coronárias	Existem tratamentos para a doença coronária além da cirurgia?	Sim, existem tratamentos para a doença coronária além da cirurgia. Um dos métodos é a medicação, que é eficaz nos casos mais simples e também usada como complemento a outros tratamentos. Outro método é a dilatação das coronárias por cateterismo com aplicação de stents, que é feita a partir de uma artéria periférica e envolve o uso de um balão que é insuflado no local do aperto e deixa um molde (stent) para manter o vaso aberto. Este tratamento é eficaz em alguns casos e tem a vantagem de não ser necessária uma operação cirúrgica de "peito aberto". No entanto, em alguns casos mais graves ou de maior risco, pode ser necessário recorrer à cirurgia coronária de "bypass", particularmente em doentes com doença mais extensa, que envolvem um maior número de artérias, ou em locais onde a dilatação seria perigosa.
1.3.1 Doenças das coronárias	Em que consiste a cirurgia coronária?	A cirurgia coronária consiste em colocar "pontes" ou bypass entre a aorta e as artérias coronárias, permitindo que o sangue irrigue as áreas em risco.
1.3.1 Doenças das coronárias	Quanto tempo demora a cirurgia coronária?	A cirurgia coronária tipicamente demora cerca de 3 horas.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
1.3.1 Doenças das coronárias	Em quanto tempo consigo retomar a minha atividade normal, após a cirurgia coronária?	Após uma cirurgia coronária, o doente geralmente consegue retomar a sua vida normal em pouco menos de um mês.
1.3.2 Doenças valvulares	Quais são as principais queixas de doentes com doenças valvulares?	As principais queixas de doentes com doenças valvulares incluem usualmente cansaço, falta de ar, palpitações. Por vezes, também podem ocorrer inchaço nas pernas (edemas), dor no peito e mesmo desmaios.
1.3.2 Doenças valvulares	Como é que as doenças valvulares são detetadas?	As doenças valvulares são quase sempre detetadas através de um ecocardiograma, que é um exame que se baseia no uso de ecos - ultrassons. Em alguns pacientes, também pode ser necessário realizar um cateterismo para medir as pressões no interior das cavidades e vasos do coração.
1.3.2 Doenças valvulares	Qual é a relação entre a idade do paciente com a existência de doenças valvulares?	A idade do paciente pode estar relacionada com a existência de doenças valvulares. Por exemplo, verifica-se que em pessoas com mais de 75 anos pode surgir uma calcificação da válvula aórtica, que causa aperto valvular conhecido como estenose. Também, em qualquer idade, podem surgir alterações degenerativas da válvula mitral, mas em pessoas mais velhas, essas alterações podem ser mais comuns.
1.3.2 Doenças valvulares	Como é que as doenças valvulares podem ser tratadas?	Na maioria dos casos, as formas moderadas e graves de doenças valvulares são tratadas cirurgicamente. A cirurgia idealmente consiste na reparação da válvula, mas se isso não for possível, a válvula é substituída. No caso das válvulas mitral e tricúspide, a reparação é possível na grande maioria dos casos (mais de 90%), enquanto na válvula aórtica essa possibilidade é mais rara.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
1.3.2 Doenças valvulares	Que tipos de próteses existem para a substituição das válvulas cardíacas? Quais as vantagens e desvantagens de cada uma?	<p>Existem fundamentalmente dois tipos de próteses para a substituição das válvulas cardíacas:</p> <p>Próteses Mecânicas (em carbono): Estas próteses têm a vantagem de durarem indefinidamente, o que significa que, em princípio, não requerem re-operações no futuro. Por esse motivo, são geralmente a escolha preferida para pessoas com menos de 65 anos de idade. No entanto, a maior desvantagem é que obrigam o paciente a tomar hipocoagulantes (medicamentos que tornam o sangue menos coagulável) e a fazer análises periódicas de controle da sua coagulação pelo resto das suas vidas.</p> <p>Próteses Biológicas: Estas são feitas com tecido porcino, bovino ou mesmo humano. Têm uma duração média de 12 a 15 anos e proporcionam uma melhor qualidade de vida, pois não exigem hipocoagulação, com os seus transtornos e riscos associados. No entanto, a desvantagem é que, por calcificarem com o tempo, geralmente requerem uma nova intervenção mais tarde. São preferíveis para pacientes idosos ou para aqueles com outras doenças associadas que não lhes permitem tomar hipocoagulantes.</p> <p>É ainda de notar que, no caso de próteses biológicas em posição aórtica, é já possível, e mais será ainda no futuro, colocar uma prótese (TAVI) dentro da pré-existente, através de um cateterismo e dispensando, assim, uma nova intervenção cirúrgica.</p> <p>A escolha entre próteses mecânicas e biológicas deve ser feita em conjunto com o cirurgião, tendo em conta a situação clínica específica de cada paciente.</p>

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
1.3.2 Doenças valvulares	Quais os riscos/-taxa de sucesso da cirurgia valvular?	A cirurgia valvular é realizada com riscos que não ultrapassam os 3%. No entanto, os resultados para a qualidade de vida dos doentes são excelentes.
1.3.2 Doenças valvulares	A substituição das válvulas requer (obrigatoriamente) cirurgia convencional?	Não, a substituição das válvulas cardíacas não requer obrigatoriamente uma cirurgia convencional. Nos últimos 10 anos, a substituição das válvulas aórticas tem sido progressivamente realizada através de procedimentos menos invasivos. Estes envolvem a utilização de cateteres introduzidos pela perna ou através de pequenas incisões no tórax, em uma técnica geralmente designada de "TAVI". Embora não seja livre de complicações, esta abordagem menos invasiva está a ser progressivamente alargada, sobretudo para doentes mais idosos ou com um risco elevado associado à cirurgia convencional. A decisão entre a cirurgia convencional e este procedimento deve ser bem individualizada, baseada na avaliação do risco e discutida entre cardiologistas e cirurgiões, naturalmente com a participação do doente.
1.3.3. Doenças do músculo cardíaco	O que significa insuficiência cardíaca?	A insuficiência cardíaca é um termo geral que se refere às situações em que o coração não tem capacidade para efetuar adequadamente as suas funções essenciais.
1.3.3. Doenças do músculo cardíaco	Como é que as doenças do músculo cardíaco podem ser detetadas?	As doenças do músculo cardíaco, também conhecidas como miocardiopatias, podem ser detetadas através de exames como o ecocardiograma e a ressonância magnética cardíaca.
1.3.3. Doenças do músculo cardíaco	Quais as principais causas das doenças do músculo cardíaco?	As doenças do músculo cardíaco, ou miocardiopatias, são normalmente causadas por infecção viral, doença coronária e enfartes, ou por causas tóxicas como o excesso de álcool e medicamentos usados no tratamento do cancro.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
1.3.3. Doenças do músculo cardíaco	Como é que as doenças do músculo cardíaco podem ser tratadas?	As doenças do músculo cardíaco podem ser tratadas através de um vasto portefólio de fármacos. No entanto, em casos mais graves, pode ser necessário recorrer à transplantação cardíaca ou à utilização de sistemas mecânicos de assistência ventricular, como as bombas implantadas no corpo que suportam a atividade do coração em paralelo.
1.3.3. Doenças do músculo cardíaco	Que riscos estão associados à transplantação?	Os riscos associados à transplantação cardíaca incluem a rejeição do órgão transplantado, o risco de infecção devido à queda da imunidade e o aparecimento de tumores malignos devido à imunidade celular deprimida.
