Linguistic-agnostic Intelligent Support Systems for probable diagnosis to aid in Clinical Settings

Gautam Ahuja

The Capstone Project Report is submitted to the
Department of Computer Science to fulfill the requirements for the degree of

Postgraduate Diploma in Advanced Studies and Research (DipASR)

in Computer Science at Ashoka University

Supervisor:
Dr. Rintu Kutum, Faculty Fellow, Department of Computer Science, Ashoka University
Internal Mentor:
Dr. Anurag Agarwal, Dean, Trivedi School of BioSciences, Ashoka University
External Mentor:
Dr. Govind K Makharia, Professor, Department of Gastroenterology and Human Nutrition Adjunct Faculty, AIIMS - New Delhi

The capstone project report was submitted to the Ashoka University on December 08, 2024 and accepted by the Department of Computer Science on December 08, 2024.

Acknowledgment

I want to thank my advisor, Dr. Rintu Kutum. I am grateful for his guidance, support, and encouragement throughout my project. I would also like to thank Dr. Anurag Agarwal and Dr. Govind K Makharia for their valuable suggestions and feedback. I want to acknowledge the help of Sanjana Ahuja, who designed the mobile interface. I would like to thank the Mphasis F1 Foundation for the internship opportunity leading to this project. I would like to thank Aaryan Nagpal and Ragul N for their assistance in the prototyping and development of the Folk Computing application under the mentorship of Prof. Ramesh Jain, Prof. Partha Pratim Das, and Dr. Rintu Kutum. I would like to thank the team from HealthUnity, namely Prof. Jia Li and David Oniani, for the suggestions and feedback on the Folk Computing Project. I would like to thank the members of the Augmented Health Systems group for their constant support, encouragement, and valuable discussion related to the project. Additionally, I would like to thank the Mphasis F1 Foundation, the Department of Computer Science, Ashoka University, and the Koita Centre for Digital Health at Ashoka (KCDH-A) for the computing infrastructure, without which this project work could not be completed. Finally, I would like to thank my external evaluators for evaluation of my capstone project in advance.

Dedicated to my parents and all family members.

Certificate

I, **Gautam Ahuja**, declare that the work entitled **Linguistic-agnostic Intelligent Support Systems for probable diagnosis to aid in Clinical Settings** is an authentic record of my work carried out at the **Department of Computer Science, Ashoka University**, under the guidance of my supervisor, **Dr. Rintu Kutum** for the fulfillment of the requirements of the degree of **Postgraduate Diploma in Advanced Studies and Research (DipASR)** in the **Department of Computer Science** at **Ashoka University**.

I further declare that this thesis has not been submitted, in part or in full, for the award of any other degree, diploma, or certificate at this or any other institution. All sources of information, ideas, or quotations that are not my own have been duly acknowledged in the bibliography.

This declaration is made with the full knowledge and understanding of the university's regulations on academic honesty and integrity.

This thesis has been approved by the thesis advisor and the examination committee, and it fulfills the requirements for the award of the degree.

Date: December 08, 2024

Gautam Ahuja

Gautam Ahuja
ASP' 25, 1020211395
Department of Computer Science
Ashoka University

Dr. Rintu Kutum
Department of Computer Science
Ashoka University

Contents

List of Figures

List of Tables

Acronyms

AI	Artificial Intelligence.
AIIMS	All India Institute of Medical Sciences.
AIM	Artificial Intelligence in Medicine.
CDS	Clinical Decision Support.
CDSS	Clinical Decision Support System.
LLM	Large Language Model.
NLP	Natural Language Processing.
OPD	Outpatient Department.
osLLM	Open Source Large Language Model.

1 Introduction

Computer-aided clinical decision-making and reasoning have long been considered a model of human behavior. For many years, it has influenced and been the subject of Artificial Intelligence study [1]. From the very inception of reasoning foundations in medical diagnosis and Clinical Decision Support nearly 65 years ago—where the reasoning was based on symbolic logic and probability understanding and optimum treatment was calculated via value theory—to current advancements in AI, decision systems are being designed to model human knowledge and augment the work of clinicians [2, 3].

This capstone project, aims to develop an application tailored to address the unique needs in the Outpatient Department (OPD) setting at the Department of Gastroenterology and Human Nutrition, All India Institute of Medical Sciences (AIIMS), New Delhi, specifically for the chief complaint of **abdominal pain**. The application is designed to assist physicians through a linguistic-agnostic conversational agent that engages directly with patients, collecting key responses from a protocol-driven questionnaire. These responses are then processed by a rule-based deterministic system to provide a probable diagnosis and organ of origin.

The remainder of this report is organized as follows:

- **Chapter 2** outlines the background and motivation, along with problem statement and objectives of the project.
- **Chapter 3** provides a detailed literature survey, identifying the state-of-the-art in CDSS and gaps.
- **Chapter 4** explains the methodology, and summarizes the work done.
- **Chapter 5** presents work done, results and discussions.
- **Chapter 6** concludes the report with future work direction.

2 Background, Motivation, Problem Statement, and Objectives

2.1 Background

2.1.1 Diagnosis of Abdominal Pain in Clinical Setting (OPD)

Abdominal pain is one of the most common and diagnostically challenging chief complaints encountered in clinical practice [4]. This serves as a symptom of a wide range of underlying conditions, encompassing both gastrointestinal and non-gastrointestinal issues. This wide range of potential causes increases diagnostic complexity, requiring a systematic and comprehensive evaluation process.