1.3.3. Doenças do músculo cardíaco	Que alternativas existem à transplantação?	Uma alternativa à transplantação cardíaca é a utilização de sistemas mecânicos de assistência ventricular, como bombas implantadas no corpo, que apoiam a atividade do coração em paralelo. Estas bombas, embora estejam associadas a complicações como hemorragia e infeções, podem melhorar a qualidade de vida do paciente e podem ser utilizadas como ponte para a realização de um eventual transplante, ou como terapêutica definitiva, se a transplantação não for uma opção viável, ou nunca houver um dador adequado.
1.3.4. Doenças do pericárdio	Qual é a doença mais associada ao pericárdio?	A doença mais associada ao pericárdio é a pericardite, que consiste na inflamação e espessamento do pericárdio, comprometendo o enchimento cardíaco. Por vezes, pode haver acumulação de líquido pericárdico em excesso. As causas são múltiplas, desde a tuberculose, a infeções virais ou a tumores. O diagnóstico é facilmente feito por eco, TAC ou ressonância.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
1.3.4. Doenças do pericárdio	Como é que a pericardite pode ser tratada?	A pericardite pode ser tratada dependendo da sua gravidade e da causa subjacente. Se houver acumulação de líquido pericárdico em excesso que comprometa a função de enchimento cardíaco, deve ser feita uma drenagem. Se o pericárdio for demasiado espesso e impossibilitar o enchimento cardíaco, será feita uma cirurgia com remoção do pericárdio – pericardiectomia.
1.3.5. Infecções no coração	O que são endocardites?	Endocardites são infecções no coração causadas por bactérias que entram na circulação do sangue, tipicamente após tratamentos dentários ou intervenções sobre a pele. As bactérias alojam-se nas válvulas ou em eventuais próteses cardíacas existentes, causando as infecções.
1.3.5. Infecções no coração	Como é que as endocardites são identificadas?	As endocardites são tipicamente identificadas através de histórias arrastadas de febre e mal-estar geral, mas o diagnóstico é confirmado por ecocardiografia (preferencialmente transesofágica) e pelo identificar da bactéria no sangue (hemocultura positiva).
1.3.5. Infecções no coração	Como é que as endocardites são tratadas?	As endocardites são tratadas inicialmente por antibióticos em associação, feitos por períodos de até seis semanas. Nalguns casos, a resposta é boa, podendo mesmo haver cura da infecção, mas noutros, sobretudo com complicações, será necessária a cirurgia para limpeza completa do local e nova substituição valvular.
1.3.5. Infecções no coração	Durante o tratamento de endocardites, quando é que pode haver indicação para cirurgia?	Durante o tratamento de endocardites, pode haver indicação para cirurgia em casos em que a infecção resiste aos antibióticos, ou em que se verifica lesão valvular decorrente com insuficiência cardíaca, ou em que se verificam embolias repetidas de material infectado (vegetações). Em casos de endocardite em prótese implantada, ou para certos agentes bacterianos (estafilococos e fungos) a opção cirúrgica deve ocorrer mais cedo.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
1.3.6. Tumores do coração	O mixoma auricular é um tumor maligno?	Não
1.3.6. Tumores do coração	Como é o mixoma auricular pode ser tratado?	O mixoma auricular pode ser tratado através de cirurgia, que é simples e quase sempre curativa.
1.4. As doenças da aorta	O que é um aneurisma?	Um aneurisma é uma distensão anormal da aorta em que o seu calibre ultrapassa o dobro do calibre normal
1.4. As doenças da aorta	Quais as principais causas de um aneurisma?	As principais causas de um aneurisma incluem a aterosclerose, a hipertensão arterial prolongada e causas genéticas, entre as quais a válvula aórtica bicúspide.
1.4. As doenças da aorta	Quais os riscos associados aos aneurismas?	Os riscos associados aos aneurismas incluem a possibilidade de romperem, o que pode ser imediatamente fatal. Este risco aumenta exponencialmente para diâmetros superiores a 55 mm. A rutura da aorta pode ainda causar uma situação grave e evolutiva em poucos dias, a dissecção da aorta, que ocorre com dor súbita e constitui uma verdadeira emergência.
1.4. As doenças da aorta	Como é que os aneurismas podem ser tratados?	Os aneurismas podem ser tratados através de cirurgia, que consiste em excisar a área dilatada e substituí-la por um tubo de Dacron, que é um tecido sintético. A extensão da substituição pode variar. Em alguns casos mais complexos, frequentemente localizados à aorta torácica descendente, a abordagem pode ser não cirúrgica, através da implantação de tubos por cateter a partir de uma artéria periférica, em procedimentos denominados EVARs. Estes visam restituir e estabilizar o vaso no segmento dilatado ou dissecado.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
1.5. As doenças do ritmo cardíaco	O que é a bradicardia?	A bradicardia é uma condição em que o coração bate mais lentamente, com frequências inferiores a 50 batimentos por minuto. Esta condição pode causar várias queixas, incluindo tonturas ou desmaios, e é frequentemente devida a doenças do tecido de condução do coração.
1.5. As doenças do ritmo cardíaco	Como se trata a bradicardia?	A bradicardia é tratada através da implantação de um pacemaker, uma intervenção realizada sob anestesia local. Durante o procedimento, dois fios, os eléctrodos, são inseridos através de uma veia na base do pescoço até ao coração, sob controlo de radioscopia. Estes eléctrodos detetam a atividade do coração e transmitem impulsos que fazem o coração bater mais rápido sempre que a frequência natural baixa para um valor inseguro. A bateria do pacemaker, com cerca de 4 a 5 cm, fica alojada sob a pele, geralmente sob a clavícula, e precisa ser substituída, em média, a cada 6 anos. O pacemaker garante que os batimentos cardíacos do paciente nunca serão inferiores à frequência programada e os sintomas desaparecem logo após a implantação. No entanto, em casos mais graves, como em casos avançados de insuficiência cardíaca, podem ocorrer arritmias ventriculares graves, com risco de morte súbita. Nestes casos, os cardiologistas podem implantar um desfibrilador, que deteta continuamente o eletrocardiograma do paciente e, em caso de arritmia perigosa, aplica automaticamente um choque elétrico, salvando assim a vida do paciente.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
1.5. As doenças do ritmo cardíaco	O que é a fibrilação auricular?	A fibrilação auricular é um tipo de arritmia que é mais comum em pessoas idosas e em doentes com doenças valvulares. Nesta arritmia, as aurículas, que estão frequentemente dilatadas, batem com frequências muito elevadas (400 a 600 vezes por minuto), e os ventrículos batem descompassadamente, com frequências que podem chegar a 150 e mais batimentos por minuto. Como consequência, perde-se a conjugação mecânica entre as aurículas e os ventrículos, o que facilita em 15 a 30% a função do coração, e geram-se facilmente coágulos dentro das aurículas. Estes coágulos são extremamente perigosos, pois podem desprender-se e causar embolias em qualquer vaso, com grande risco de atingir o cérebro, causando acidentes vasculares, por vezes devastadores.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
1.5. As doenças do ritmo cardíaco	Como é que a fibrilação auricular pode ser tratada?	A fibrilação auricular pode ser tratada de várias formas. Nos casos em que a arritmia não está estabelecida há muito tempo, o tratamento por cateterismo pode ser eficaz. Este envolve a ablação com diversas fontes de energia no interior das cavidades do coração. A cirurgia também pode ser usada para tratar estas arritmias, tipicamente em concomitância com cirurgia valvular ou coronária. A técnica cirúrgica consiste em aplicar diretamente linhas de energia nas aurículas, de forma a isolá-las completamente dos ventrículos, abolindo assim a arritmia em cerca de 75% dos casos. Para os restantes doentes, com fibrilação auricular crónica estabelecida há muito tempo, o tratamento pode envolver o uso de anticoagulantes para reduzir o risco de formação de coágulos e acidentes vasculares. Em casos selecionados, a oclusão do apêndice da aurícula esquerda (local onde se formam habitualmente coágulos) pode ser feita durante uma cirurgia valvular para minimizar o desprendimento de coágulos, ou em outras situações, por cateterismo com a implantação de uma "rolha" que encerra o apêndice.