The assessment of abdominal pain typically follows a systematic evaluation strating with a physician-patient interaction. This interaction incorporates multiple clinical dimensions. These dimensions provide clues that aid in narrowing down possible diagnoses. The key aspects considered during an evaluation include:

- **Location of Pain**: The abdomen is anatomically divided into nine regions (as shown in Figure 2.1 (A) below) — epigastric, umbilical, hypogastric, bilateral hypochondriac, bilateral lumbar, and bilateral iliac regions [5]. The specific region where pain is reported often serves as an essential diagnostic clue [4].

- **Presence of Danger Signs**: Symptoms like lightheadedness, altered sensorium, and respiratory distress may signal critical underlying conditions requiring immediate attention.

- **Severity of Pain**: Pain intensity (1-10) is subjectively reported by the patient but often correlates with the urgency of the clinical situation.

- **Onset of Pain**: The onset of pain, whether the pain arise suddenly over minutes-hours (acute) or the pain arose gradually over hours-days (insidious).

- **Character of Pain**: Pain can be classified as burning, stabbing, pin-pricking, constricting, throbbing, dull aching, or non-specific. Each pain type is linked to a distinct set of possible diagnoses.

- **Duration of Pain**: The duration of pain can be classified as either less than 3 months (acute) or more than 3 months (chronic).

- **Radiation of Pain**: The direction of pain radiation (to the back, shoulder, groin/inner thigh, arms, or neck) provides vital diagnostic insight.

- **Aggravating Factors**: Activities, such as eating, bending forward or sideways, passing stool, passing urine, menstruation, deep inspiration, or walking and exercise, can exacerbate abdominal pain.

- **Associated Symptoms**: Symptoms like fever, nausea, vomiting, constipation, diarrhea, jaundice, and changes in bowel habits serve as crucial diagnostic adjuncts. Additionally, systemic symptoms like weight loss, loss of appetite, and signs of anxiety or depression may signal specific gastrointestinal disorders.

- **Comorbidities**: Chronic illnesses such as diabetes, cardiovascular disease, kidney disease, or a history of gallstones can affect abdominal pain etiology.

- **Surgical History**: Prior surgical procedures on the gallbladder, intestines, kidneys, or uterus can influence the current presentation of abdominal pain.

- **Gender-Specific Considerations**: Conditions related to the female reproductive health, such as abnormal vaginal bleeding, menstrual irregularities, and foul discharge, play a role in the evaluation of abdominal pain.

The process of evaluating abdominal pain begins with a physician-patient interaction where history is taken to collect information on the above-mentioned aspects. This is followed by a physical examination, as shown in Figure 2.1 (B) below, where physicians use visual inspection and hands-on techniques such as palpation, percussion, and auscultation. The examination is often performed with the patient in a supine position with bent knees to relax the abdominal muscles and facilitate assessment [5, 6].

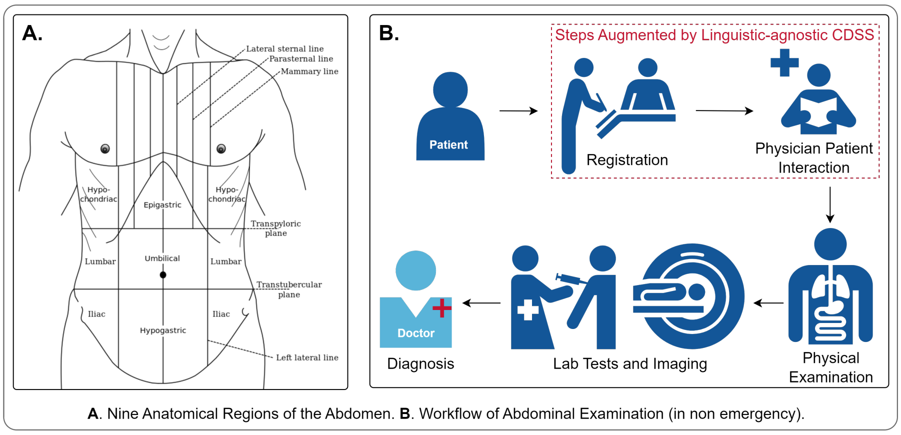

Figure 2.1: A. The nine regions of the abdomen. B. The step-by-step process of a patient undergoing an evaluation of abdominal pain. The dotted red box show where the linguistic-agnostic conversational application will be augmented.

If the initial evaluation does not provide sufficient diagnostic clarity, physicians may order laboratory tests (e.g. blood work, urine analysis, etc.) or imaging studies (e.g. ultrasound, MRI, or CT scan) to gather further evidence. Based on the collective information from history, physical examination, and diagnostic tests, physicians generate a differential diagnosis — a list of possible conditions that could explain the symptoms. This process continues until the final diagnosis is reached [7].

The application developed as part of this capstone project aims to augment the physician-patient interaction phase of abdominal pain evaluation workflow. By integrating a conversational agent, the system collects essential patient information regarding the dimensions mentioned earlier to identify a probable diagnosis and organ of origin using a deterministic approach. The final decision and subsequent physical examination remain under the physician's control.

2.1.2 Brief Introduction to Clinical Decision Support Systems (CDSS)

Clinical Decision Making

Effective clinical decision-making requires:

- Up-to-date Medical Knowledge
- Access to Accurate and Complete Patient Data
- Good Decision-Making Skills

However, several significant changes have emerged as the healthcare landscape has evolved:

- Exponential Growth of Medical Knowledge
- Rapid Accumulation of Patient Data
- Clinical Data Capture and Documentation Burden

Given the changes and the rapid advancements in health, assisting clinicians is becoming more and more important. Thus the role of Clinical Decision Support Systems (CDSS) becomes essential in modern healthcare [8]. With the adoption of Artificial Intelligence (AI) in Medicine, existing physicians can handle **more cases** as systems become **more automated** [9].