1.6. As cardiopatias congénitas	Existem cardiopatias que possam estar presentes desde o nascimento?	Sim, existem cardiopatias que podem estar presentes desde o nascimento, essas são denominadas de cardiopatias congénitas.
1.6. As cardiopatias congénitas	Quais são exemplos de cardiopatias congénitas?	Exemplos de cardiopatias congénitas são as comunicações interauriculares (CIAs) e a válvula aórtica bicúspide. Adicionalmente, existe uma condição chamada foramen ovale patente, que é uma pequena comunicação interauricular.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
2.1 Preparando a cirurgia - exames médicos	Que exames médicos podem ser pedidos antes de realizar a operação?	Antes de realizar a operação podem ser pedidos vários exames médicos, tais como Radiografia de Tórax, TAC de tórax, TAC da aorta e das artérias coronárias, Electrocardiograma, Ecocardiograma, Cintigrafia de perfusão, Ressonância magnética nuclear e funcional, Cateterismo cardíaco e outros exames como testes de função respiratória e Eco Doppler das carótidas e da circulação cerebral.
2.1 Preparando a cirurgia - exames médicos	Para que serve uma radiografia ao tórax?	Uma radiografia ao tórax serve para avaliar o tamanho do coração e para estudar os pulmões.
2.1 Preparando a cirurgia - exames médicos	Para que serve um TAC ao tórax?	Um TAC ao tórax serve para esclarecer as dimensões do coração, ver quais as cavidades dilatadas, medir o calibre da aorta, avaliar a espessura do pericárdio, e avaliar os pulmões, excluindo quaisquer doenças associadas.
2.1 Preparando a cirurgia - exames médicos	Para que serve um TAC da aorta e das artérias coronárias?	O TAC da aorta e das artérias coronárias serve para reconstruir espacialmente os grandes vasos intratorácicos, diagnosticar aneurismas e dissecções e avaliar o grau de aperto causado por eventuais lesões de aterosclerose.
2.1 Preparando a cirurgia - exames médicos	Para que serve um Electrocardiograma?	Um electrocardiograma serve para estudar o ritmo do coração e avaliar sinais de falta de irrigação (isquemia) ou mesmo sinais de um enfarte do miocárdio. Em alguns casos, um ECG em esforço, como a Prova de Esforço, pode ser necessário para revelar sinais de má irrigação coronária (isquemia), mas estes testes não são totalmente fiáveis
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
2.1 Preparando a cirurgia - exames médicos	Para que serve um Ecocardiograma?	Um ecocardiograma serve para avaliar com muito detalhe a contração do músculo cardíaco, a estrutura e o funcionamento das válvulas cardíacas, e a presença de líquido no pericárdio (derame pericárdico). É possível estudar fluxos anormais de sangue que passam através das válvulas, quer entre cavidades cardíacas. Em certos casos, pode ser necessário realizar um ecocardiograma transesofágico para precisar melhor algumas estruturas, avaliar a válvula mitral e em casos de infecções, como endocardite, visto que este exame fornece superioridade de imagem e maior fiabilidade diagnóstica.
2.1 Preparando a cirurgia - exames médicos	Para que serve uma Cintigrafia de perfusão?	A Cintigrafia de perfusão serve para distinguir, no coração, as zonas vivas e bem irrigadas das zonas que apresentam falta de irrigação (isquemia) ou já sofreram enfarte, ajudando os especialistas a orientar as áreas a irrigar com bypass ou por cateterismo.
2.1 Preparando a cirurgia - exames médicos	Para que serve uma Ressonância magnética nuclear (RMN) e funcional (RMNf)?	Uma Ressonância Magnética Nuclear (RMN) e funcional (RMNf) serve para estudar a função do miocárdio e para o estudo do coração e vasos intratorácicos, através da sua enorme capacidade de resolução das estruturas cardíacas e vasculares. Além disso, a RMN é utilizada para substituir a cintigrafia de perfusão, adicionando sensibilidade relativamente a esta e tem a vantagem de não sujeitar o doente a radiações.
2.1 Preparando a cirurgia - exames médicos	Para que serve um Cateterismo cardíaco?	Um Cateterismo Cardíaco serve para medir pressões dentro das cavidades e visualizar o trajecto das artérias coronárias (coronariografia), permitindo saber exactamente onde se localizam os entupimentos das coronárias, que é um passo essencial antes de qualquer cirurgia coronária.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
2.2 Preparação para a operação	Tenho diabetes, devo ter alguma preocupação adicional antes da operação?	A sua diabetes deve estar o mais optimizada possível antes da cirurgia cardíaca, com medicação adequada que o seu cardiologista deverá ter prescrito.
2.2 Preparação para a operação	Sou fumador, como devo preparar-me para a operação?	Se é fumador, deverá suspender o tabaco quanto mais cedo melhor, antes da sua operação. Além disso, se tiver mais idade ou maior risco respiratório (bronquíticos e fumadores), deverá fazer exercícios de preparação respiratória nas semanas que antecedem a operação, com vista a acelerar a recuperação dos pulmões após a cirurgia.
2.2 Preparação para a operação	Tenho que fazer um procedimento dentário antes da operação, que cuidados adicionais devo ter ao preparar-me para a cirurgia?	Se tiver que fazer um procedimento dentário antes da operação, deve ter em conta que os tratamentos dentários devem ser realizados sempre sob cobertura antibiótica (profilaxia antibiótica) para prevenir endocardites (infecções dentro do coração). Além disso, deverá aguardar pelo menos uma semana após qualquer extracção dentária antes de ser operado, de modo a ter a gengiva totalmente cicatrizada.
2.2 Preparação para a operação	Como devo gerir a minha medicação atual durante a preparação para a operação?	De um modo geral, toda a sua medicação deverá ser mantida até à véspera da operação. No entanto, na maior parte dos casos e por indicação do seu cirurgião, a aspirina e seus derivados, bem como outros medicamentos antiagregantes das plaquetas, serão suspensos uma semana antes da operação. Os medicamentos anticoagulantes (ditos cumarínicos), como o Varfine® e o Sintrom®, deverão ser suspensos, sempre sob vigilância médica, quatro dias antes da operação. Estas medidas visam reduzir a hemorragia durante e depois da operação e minimizar o consumo de sangue.
2.3 Admissão no hospital	Quando serei admitido no hospital?	Será admitido no hospital na tarde do dia anterior à operação.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
2.3 Admissão no hospital	Que exames devo levar para a minha admissão no hospital?	Deve levar todos os seus exames, especialmente os últimos realizados: radiografia do tórax, electrocardiograma, ecocardiograma e análises, bem como, o relatório do cateterismo, as películas e os CDs de imagens.
2.3 Admissão no hospital	O que devo levar (itens pessoais e afins)?	Deve levar os artigos de uso pessoal de quarto, um bom livro, revistas e a sua música favorita. No entanto, deve evitar levar valores pessoais.
2.3 Admissão no hospital	Como vai funcionar o período de admissão?	Após chegar ao hospital e realizar as formalidades de admissão, será conduzido ao seu quarto. Receberá visitas dos enfermeiros do piso e poderá ter de se deslocar ao Rx, ao Laboratório e ao Ginásio onde o instruirão sobre o uso de dispositivos e a realização de exercícios de respiração. Poderá ter de remover os pêlos em excesso na área cirúrgica, tomará um duche com anticéptico e fará um clister de limpeza. O cirurgião e o anestesista visitá-lo-ão para explicar como tudo se desenrolará e pedir-lhe-ão para assinar uma autorização informada para a sua operação e para a anestesia. Será prescrita uma medicação pré-anestésica e será dada uma boa sedação nocturna. Deve respeitar as instruções de jejum que lhe forem dadas e não beber ou comer durante as 6-8 horas antes da intervenção. Caso esteja internado em quarto particular, os seus familiares poderão ficar consigo.
2.3 Admissão no hospital	Durante a minha admissão no hospital, terei que assinar algum documento?	Sim, durante a sua admissão no hospital, o cirurgião e o anestesista pedir-lhe-ão para assinar uma autorização informada para a sua operação e para a anestesia.
2.3 Admissão no hospital	Os meus familiares podem ficar comigo no hospital?	Sim, caso esteja internado em quarto particular, os seus familiares poderão ficar consigo.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
2.3 Admissão no hospital	Como vai funcionar a noite antes da intervenção?	Na noite antes da intervenção, será prescrita uma medicação pré-anestésica e será dada uma boa sedação nocturna para que passe uma noite bem tranquila. Deve respeitar as instruções de jejum que lhe forem dadas e não beber ou comer durante as 6-8 horas antes da intervenção.
3.1 O dia da operação	Como vai ser o dia da operação?	O dia da operação será iniciado com o acompanhamento do paciente ao Bloco Operatório, cerca de meia hora antes da cirurgia, onde será deitado numa maca. Na sala de operações, verá grandes luzes, muito equipamento e toda a gente vestida de azul ou verde, usando máscara e barrete. Será ligado por fios ao monitor cardíaco, o anestesista canalizará, sob anestesia local e sem que lhe doa, uma artéria e uma veia nos seus pulsos, por onde medirá sua pressão arterial enquanto lhe injecta a anestesia. Adormecerá rapidamente e acordará já na Unidade de Cuidados Intensivos, sem recordar do que se passou no Bloco Operatório. A família poderá visitá-lo cerca de meia hora depois na Unidade de Cuidados Intensivos, onde poderão ver alguns tubos de drenagem e o tubo do ventilador. Não terá dores.
3.1 O dia da operação	Onde pode ficar a minha família/acompanhantes durante a operação?	Os seus acompanhantes deverão aguardar no local que lhes indicarem para que, logo que a operação termine, o cirurgião os possa facilmente contactar e dizer como decorreu a intervenção.
3.1 O dia da operação	Em quanto tempo posso receber visitas?	Cerca de meia hora depois da operação, na Unidade de Cuidados Intensivos, poderá receber a visita de não mais de dois dos seus familiares, mas estará ainda a recuperar da anestesia nessa fase tão precoce.
3.1 O dia da operação	Vou ter dores quando acordar após a cirurgia?	Não deverá ter dores quando acordar após a cirurgia.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
3.2 Na unidade de cuidados intensivos	Como vão ser as primeiras horas na unidade de cuidados intensivos?	Nas primeiras 3 a 6 horas na unidade de cuidados intensivos, o paciente poderá manter-se a dormir devido à anestesia que demora algum tempo a ser eliminada. Não sentirá dores, mas não poderá movimentar-se muito. Poderá mesmo ver e ouvir antes de conseguir mexer-se. Nesta fase, estará ligado a um ventilador que o ajuda a respirar, por isso não deve tentar falar ou lutar com o respirador. Quando estiver completamente acordado, verá pessoas à sua volta, como enfermeiros e o cirurgião residente, que manipularão os tubos que o rodeiam. Poderá ouvir apitos intermitentes e notará uma atitude atenta em todos ao seu redor. De vez em quando, poderá sentir tosse, quando aspiram o tubo dentro da sua traqueia.
3.2 Na unidade de cuidados intensivos	Posso dormir após ser retirado o tubo de ventilação?	Não, após a retirada do tubo de ventilação, não deverá dormir. Deverá procurar manter-se desperto, respirar fundo e tossir de vez em quando.
3.2 Na unidade de cuidados intensivos	Quando posso voltar a beber líquidos?	Poderá voltar a beber líquidos mais tarde, não antes das 24 horas pós-operatórias, e quando os intestinos voltarem a entrar em actividade. Nessa altura, será retirada a sonda através da qual tem sido alimentado.
3.2 Na unidade de cuidados intensivos	Quando serão retirados os drenos torácicos? Como vai funcionar este processo?	Os drenos torácicos serão retirados mais ou menos às 36 horas no pós-operatório. Este processo poderá causar uma sensação algo desagradável, mas não especialmente dolorosa, já que lhe será dado um analgésico forte. Após esta fase, sentirá um grande alívio e respirará ainda melhor.
3.2 Na unidade de cuidados intensivos	Quando irei ser transferido para o quarto?	Será transferido para o seu quarto tipicamente a partir do segundo dia após a operação.
3.2 Na unidade de cuidados intensivos	Quanto tempo ficarei na UCI?	A estadia na Unidade de Cuidados Intensivos é em média inferior a 2 dias. No entanto, pessoas mais idosas ou se surgirem percalços, podem permanecer, por segurança, mais tempo.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
3.3 O regresso ao quarto	Como vai funcionar o regresso ao quarto?	No regresso ao quarto, começa uma importante fase da recuperação. Ainda que possa experimentar algum desânimo, deverá tentar combater essa sensação, levantando-se, andando e cuidando da sua higiene, tendo o cuidado de não molhar os pensos. Nas primeiras vezes que se levantar, é aconselhável pedir ajuda ao pessoal de enfermagem para evitar quedas. Provavelmente será ligado a uma telemetria, um pequeno aparelho que transmite o seu ECG de forma constante, mas que não impedirá a sua deslocação livre. Receberá também um colete que deverá usar, principalmente no primeiro mês, para auxiliar em situações como a tosse, movimento e suporte do esterno. Deverá respirar uma atmosfera húmida e usar meias elásticas. As meias devem ser colocadas de manhã e retiradas ao deitar para evitar inchaço nas pernas e deve movimentar as pernas frequentemente para evitar a formação de coágulos. Os enfermeiros medirão o pulso e a pressão arterial diversas vezes ao dia e serão realizados exercícios respiratórios.
3.3 O regresso ao quarto	Poderei ficar acompanhado no quarto?	Sim, poderá ficar acompanhado no quarto, conforme as regras das instituições e se assim o entender.
3.3 O regresso ao quarto	Quando poderei voltar a andar?	Poderá começar a levantar-se e andar logo que comece a fase da recuperação no quarto. No entanto, nas primeiras vezes, é aconselhável que peça ajuda ao pessoal de enfermagem para evitar desequilibrar-se e cair.
3.3 O regresso ao quarto	Quando poderei voltar a cuidar da minha higiene?	Poderá começar a cuidar da sua higiene assim que inicie a fase de recuperação no quarto. No entanto, deve ter o cuidado de não molhar os pensos.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
3.3 O regresso ao quarto	O que é um dispositivo de telemetria?	Um dispositivo de telemetria é um pequeno aparelho, comparável a um gravador de bolso, que transmite em permanência o seu ECG para uma central onde é constantemente controlado. Este dispositivo é utilizado para garantir maior segurança após uma operação, permitindo monitorizar continuamente o estado do coração do paciente.
3.3 O regresso ao quarto	Estou a recuperar no quarto, para que serve o colete que me entregaram?	O colete que lhe foi entregue tem vários propósitos no processo de recuperação. Este vai ajudar a suportar o osso do peito (o esterno), reduzindo a dor, permitindo mobilizar a expectoração e facilitando a cicatrização do esterno. Além disso, o colete também ajudará ao tossir e ao mover-se. Note que o colete não deve ser utilizado para o apertar, mas usado como suporte nestas situações.