Clinical Decision Support (CDS)

Clinical Decision Support (CDS) refers to a broad set of tools, systems, and interventions aimed at enhancing clinical decision-making. By providing clinicians, healthcare workers, and patients with **situation-specific knowledge**, CDS supports a range of critical healthcare activities such as diagnosis, risk assessment, prognosis, and treatment selection [10]. This support can be delivered through various mediums, including reference guidelines, alerts, and evidence-based recommendations.

The goal of CDS is to **bridge the gap between clinical knowledge and patient care**. As medical knowledge expands and healthcare systems become more data-driven, CDS ensures that healthcare providers have access to timely and relevant information. This, in turn, enables more efficient and informed decision-making [10].

Clinical Decision Support System (CDSS)

A Clinical Decision Support System (CDSS) is a specialized type of CDS that employs **software-based tools** to directly aid healthcare providers in making clinical decisions. Unlike general CDS, which may include static guidelines or references, CDSS is dynamic, **patient-centric**, interactive and assist physicians. By integrating patient-specific information with a computerized knowledge base, CDSS generates recommendations, alerts, and assessments that assist in diagnosis, treatment, and care planning [11].

> *"Any software designed to aid in clinical decision-making by matching patient characteristics to a computerized knowledge base to generate tailored patient-specific assessments or recommendations"*
>
> - Hunt et al. [11]

Unlike fully autonomous AI models, CDSS systems are designed to work **in collaboration with healthcare professionals**, offering them interpretive insights and enabling human oversight.

Types of Clinical Decision Support System (CDSS)

As per [8] CDSS can be broadly classified into two categories as shown in Figure 2.2 below:

- **Knowledge-based CDSS**: The key components of a knowledge-based AI-CDS system include a knowledge base such as expert-derived rules and an inference mechanism for clinical application such as chained inference for rules.

- **Data Derived CDSS**: The key components of a data-derived AI-CDS include a model, such as a data-derived neural network, and an inference mechanism for clinical application, such as forward propagation in a neural network model..

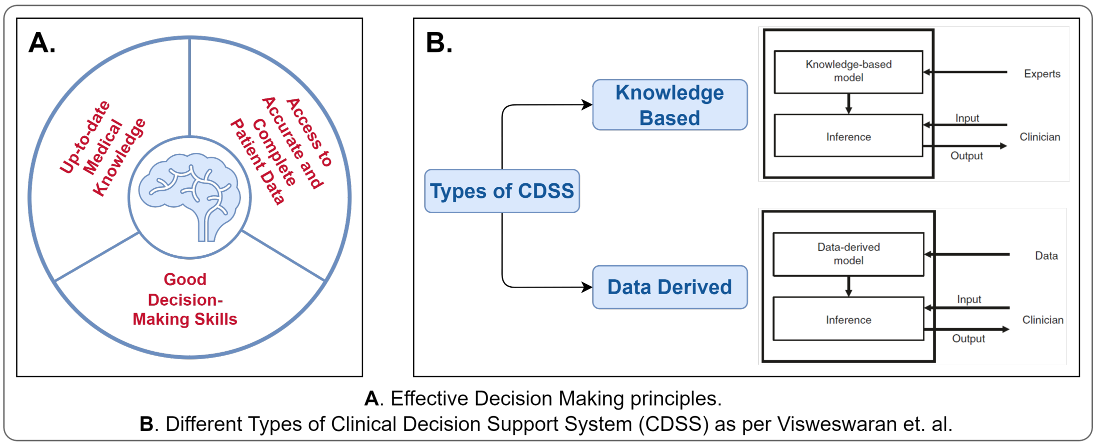

Figure 2.2: A. Illustrates the principles of decision making process followed by a physician. B. Illustrates the two types of CDSS systems, Knowledge-based and Data-derived.

2.1.3 Brief Introduction to Conversational AI in Healthcare

The efforts to use natural language system for problem solving using human input can be traced back to the 1960s [12] and have since evolved significantly. Dialogue systems can be classified into two categories — task oriented chatbots for specific tasks and open-domain conversational AI for general conversation. Today, conversational AI systems, both mobile and web-based, enbale human-computer interaction using natural, human-like dialogue. Unlike chatbots, which rely on scripted responses to user queries, conversational AI utilizes NLP and LLMs to enable more dynamic, context-aware, and adaptive responses [13]. Chatbots follow rigid workflows, while conversational AI systems understand intent, manage multi-turn dialogues, and evolve with user interactions.

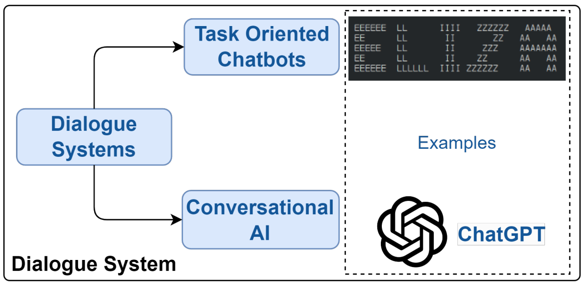

Figure 2.3: Types of Dialogue Systems. Task-oriented chatbots such as ELIZA follow a rigid workflow, while conversational AI such as ChatGPT understand intent, manage multi-turn dialogues, and evolve with user interactions.

The origin of using chatbots in healthcare can be traced back to the development of ELIZA. Joseph Weizenbaum developed one of the first medical chatbots in the 1966 at the Massachusetts Institute of Technology (MIT) [14, 15]. ELIZA employed pattern-matching rules to simulate therapist-like conversations with users. The system was designed to respond to user inputs by reflecting the user's statements back as questions.