3.3 O regresso ao quarto	Como posso evitar a formação de coágulos?	Para evitar a formação de coágulos, é importante que se movimente com frequência, especialmente as pernas. Além disso, deve usar as meias elásticas que lhe serão fornecidas. Estas meias serão colocadas de manhã, ao levantar e retiradas só ao deitar, para evitar que as pernas fiquem edemaciadas (com inchaço).
3.4 As queixas frequentes	Quais são as queixas mais frequentes após a operação?	As queixas mais frequentes após a operação são a dificuldade de dormir, a "prisão de ventre", dores ósseas e musculares, e uma diminuição temporária do apetite.
3.4 As queixas frequentes	Existe algum motivo relacionado com a operação para ter sintomas de prisão de ventre?	Sim, a operação pode causar "prisão de ventre" devido ao facto de os intestinos poderem levar, por vezes, alguns dias para retomar a atividade normal.
3.4 As queixas frequentes	Quando é que o meu apetite vai voltar?	O seu apetite voltará pouco depois de abandonar o Hospital.
3.5 Chegou o momento da alta	Quando é que vou ter alta?	A alta ocorre, tipicamente, a partir do sexto dia pós-operatório.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
3.5 Chegou o momento da alta	Como vai funcionar a alta?	A alta será dada pelo seu cirurgião, que lhe entregará uma carta para o seu cardiologista, um relatório da sua operação e uma nota de alta. Também lhe dará as instruções necessárias, indicando o plano de consultas e a medicação que ficará a tomar. Deverá ir para casa no carro de um familiar e deve evitar grandes deslocações, nomeadamente de autocarro ou de avião, nas primeiras duas semanas após a operação
3.5 Chegou o momento da alta	Ainda não me sinto confiante para ir para casa, posso confiar na alta?	Sim, pode confiar na alta. Se o seu cirurgião lhe está a dar alta é porque acha mesmo seguro fazê-lo. No entanto, se tiver qualquer dúvida, ou no caso, pouco provável, de não se sentir bem, pode sempre telefonar para o seu cirurgião ou para o hospital para esclarecer as suas dúvidas.
3.5 Chegou o momento da alta	Posso conduzir para casa no dia da alta?	Não, não deve conduzir para casa no dia da alta. Deverá ir para casa no carro de um familiar. Não necessita, obrigatoriamente, de uma ambulância.
4.1 A rotina da vida diária	Como vai funcionar o dia-a-dia em casa?	Ao chegar a casa, deve tentar levar uma vida o mais normal possível. Deve levantar-se à hora habitual, tomar banho e arranjar-se, mas sem molhar as suturas até que os pontos sejam retirados. Deve vestir-se normalmente e não permanecer em pijama durante o dia. Pode sair de casa e andar, sem se cansar, evitando extremos de temperatura e descansando a intervalos regulares. Deve procurar o convívio de amigos e moderar todas as suas atividades. Se se sentir triste ou deprimido, deve lembrar-se que é uma situação passageira e que irá superá-la à medida que for recuperando a sua condição física. Deve tentar readquirir rapidamente o seu amor próprio, arranjando-se bem, saindo de casa e voltando ao seu ciclo social o mais depressa possível. Deve manter o cirurgião que o operou informado sobre qualquer alteração no seu estado de espírito.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
4.1 A rotina da vida diária	Tenho me sentido deprimido após a operação, é uma situação comum com outros pacientes?	Sim, é comum que alguns doentes (não muitos!) possam sentir-se deprimidos nas primeiras semanas após a operação. Este sentimento é normal e ocorre no "desacelerar" da enorme tensão sentida antes da cirurgia, numa fase em que o doente se sente cansado, tem algumas dores e se acha fisicamente um pouco limitado. No entanto, é importante lembrar que se trata de uma situação passageira e sem gravidade, que outros doentes já experimentaram e que vencerá sozinho com a sua força de vontade, mas também à medida que for recuperando a sua condição física.
4.1 A rotina da vida diária	Como posso evitar a depressão após a cirurgia?	Para evitar a depressão após a cirurgia, deve tentar levar uma vida o mais normal possível. Levante-se à hora habitual, tome o seu banho e arranje-se, vista-se normalmente e não permaneça em pijama durante o dia. Pode sair de casa e andar, sem se cansar, evitando extremos de temperatura e descansando a intervalos regulares. Procure o convívio de amigos e use de moderação em todas as suas actividades. Se se sentir triste e deprimido, lembre-se que é uma situação passageira e sem gravidade, que outros doentes já experimentaram e que vencerá sozinho com a sua força de vontade, mas também à medida que for recuperando a sua condição física. Deve tentar readquirir rapidamente o seu amor próprio, arranjando-se bem, saindo de casa e voltando ao seu ciclo social o mais depressa possível. Não hesite em manter o cirurgião que o operou ao corrente sobre as alterações que porventura note no seu estado de espírito.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
4.1 A rotina da vida diária	A quem devo comunicar alterações que sinto na minha personalidade após a cirurgia?	Deve comunicar qualquer alteração que sinta na sua personalidade após a cirurgia ao cirurgião que o operou. Ele estará sempre ao seu dispor para falar consigo e para o ajudar, por isso não hesite em mantê-lo ao corrente sobre as alterações que porventura note no seu estado de espírito.
4.1 A rotina da vida diária	Quanto tempo levarei até estar totalmente recuperado?	Tipicamente, levará cerca de um a dois meses a sentir-se totalmente recuperado e a experimentar os verdadeiros benefícios da sua cirurgia.
4.2 O suporte familiar e dos amigos	Durante a recuperação em casa, poderei realizar as tarefas diárias ou poderei precisar de ajuda?	Durante a recuperação em casa, poderá precisar de ajuda para realizar as tarefas diárias. Isso inclui preparar as suas refeições, cuidar das crianças, fazer a sua higiene pessoal, ir às compras, arrumar a casa, tratar do jardim, levar o cão à rua, entre outros. Não poderá levantar pesos de nenhum tipo antes que o esterno (osso do peito) esteja firmemente fixado, o que pode demorar até 4 semanas. Não deverá conduzir durante o primeiro mês após a intervenção e para esforços mais intensos, por exemplo cavar ou jogar golfe, não antes de 2 meses. Portanto, é recomendável que faça planos com a família e os amigos para ter ajuda em casa após a sua cirurgia.
4.2 O suporte familiar e dos amigos	Quando poderei voltar a conduzir?	Não deverá conduzir durante o primeiro mês após a intervenção.
4.2 O suporte familiar e dos amigos	Quando poderei fazer tarefas mais intensas?	Para tarefas mais intensas, deverá esperar pelo menos 2 meses após a intervenção.
4.3 A Dieta	Vou ter algum tipo de restrição alimentar enquanto estiver a recuperar em casa?	Não, não é obrigado a seguir qualquer dieta específica por ter sido operado ao coração. No entanto, deve evitar excessos alimentares em geral e deve consumir com moderação o sal, gorduras, açúcar, café e álcool. Pode consumir normalmente dois cafés por dia e um copo de vinho a cada refeição. É importante que mantenha o seu peso, evitando engordar.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
4.3 A Dieta	Sou fumador, quando poderei voltar a fumar?	Não deverá voltar a fumar em qualquer caso. Se é doente coronário, é absolutamente proibido fumar.
4.4 O banho	Enquanto estiver a recuperar em casa, deverei ter cuidados adicionais com o banho?	Sim, enquanto estiver a recuperar em casa, deverá ter cuidados adicionais com o banho. Poderá tomar banhos, preferencialmente de chuveiro, e se se sentir mais seguro, poderá sentar-se num banco, tendo cuidado com as derrapagens e as quedas. Até remover os agraços metálicos ou pontos na pele, as incisões não devem ser molhadas, devendo ser cobertas com um penso impermeável ou um segmento rectangular de plástico preso superiormente com adesivo.
4.5 Cuidados a ter com as suturas	Que cuidados devo ter com as suturas/-costuras?	As suturas devem ser deixadas ao ar e não requerem cuidados especiais. No entanto, se ainda houver áreas não completamente cicatrizadas, como o local dos drenos, essas áreas podem ser limpas duas vezes por dia com Betadine, por exemplo, e se quiser, cobertas com um penso pequeno. Para não atrasar o processo de cicatrização, só deve molhar as suturas após a remoção dos pontos. Pode tomar banho, mas deve ter cuidado para não humedecer as suturas. Pode expor as suturas ao sol, mas durante o primeiro ano, deve cobri-las com uma camada espessa de creme gordo como o Nivea. Se surgirem sinais inflamatórios (tumefacção, rubor, calor ou dor) ou de supuração (pus), deve notificar imediatamente.
4.6 As meias elásticas	Para que servem as meias elásticas?	As meias elásticas servem para facilitar a circulação e evitar inchaços nas pernas.
4.6 As meias elásticas	Estas meias são recomendáveis para todos os pacientes?	Não, as meias elásticas não são recomendáveis para todos os pacientes. Por exemplo, estas podem ser difíceis de colocar e não favorecer a cicatrização das pernas em alguns casos, como nos diabéticos que tiveram a veia safena removida.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
4.6 As meias elásticas	Sou diabético, devo usar meias elásticas?	Em alguns casos, como nos diabéticos, as meias elásticas podem não favorecer a cicatrização das pernas em caso de remoção da veia safena. Portanto, se for diabético, pode não ser recomendado o uso de meias elásticas.
4.6 As meias elásticas	Quando é que as meias elásticas devem ser usadas?	As meias elásticas devem ser usadas nas primeiras semanas do período inicial da recuperação, particularmente se estiver imóvel a maior parte do tempo.
4.6 As meias elásticas	Como posso prevenir o inchaço das pernas?	Para prevenir o inchaço das pernas, é recomendado o uso de meias elásticas, especialmente nas primeiras semanas de recuperação e se estiver imóvel a maior parte do tempo. Além disso, recomenda-se manter as pernas levantadas quando estiver sentado e evitar ficar em pé por demasiado tempo consecutivo. Contudo, andar precocemente é o melhor remédio para prevenir o inchaço das pernas.
4.7 A medicação	Como vai funcionar a medicação após a cirurgia?	Após a cirurgia, o paciente tomará menos medicamentos para o coração, mas é provável que lhe sejam prescritos medicamentos de vários tipos. Estes podem incluir antiagregantes plaquetários como a Aspirina ou Clopidogrel para manter os "bypass" a funcionar por mais anos, beta-bloqueantes para manter a frequência cardíaca baixa, Amiodarona (Cordarone) um antiarrítmico que tomará por uma ou duas semanas para prever arritmias, diuréticos como a Furosemida (Lasix®) para ajudar a perder líquidos e reduzir o inchaço dos membros inferiores, estatinas para manter o colesterol controlado, hipocoagulantes como o Varfine® ou o Sintrom® (cumarínicos) se tiver próteses mecânicas implantadas ou arritmias permanentes, e medicamentos para controlar a tensão arterial e a diabetes. Alguns destes medicamentos podem ser tomados indefinidamente, enquanto outros serão temporários.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
4.7 A medicação	A medicação prescrita é compatível com a que estava a tomar anteriormente?	É aconselhado que o paciente pergunte ao médico no momento da alta sobre a compatibilidade da sua medicação anterior com a nova.
4.7 A medicação	De quanto em quanto tempo terei consultas de acompanhamento pós cirurgia?	Após a cirurgia, deverá visitar o seu cirurgião a intervalos regulares durante o primeiro ano, especificamente aos 3, 6 e 12 meses
4.8 Ocorrência de dores ou de febre	É normal ter dores no tórax quando faço certos movimentos?	Sim, é normal e referido com frequência por doentes a percepção de dores no tórax com características "reumáticas" e em relação com os movimentos.
4.8 Ocorrência de dores ou de febre	Como posso atenuar as dores no tórax?	As dores no tórax podem ser atenuadas com medicação analgésica do tipo do Bem-U-Ron® ou, se mais severas, do tipo do Nolotil®. No entanto, a longo prazo, é claramente preferível uma mobilização gradual correcta e a reintrodução bem dirigida da actividade física. Se as dores forem particularmente incomodativas ou anormalmente incapacitantes, deverá contactar o médico.
4.8 Ocorrência de dores ou de febre	Estou a recuperar em casa e tive febre, como devo atuar?	Se estiver a recuperar em casa e tiver febre, deve registá-la e avisar o médico. Se, em qualquer fase da convalescença, surgir febre alta ou queixas do foro cardíaco ou outras que ache dever referir, não hesite em contactar o médico.
4.9 Consultas	Quando vão ser retirados os pontos?	Os pontos vão ser retirados cerca de 10 dias após a operação.
4.9 Consultas	Quando vai ser a primeira consulta pós-operatória?	A primeira consulta pós-operatória vai ser tipicamente um mês após a operação.
4.9 Consultas	Quem vai seguir o meu caso após a operação?	Após a consulta cirúrgica do primeiro mês de pós-operatório, será o seu Cardiologista que vai continuar a seguir o seu caso.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
4.10 A telemonitorização (monitorizar a distância)	Para que serve um dispositivo de telemetria?	Um dispositivo de telemetria serve para monitorizar a distância o electrocardiograma do paciente, permitindo detectar precocemente arritmias e tratá-las mesmo antes que o paciente se aperceba delas. Este sistema é usado para a segurança do paciente durante todo o internamento.
4.10 A telemonitorização (monitorizar a distância)	Que dispositivos posso usar para monitorizar a minha recuperação a partir de casa?	Pode usar várias tecnologias para monitorizar a sua recuperação a partir de casa. Existem balanças e outros dispositivos médicos que comunicam com o telemóvel ou têm internet, e enviam as medições para a equipa clínica do hospital. Também pode usar pulseiras ou relógios que têm sensores, que medem a actividade física (número de passos) e o batimento cardíaco, e podem comunicar estas medições para o hospital. Além disso, pode ser-lhe pedido para enviar fotografias da ferida cirúrgica ou responder a perguntas relacionadas com o seu cansaço, dor ou falta de ar, através de um telemóvel.
4.11 O retomar da atividade	Quanto tempo é necessário para recuperar a forma física?	No mínimo, são necessárias quatro semanas para recuperar a forma física anterior.
4.11 O retomar da atividade	Como vai funcionar o regresso à atividade normal?	O regresso à atividade normal deverá ser conduzido pelo próprio indivíduo, sempre de modo gradual e progressivo, especialmente após o primeiro mês de pós-operatório. Deve-se seguir os princípios de não forçar a atividade e fazer o que achar que está ao seu alcance, interrompendo a atividade antes de se cansar. Na maior parte dos casos, pelo segundo mês de pós-operatório, espera-se que tenha sido atingido o grau de recuperação esperada. A atividade profissional pode, em regra, ser recomeçada a partir do segundo mês após a operação.