Since then, Natural Language Processing (NLP) has progressed to the current state of Large Language Models (LLMs). These are advanced **AI models** specifically trained to process, understand, and generate text. This evolution has facilitated the rise of advanced conversational AI systems that can understand, generate, and respond to human language with the help of conversational agents. These agents can understand the context of a conversation, conduct multi-turn dialogue transactions, and establish a more natural interaction with users [16]. Since the release of ChatGPT, numerous LLMs—both open-source and proprietary—have been developed at unprecedented speed [17]. These models have been applied to various domains, including healthcare.

2.1.4 Brief Introduction to Linguistic-agnostic Systems in Healthcare

Computational linguistics focuses on **modeling human language** using computational techniques, enabling machines to understand and process human language. The process of language understanding can be broken down into multiple hierarchical levels, each representing a critical step in the language comprehension pipeline [18]. These levels, illustrated in Figure 2.4, are as follows:

- **Phonology**: This level focuses on sound patterns in language. One of the most notable linguistic tasks at this level is **Automatic Speech Recognition (ASR)**, also known as speech-to-text, where speech waveforms are converted into textual data. The concept of **linguistic-agnostic** systems relates to the ability of ASR models to understand diverse accents, dialects, and speech variations without being restricted to a particular language or pronunciation. For example, a linguistic-agnostic system can process English spoken by people from different regions (e.g., American, British, and Indian accents) with equal accuracy.

- **Morphology**: Morphology focuses on how words are formed by combining **morphemes**, the smallest units of meaning. For example, in the word "unhappiness," the morphemes are "un-", "happy", and "-ness".

- **Syntax**: Syntax refers to the grammatical structure of language and focuses on how words are arranged in a sentence. It identifies the **relationships between words** and builds a hierarchical representation of the sentence, allowing for proper interpretation
- **Semantics**: Semantics deals with **word meanings** and their relationships within sentences. It enables systems to understand the meaning of words, phrases, and sentences after receiving inputs from the phonology, morphology, and syntax layers
- **Pragmatics**: This layer focuses on how context influences meaning by going beyond the literal meaning of words.
- **Generation**: While the above layers focus on language understanding, this layer focuses on **realistic language generation** given a computational representation

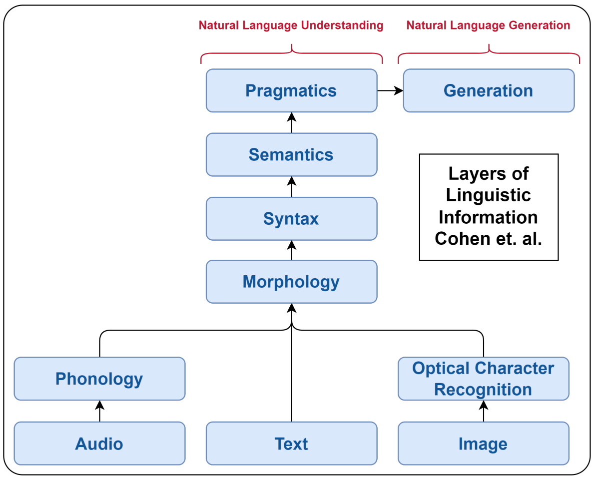

Figure 2.4: Linguistic Levels

Accurate interpretation of human language—especially spoken language—is one of the most critical factors that influence the success of **human-computer interaction**. To achieve a **linguistic-agnostic system**, a large, diverse dataset is required to capture regional, phonetic, and dialectal variations in speech.

India's linguistic diversity presents a unique challenge for developing linguistic-agnostic systems. With more than 1,652 "mother tongues" and 22 scheduled languages, poses significant challenges for healthcare access across the health literacy spectrum [19]. Patients often speak in their native languages and dialects, and may find it challenging to describe symptoms or comprehend medical advice in a language unfamiliar to them. Addressing this challenge requires linguistic-agnostic systems capable of understanding and responding in multiple languages, dialects, and speech patterns.

Efforts to bridge linguistic barriers have seen notable contributions within India. The Bhashini initiative [20], under India's National Language Translation Mission, aims to create linguistic-agnostic systems that allow seamless access to digital services in 22 Indian languages. Using AI-driven voice and text-based translations, Bhashini strives to reduce the language divide across India's vast population. AI4Bharat [21], a research lab at IIT-Madras, advances AI technology for Indian languages, focusing on speech synthesis, automatic speech recognition, and natural language understanding. Their open-source models were trained on diverse Indian languages and dialects data collected from over 400 districts in India with over 15,000 hours of transcribed data, encompassing all 22 scheduled languages of India.

The Folk Computing project, with its visions to allow speech input and output to enhance the accessibility of health related information in multiple languages, is a step towards improving accessibility. It is an Android based application with a chat interface powered by LLMs, enabling users to access the health model in multiple languages. The inital version of the application was developed for Hindi, Tamil, Telugu, and Bengali languages at Ashoka University [22]. A version with Chinese language support was also developed in collaboration with the HealthUnity organization. The demonstration of the application are available on the Folk Computing website here. The complete link to the project is available in Appendix 5.

2.2 Motivation

Assisting Physicians in OPD settings

The **Department of Gastroenterology and Human Nutrition** at the All India Institute of Medical Sciences (AIIMS), New Delhi, was established in 1971 to create qualified gastroenterologists for the country. It is one of the 49 teaching departments and centers at AIIMS, New Delhi.

According to the 67th AIIMS Annual Report (2022-2023) [23], the department managed a total of **1,35,944 outpatient department (OPD) cases**, of which **42,586 were new cases** and **93,358 were follow-up cases**. This is a significant increase from the previous year's figures of 17,790 new cases and 35,622 follow-up cases as shown in Table 2.1 below. In total, the Main Hospital at AIIMS catered to **10,39,523 patients** through its General OPD, Specialty Clinics, and Emergency Department.

Year	New Cases	Follow-up Cases	Total Cases
2020-2021	7,920	11,956	19,876
2021-2022	17,790	35,622	53,412
2022-2023	42,586	93,358	1,35,944

Table 2.1: Number of OPD Cases (new and follow-up) observed at the Department of Gastroenterology and Human Nutrition, AIIMS, New Delhi through the years 2020-2023.

The routine gastroenterology OPD operates from Monday to Friday, 8:30 a.m. to 1:00 p.m. [24]. Given this limited time frame of 270 minutes daily over 5 working days, approximately **8,500 new cases are handled per day**, with multiple physicians addressing various chief complaints. As discussed in Section 2.1.1, one of the most diagnostically challenging chief complaints is **abdominal pain**. A complete protocol of structured questions helps filter the probable diagnosis and the organ of origin. **Automating** this process will significantly assist physicians in managing the high volume of patients the OPD setting.

This challenge forms the motivation for the development of an autonomous system. The application aims to augment the physician-patient interaction phase of abdominal evaluation through a conversational agent and assist physicians by generating probable diagnoses and identifying the organ of origin before physical examination. When deployed, the system can handle multiple patients simultaneously via multiple devices. It can conduct the initial evaluation and generate a probable diagnosis and organ of origin. The output will be printed for the patient, who can then present it to the physician for further evaluation.

Another significance is the drive to automate and digitze the health data for India's linguistically diverse population. AIIMS serves as a **melting pot for linguistically diverse populations** from across India. This is reflected in the geographical distribution of inpatients at AIIMS during the year 2021-2022, as shown in the Figure 2.5 below [23]. This diversity spans multiple languages, dialects, accents, regional pronunciations, and dialectical variations.

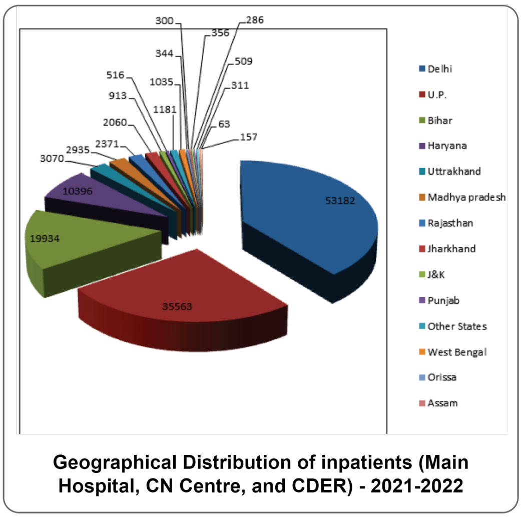

Figure 2.5: A pie chart showing the geographical distribution of inpatients at AIIMS, New Delhi, during the year 2021-2022. The majority of inpatients come from the states of Delhi, Uttar Pradesh, and Bihar.

This diversity in linguistic backgrounds presents a unique opportunity to digitize **voice-based** data related to health from from system-patient interactions. Such digitlization may provide, in the future, valuable insights into the linguistic variations in healthcare settings in India.

2.3 Problem Statement

The volume of patients at OPD settings related to gastrointestinal complaints is large. Given the time constraints (270 minutes) and the large influx of patients (around 8,500 cases per day), automation and digitlization of the initial physician-patient interaction, specifically for evaluating the chief complaint of abdominal pain, will significantly assist physicians in managing the high volume of patients in the OPD setting.

2.4 Objectives

The objectives are:

1. **Development of an application** to assist physicians in the evaluation of abdominal pain in the OPD setting. This application will collect patient responses through a structured, protocol-driven questionnaire and generate a report with a probable diagnosis and organ of origin through a deterministic rule-based system.

2. **Implementing** a conversational agent that interacts with patients to help collect responses for the questionnaire.

3. **Evaluation** and assessment of the impact of the rule-based deterministic system with the conversational agent in the OPD setting for abdominal pain evaluation. The evaluation will be performed by comparing no system (baseline), option and click based system with help of a healthcare staff, and a fully conversational system.

3 Literature Review

3.1 Clinical Decision Support Systems (CDSS)

3.1.1 A Brief History of the Field

The field of diagnostic reasoning in medicine became an early focus of AI, demonstrating that AI methods could approximate human performance in tasks requiring extensive domain knowledge. Early systems were designed to model human cognition explicitly, prioritizing interpretability over mere optimization of accuracy. Such systems were particularly adept at explaining their reasoning, contrasting with modern AI systems optimized solely for prediction accuracy, often at the cost of transparency [1].

One of the landmark systems in this domain was **MYCIN**, introduced in the 1970s [18, 25, 26]. MYCIN was developed to assist physicians in selecting appropriate antimicrobial therapies for severe infections. Its key components included:

- **Consultation Program**: Acquired patient data and provided treatment recommendations.
- **Explanation Program**: Generated English-language explanations, detailing why certain questions were asked and how conclusions were reached.

MYCIN's knowledge of infectious diseases was represented as production rules—conditional statements linking observations to inferred outcomes. It introduced backward chaining, a reasoning strategy that began with a hypothesis and worked backward to validate it using available evidence. This allowed MYCIN to answer "WHY" questions, making its decision-making process interpretable and user-friendly [26, 27].

Before MYCIN, the **Leeds Abdominal Pain System**, developed in the late 1960s, marked earlier attempt at diagnostic reasoning. F. T. de Dombal and his colleagues at the University of Leeds created decision aids for diagnosing abdominal pain based on Bayesian probability theory. This work laid the foundation for probabilistic reasoning in medical decision-making [27].