4.11 O retomar da atividade	Quando posso voltar ao trabalho?	Pode voltar ao trabalho a partir do segundo mês após a operação.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
4.11 O retomar da atividade	Como vai ser, idealmente, o ritmo da minha recuperação?	<p>O ritmo ideal da sua recuperação será:</p> <p>Passada 1 semana: Deve vestir-se todos os dias, comer à mesa, subir/descer alguns degraus de escadas, caminhar diariamente ao ar livre e fazer pequenas tarefas domésticas como pôr a mesa e dobrar roupa. Passadas 2 semanas: Pode começar a cozinhar, aumentar pouco a pouco a intensidade de algumas tarefas domésticas como limpar o pó e lavar a loiça, e aumentar gradualmente a distância que pode percorrer a pé. Passadas 3 semanas: Pode apanhar o autocarro ou o metro, subir ruas levemente inclinadas, ir ao cinema ou a um restaurante e usar a sua bicicleta estática. Passado 1 mês: Pode começar a conduzir o seu carro, ir às compras sem exagerar nos pesos e fazer uma caminhada mais longa. Passadas 6 semanas: Pode passar o aspirador em casa, varrer os espaços interiores e exteriores, passear o seu cão com a trela e fazer uma caminhada de pelo menos 40 minutos diariamente. Passados 3 meses: Pode praticar exercício físico como corrida, ciclismo, ténis, natação, etc., fazer trabalhos domésticos mais exigentes e tratar da sua jardinagem como plantar, cavar e limpar.</p>
		Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
5.1 Hipocoagulação e antiagregação	Fui indicado para hipocoagulação, como devo agir?	<p>Deve seguir todas as indicações médicas rigorosamente. A dose hipocoagulante necessária deve ser ajustada com base na análise do Tempo de Protrombina expressa em INR. Esta análise deve ser realizada regularmente, a frequência a ser definida pelo seu cardiologista (normalmente uma vez por mês e sempre que o INR se altere ou a medicação com Varfine seja alterada). O valor desejável de INR varia para cada paciente. Na maioria dos casos, este valor está entre 2,5 e 3,5. Se o valor de INR for inferior ou superior ao estabelecido, pode haver riscos para a sua saúde, como por exemplo risco de hemorragias. Estes medicamentos devem ser tomados sempre à mesma hora do dia e nunca interrompidos sem supervisão médica. Adicionalmente, deve ter cuidado com as interações entre medicações. Medicamentos como a Aspirina podem alterar a eficácia do Varfine. Portanto, sempre consulte o seu médico antes de começar a tomar qualquer novo medicamento.</p>
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
5.1 Hipocoagulação e antiagregação	Fui indicado para hipocoagulação que cuidados extra devo ter?	Deve assegurar-se que a dose do medicamento hipocoagulante é ajustada através da análise do Tempo de Protrombina expresso em INR. Normalmente este é um acompanhamento que é feito pelo seu Cardiologista que ajustará a medicação e definirá a regularidade das análises, que usualmente são realizadas uma vez por mês e sempre que houver alteração do INR ou da medicação. Deve ter sempre em atenção que INR muito baixos ou muito elevados em relação ao que foi definido para si, representam um risco para a sua saúde. Valores muito baixos comprometem a hipocoagulação desejada e no caso de ter uma prótese mecânica cardíaca podem representar um grave risco para a sua vida, enquanto valores muito elevados aumentam a propensão para hemorragias. Caso observe sangue na urina, na expectoração, pelo nariz ou nódoas negras no corpo, deve notificar prontamente o seu médico. Além disso, é importante que tome a medicação sempre à mesma hora e nunca a interrompa sem indicação e supervisão médica. Complementarmente, deve estar ciente de que medicamentos como a Aspirina e outros anti-inflamatórios podem influenciar a eficácia do Varfine. Portanto, antes de iniciar qualquer novo tipo de medicação, deve sempre consultar o seu médico.
5.1 Hipocoagulação e antiagregação	Fui indicado para hipocoagulação e os valores que medi foram abaixo dos indicados, como devo prosseguir?	Se os valores que medi são inferiores aos que lhe foram indicados, isso pode colocar em risco a eficácia da hipocoagulação e, no caso de ter uma prótese mecânica cardíaca, pode representar um grave risco para a sua saúde e vida. Deverá proceder a uma pronta notificação médica para que o seu médico possa reajustar a sua medicação de acordo com a necessidade.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
5.1 Hipocoagulação e antiagregação	Fui indicado para hipocoagulação e os valores que medi foram acima dos indicados, como devo prosseguir?	Se os valores que mediu são superiores aos que lhe foram indicados, isso pode predispor a hemorragias, o que pode manifestar-se na forma de sangue na urina, na expectoração, pelo nariz, ou sob a forma de derrames cutâneos (nódoas negras) e oculares. Se observar qualquer um destes sintomas, deve proceder a uma pronta notificação médica para que o seu médico possa reajustar a sua medicação de acordo com a necessidade
5.1 Hipocoagulação e antiagregação	Fui indicado para hipocoagulação, devo ter atenção à restante medicação?	Sim, deve ter atenção à restante medicação que toma. Alguns medicamentos, como a Aspirina® e seus derivados, entre outros anti-inflamatórios, podem potenciar ou diminuir a eficácia hipocoagulante do Varfine®. Por isso, sempre que for iniciar um novo medicamento, deve falar com o seu médico para que este avalie possíveis interações medicamentosas. Além disso, os medicamentos que lhe foram receitados para hipocoagulação devem ser tomados sempre à mesma hora (por exemplo, às 19 horas) para não esquecer, e nunca devem ser interrompidos a não ser por indicação e sob controlo médico.
5.1 Hipocoagulação e antiagregação	Para que serve um antiagregante?	Consultar possível gralha no livro
5.2 Profilaxia de endocardite bacteriana	Fiz a cirurgia há 3 meses e vou extrair um dente. Devo ter algum cuidado adicional?	Sim, deve ter um cuidado adicional. Por 6 meses, o paciente está indicado para fazer profilaxia da infecção no coração (endocardite) mediante procedimentos dentários que ocorram após a operação cardíaca. Assim, antes da extração do dente, deverá ser administrado um tratamento com antibióticos, conforme indicado pelo seu médico. Este tratamento tem como objetivo prevenir a entrada de bactérias em circulação.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
5.3 Antibioticos a usar	Que antibioticos devo usar em caso de necessidade?	A resposta depende do tipo de intervenção que precisa ser realizada. Se for uma intervenção na boca, vias respiratórias ou dentes, deve usar a Amoxicilina, ou Eritromicina se for alérgico à Penicilina. Se for uma intervenção nas vias urinárias ou no tubo digestivo, a opção seria a Ampicilina e a Gentamicina ou, em caso de alergia à Penicilina, Vancomicina e Gentamicina. Em todos os casos, as dosagens e tempos de administração variam.
5.4 Sinais de alarme - o que fazer?	Que queixas ou sintomas devo ter uma atenção reforçada durante a recuperação?	Durante a convalescença, deve prestar especial atenção aos seguintes sinais de alarme: Febre alta Sensação de falta de ar, em esforço ou quando se deita; Batimentos cardíacos acelerados e/ou persistentemente irregulares; Edemas (inchaços) nos membros inferiores, que persiste e não desaparece durante a noite, enquanto está deitado; Qualquer dor súbita no tórax, particularmente intensa, mantida e que não desapareça com a medicação prescrita; Desmaios ou perdas súbitas de consciência; Aparecimento de sangue, serosidade ou pus em qualquer das suturas, em particular na do peito; Nódoas negras injustificadas – sangramento excessivos (urina, nasal, gengivas). É importante consultar o cirurgião ou o cardiologista caso sinta algum desses sintomas.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
5.4 Sinais de alarme - o que fazer?	Que sintomas são considerados anormais?	Sintomas como febre alta, falta de ar, batimentos cardíacos acelerados e/ou persistentemente irregulares, inchaços nos membros inferiores que persistem e não desaparecem durante a noite, dor súbita no tórax particularmente intensa, mantida e que não desapareça com a medicação prescrita, desmaios ou perdas súbitas de consciência, aparecimento de sangue, serosidade ou pus em qualquer das suturas, em particular na do peito e nódos negros injustificadas e sangramentos excessivos na urina, nariz ou gengivas são considerados sintomas anormais durante a convalescença e devem ser imediatamente reportados ao cirurgião ou cardiologista.