The initial wave of AI-driven diagnostic systems emphasized the trade-off between accuracy and interpretability, often prioritizing the latter. Over time, the focus is shifting to integrating AI into a larger system involving human decision-makers, aiming to improve the quality, efficiency, and safety of clinical practice.

3.1.2 The Era of Knowledge-Based Systems (KBS)

Following systems like MYCIN and the Leeds Abdominal Pain System, the next phase of AI in healthcare saw the rise of Knowledge-Based Systems (KBS). These systems sought to replicate human reasoning in complex medical scenarios by formalizing knowledge into computational representations. The core ideas of KBS were:

- **Representing knowledge using**:
  - Formal methods: Mathematical frameworks for precise reasoning.
  - Ontological commitments: Hierarchies of concepts organized logically.
  - Fragmentary theories of reasoning: Integrating logic, psychology, biology, statistics, and economics.
- **Ensuring** that representations were computationally efficient and intuitive for human practitioners to understand and modify.

Prominent approaches to knowledge representation included Rules and Patterns—Logical or heuristic rules for decision-making. Probabilistic Models—Methods like Naive Bayes, Bayesian Networks, and Influence Diagrams. Causal Mechanisms—Explaining outcomes through cause-and-effect relationships. Fuzzy Logic—Handling uncertainty and imprecision in medical reasoning.

The success of KBS depended on robust methods for acquiring and organizing medical knowledge. Techniques included:

- **Taxonomic Ontologies**: Organizing concepts into hierarchical structures, specifying their super- and sub-categories.
- **Knowledge Graphs**: Capturing relationships between concepts for intuitive reasoning.
- **Textual Co-occurrence Analysis**: Identifying relationships from sentences, paragraphs, or articles.
- **Unified Medical Language Systems (UMLS)**: Leveraging the Metathesaurus to bridge semantic gaps without constructing exhaustive ontologies.
- **Triple-Based Models**: Representing knowledge as ([concept]-[relation]-[concept]) triples, enabling systems to address "which," "why," and "does" questions [28].

Some other methods included conception and prototypical design of a internet based decision-support server [29] with a option-based input method and a knowledge-based, terminologically driven data processing system for decision support. These methods were observed to have a significant impact on user-friendliness and performance of the system.

3.2 State of the Art

The evolution of AI technologies in healthcare has driven a shift toward more dialogue-based approaches, taking advantage of LLMs to facilitate interactive conversations. The development of such technologies is particularly critical in healthcare, where the delivery of complex information must be accessible. These systems can play a transformative role for individuals with low health literacy or limited familiarity with technology, helping them navigate complex medical advice and treatment options effectively [18].

Recent advancements in conversational LLMs like ChatGPT-4 have demonstrated remarkable capabilities in reasoning, planning, and using contextual information. While GPT-4 is trained on vast amounts of text data, and primarily developed for general domain, it lacks the specialized medical knowledge required for clinical decision-making. Clinical data tends to be proprietary, sensitive, and subject to strict privacy regulations, making it challenging to access and use for training AI models. Because of the interactive nature of the system, the user can request more detail regarding the response by asking follow-up questions or asking for more concise responses. These systems, often, try to provide final responses too early in order to get "to the point" more rapidly, making them not ideal for medical diagnostics purposes [30].

On the other hand, systems like Google's Articulate Medical Intelligence Explorer (AMIE): A research AI system for diagnostic medical reasoning and conversations [31] represent a new generation of AI tailored specifically for medical diagnostics and reasoning. Google's AMIE employs a multi-agent modular architecture to facilitate diagnostic reasoning akin to physician consultations. It simulates primary care physician (PCP) consultations and Objective Structured Clinical Examinations (OSCE), enabling explainable, auditable decision support. Unlike earlier black-box AI systems, AMIE's modular structure promotes system transparency.

In the Indian context, initiatives like Bhashini and AI4Bharat have contributed significantly to the development of linguistic-agnostic models as discussed in section 2.1.4. The initiaves like Folk Computing further provide examples of linguistic-agnostic conversational systems.

A notable emerging trend also being observed where graphs-based models and graph representation learning [32] is being used in medicine to manage and analyze the vast, multimodal datasets generated in healthcare. Graph-based models are particularly well-suited for capturing complex relationships between entities, such as diseases, symptoms, treatments, and patient demographics.

3.3 Gap Analysis

A significant limitation of current approaches is their black-box nature and full autonomy in diagnostic decisions [17]. There is a pressing need for auditable and traceable systems where AI assists rather than replaces clinical decision-making [18]. The ideal solution should provide explainable recommendations while keeping diagnostic control firmly in the hands of healthcare professionals, ensuring accountability, and maintaining the critical role of human expertise in patient care as explained by the three pillars in Figure 3.1 [33]. This is particularly important in specialized medical departments where standardized protocols and human oversight are essential for patient safety [18].

While recent advancements in LLMs have shown impressive capabilities in medical domains, there remains a critical gap in protocol-driven and specialized Clinical Decision Support Systems. Current general-purpose AI systems lack the specialized knowledge and rules required for specific medical departments, potentially leading to misinformation and compromised patient care. Additionally, most existing systems operate as standalone solutions rather than taking a system-level perspective that considers integration with existing clinical workflows and protocols, incorporating both **rule-based** and **data-driven** approaches while also ensuring **explainability** and **transparency** in decision-making.

From the linguistic's perspective, despite advancements, there remain critical gaps. Current models primarily support formal, text-based Hindi but struggle with spoken Hindi, which includes colloquial expressions and regional variations. Pure Hindi, as used in textbooks, often diverges from the Hindi spoken by contemporary populations. Current ASR models fail to generalize across accents and non-standard usage. Voice-based systems are further limited in handling India's vast spectrum of dialects, which vary significantly by region and community.

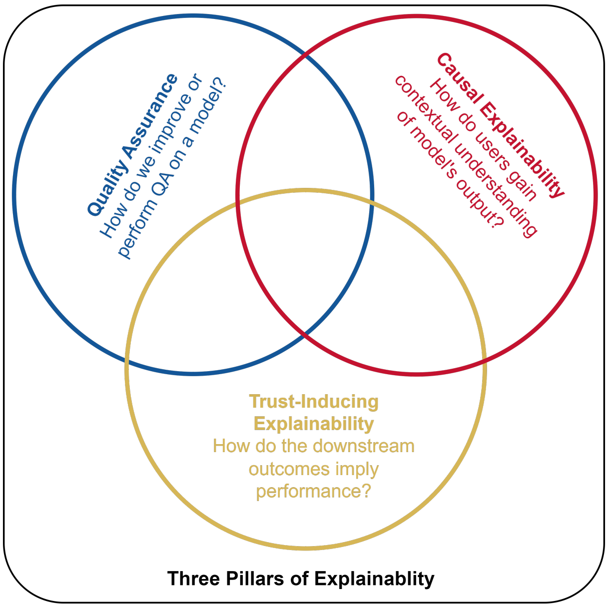

Figure 3.1: Three pillars of explainable AI

The development of this application will address the gap of assisting clinicians in the OPD setting by providing a protocol-driven, rule-based system with conversational capabilities to help the increase patient throughput. The deterministic nature of the system ensures transparency and auditability.

4 Methodology

The implementation of an application was done in various steps and several technologies were used for the same. The core component of the application is the protocol-driven questionnaire along with the data for probable diagnoses and organ of origin. The questionnaire and its associated data of probable diagnoses and organ of origin were acquired from the physicians at AIIMS, New Delhi. The questions was then processed, refined and categorized into four categories to be used in the application. The data associated with the probable diagnoses and organ of origin was also processed and refined into a structured format to be used in the application.

A step-by-step workflow was designed to implement questionnaire into the application which was developed for two platforms — mobile (Android based) and web-based. A UI/UX design for frontend and user interface was developed for the mobile application and the web-based application. The mobile application was developed using Android Studio and the web-based application was developed using open-source technologies like Python and Streamlit. A backend was designed to handle the data and the logic of the application. The backend was designed to be modular and easily inspectable. The backend hosts the processed data for probable diagnoses and organ of origin, and the conversational agent for the questionnaire. The backend was deployed on a local server.

The final workflow of evaluation of abdominal pain in OPD setting augmented by the above application was also designed.

5 Work Done, Results & Discussions

5.1 Protocol

The protocol defines the flow of patient interaction through a set of questions designed to capture essential diagnostic information related to abdominal pain. The initial protocol consisted of a 12-question diagnostic questions that addressed key clinical dimensions of abdominal pain, as described in Section 2.1.1. These questions explored aspects like pain location, intensity, aggravating factors, associated symptoms, etc. However, for a more controlled and streamlined workflow, the questionnaire was later broken down to include 17 refined questions. By refining these questions, a categorization of questions was made possible. For instance, questions related to menstrual cycle health were separated from the general set of symptoms to provide more focused and relevant insights for female patients. The 17 questions were grouped into the following categories:

- **Discriminators**: Critical questions are used to differentiate between emergency cases and regular OPD cases. If the answer to any of these questions indicates a potential emergency, the patient is automatically redirected to the emergency department instead of continuing through the normal OPD process.
- **Demographic**: Questions related to the patient's demographic information, such as age and gender.
- **Gender-Specific**: Questions specific to female health issues — menstrual health — that play a significant role in the diagnostic process for women.
- **General**: These are common questions applicable to all patients, regardless of gender, and form the core of the diagnostic questionnaire.

Figure 5.1 shows the breakdown of 17 questions into 4 categories. A tabular version of the 17 questions and their respective answers can be found in Appendix 2.

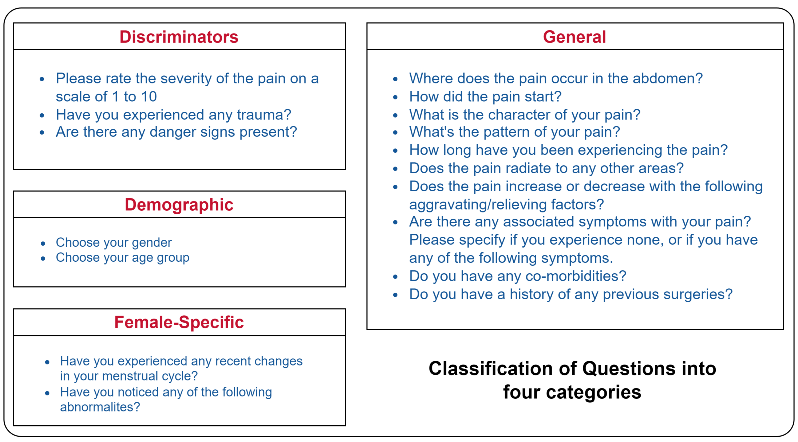

Figure 5.1: Breakdown of the 17 Questions into Categories

5.2 Probable Diagnosis and Organ of Origin

The core function of the application is to map patient responses to probable diagnoses and the organ of origin. The process for determining these outcomes relies on a pre-defined list of possible diagnoses and corresponding organs of origin.

5.2.1 Diagnosis Mapping

The physicians provided a list of 29 possible diagnoses related to abdominal pain. Each diagnosis is associated with a specific set of responses to the 17 questions from the protocol. To identify a potential diagnosis, the patient's responses are compared against this set of known answer combinations. If the patient's responses align with one of these pre-defined answer sets, a match is established for the probable diagnosis. The Figure 5.2 below depicts the bar chart for the number of possible combinations of answers for the set of 29 probable diagnoses, the top diagnosis can have more than **5,000 possible combinations of answers**.