5.4 Sinais de alarme - o que fazer?	Tive febre alta, devo preocupar-me ou é algo normal durante a recuperação?	Se tiver febre alta durante a convalescença, deverá consultar o seu cirurgião ou cardiologista o mais rápido possível, pois este sintoma é considerado como um dos sinais de alarme e poderá significar que algo está a correr fora da rotina e carece de investigação ou intervenção pronta.
6.1 Deixe de fumar	Como posso deixar de fumar?	Para deixar de fumar, pode recorrer a medicamentos específicos para esse fim. Além disso, procurar ajuda para aprender a modificar os seus hábitos de fumar pode ser uma boa opção.
6.2 Controle o colesterol	Como posso controlar/baixar o colesterol?	Para controlar ou baixar o colesterol, deve seguir uma dieta pobre em gorduras. Além disso, pode recorrer a medicamentos do tipo das Estatinas, que o seu cardiologista provavelmente lhe prescreverá.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
6.3 Tenha uma alimentação saudável	Vou ter alguma restrição alimentar durante a minha recuperação?	Durante a sua recuperação, não terá necessariamente uma dieta restritiva. No entanto, deve evitar excessos alimentares em geral e consumir com moderação o sal, gorduras, açúcar, café e álcool. Deve manter uma alimentação variada, sendo que legumes e peixe parecem ser mais saudáveis que a carne e as gorduras. Deve também evitar o excesso de sal e açúcares, mesmo que não tenha tensão alta ou diabetes. É importante manter o peso e evitar engordar.
6.4 Pratique atividade física	Existe alguma atividade física aconselhável durante a minha recuperação?"	Sim, durante a sua recuperação é aconselhável praticar exercício físico regular. Uma atividade particularmente saudável é andar a pé em marcha acelerada, cerca de 5km/h, por períodos entre os 30 a 60 minutos, pelo menos 5 vezes por semana. Este exercício deve ser feito de forma gradual e descontraída, sem esforçar os seus limites.
6.5 Aprenda a gerir o stress	Como posso controlar o stress durante a minha recuperação?	Para controlar o stress durante a sua recuperação, é aconselhável aprender a relaxar mental e fisicamente. Um dos melhores exercícios de relaxamento que pode fazer é controlar a sua respiração. Ao concentrar-se na sua respiração, deixará de pensar em qualquer outra coisa. Este exercício deve ser praticado antes e depois da sua cirurgia.
6.6 Tenha qualidade de sono	Não consigo dormir mais que 4 horas por noite, o que devo fazer?	Se está a ter problemas em adormecer ou dormir pelo menos 6 a 8 horas durante a noite, deve contactar o seu médico. O sono é essencial para o bem-estar físico, psicológico e emocional, por isso é importante garantir que está a descansar o suficiente.
6.7 Exercite a memória	Existem algumas técnicas recomendadas para exercitar a memória durante a minha recuperação?	Sim, leitura, fazer palavras cruzadas, jogar xadrez, costura, tricot, entre outros.

Continued on next page

Table II.1 – continued from previous page

Section	Query	Answer
6.8 Não caia em depressão	Recomendações para evitar sintomas de depressão?	Para evitar sintomas de depressão, é importante não deixar de conversar com o seu médico e com os demais profissionais da equipa sobre o que sente. Eles poderão dar-lhe todo o apoio necessário. É normal sentir insegurança, medo e preocupação com o futuro no período pós-operatório, mas é importante partilhar esses sentimentos com os profissionais de saúde para que possam ajudá-lo da melhor forma.
6.9.1 Tomar banho	Quando posso tomar banho sem quaisquer restrições?	Pode tomar banho sem quaisquer restrições no dia seguinte à remoção das suturas ou agrafos.
6.9.2 Atividade sexual	Quando posso retomar a minha atividade sexual?	A retoma da atividade sexual após a cirurgia irá variar de doente para doente. No entanto, em média, deverá evitar ter relações sexuais antes que tenham decorrido duas semanas após a operação. Após as quatro semanas, não deverá haver quaisquer limitações sexuais. É importante discutir qualquer dificuldade com o seu cirurgião ou cardiologista.
6.9.3 Atividade profissional	Quando posso retomar a minha atividade profissional?	Poderá retomar a sua atividade profissional quando se sentir bem para tal, tipicamente 2 meses após a cirurgia. No entanto, isto depende da opinião do seu médico cardiologista, do tipo de trabalho que faz e de como se está a recuperar.
6.9.4 Condução automóvel	Quando posso voltar a conduzir?	Deve esperar pelo menos 4 semanas após a sua cirurgia antes de voltar a conduzir. Isto é para evitar solicitar mecanicamente o osso do peito (esterno), cuja cicatrização leva em média 4 semanas.
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
6. Como zelar pela manutenção da sua saúde	Que cuidados devo ter para acelerar a minha recuperação?	<p>Para acelerar a sua recuperação, deve seguir os seguintes cuidados:</p> <p>Deixe de fumar. O tabaco é prejudicial para as suas coronárias e pulmões e pode comprometer o resultado da operação.</p> <p>Controle o seu colesterol. Mantenha o colesterol em valores inferiores a 190 mg/dl através de uma dieta pobre em gorduras e com a ajuda de medicamentos como as Estatinas.</p> <p>Tenha uma alimentação saudável. Evite excessos alimentares e consuma com moderação o sal, gorduras, açúcar, café e álcool. Mantenha uma alimentação variada, com legumes e peixe.</p> <p>Pratique atividade física regular, como andar a pé em marcha acelerada por períodos entre os 30 a 60 minutos, pelo menos 5 vezes por semana.</p> <p>Aprenda a gerir o stress. Pratique exercícios de relaxamento e controle a sua respiração.</p> <p>Tenha qualidade de sono. Se tiver problemas em adormecer ou dormir pelo menos 6 a 8 horas durante a noite, deve contactar o seu médico.</p> <p>Exercite a memória. Leitura, palavras cruzadas, xadrez, costura, tricot, etc., podem ajudar a melhorar a memória e a concentração.</p> <p>Evite cair em depressão. Converse com o seu médico e com os profissionais da equipa sobre o que sente.</p> <p>Retome a sua rotina. Pode tomar banho no dia seguinte à remoção das suturas ou agrafos, retomar a atividade sexual após duas semanas, regressar ao trabalho quando se sentir bem para tal, tipicamente 2 meses após a cirurgia, e não conduzir durante pelo menos 4 semanas após a cirurgia.</p>
Continued on next page		

Table II.1 – continued from previous page

Section	Query	Answer
8. Prevenindo as doenças do coração	Como posso prevenir as doenças do coração?	Para prevenir doenças do coração, pode modificar ou controlar vários factores de risco, tais como deixar de fumar, melhorar os seus hábitos alimentares, controlar a hipertensão arterial, gerir o stress, ansiedade e depressão, controlar a diabetes e combater o sedentarismo. Os antecedentes familiares de doença coronária são um factor que não pode ser modificado, mas estar ciente disso pode ajudar a tomar medidas preventivas.





2024 Developing End-to-End, Deep Learning-Based Chatbots for Healthcare Support in Portuguese: Miguel Santos