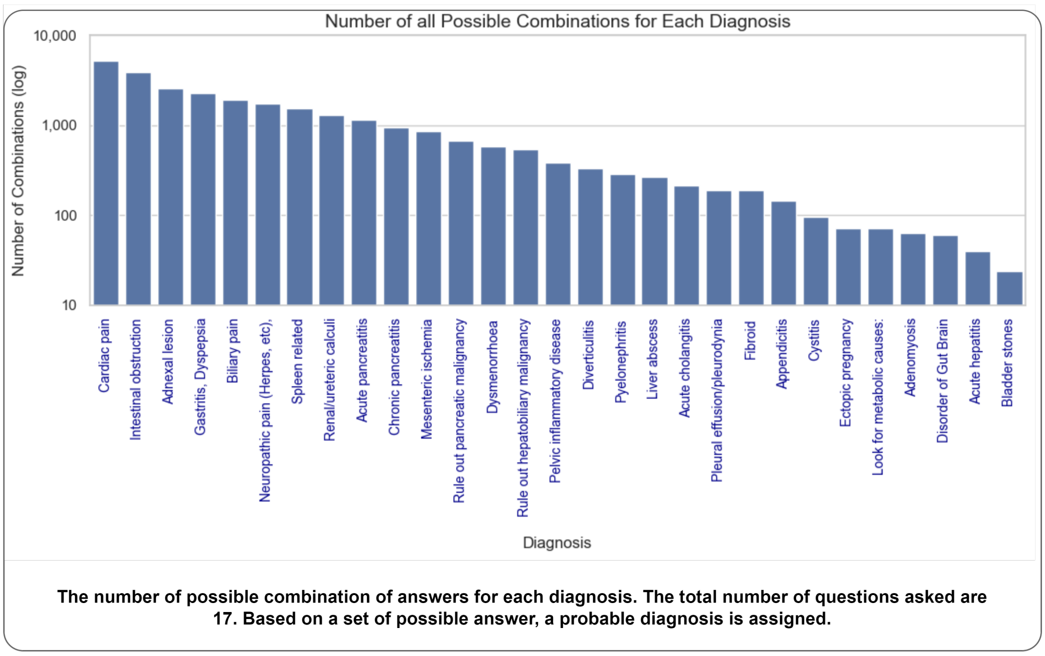

Figure 5.2: Possible Combinations of Answers

5.2.2 Organ-of-Origin Identification

The application performs the identification of the organ of origin for abdominal pain. This identification is done via the mapping established by linking 29 diagnoses to a set of 19 possible organs of origin. Once a probable diagnosis is identified, the system determines the corresponding organ of origin using a deterministic mapping approach. This allows the system to be interpretable.

The relationships between questions, diagnoses, and organs of origin are represented as a graph structure. Each diagnosis is connected to potential answers for the 17 questions, forming distinct paths in the graph. The graphical representation in Figure 5.3 provides a visual explanation of how diagnoses are reached based on patient responses. The pink nodes (center) correspond to the 29 diagnoses, while the yellow nodes represent the 19 possible organs of origin (center below). The nodes representing answers to the 12 original questions are displayed in 12 distinct colors, each corresponding to a specific question. A tabulated mapping of organs of origin to diagnoses can be found in Appendix 1.

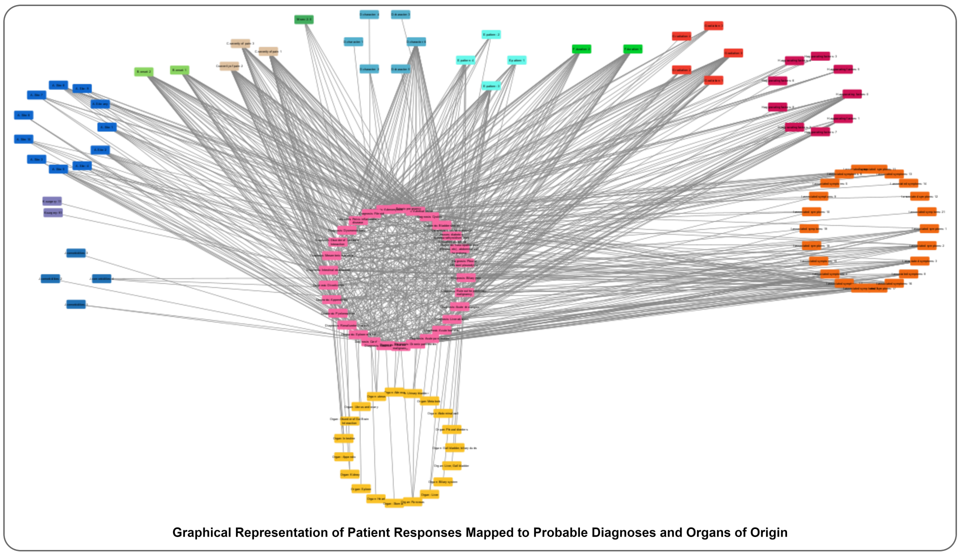

Figure 5.3: Graphical Representation of the Mapping between 12 Questions, 29 Probable Diagnoses, and 19 Organs of Origin

5.2.3 Step-by-Step Workflow

The workflow of the questionnaire begins with the patient's interaction with the system through a structured questionnaire. As shown in Figure 5.4, it starts with discriminator questions to identify emergencies; if flagged, the patient is directed to the emergency department. Otherwise, the patient proceeds to demographic questions to gather basic information regarding age and gender, followed by general questions covering universal symptoms and history. Female patients are asked additional questions related to menstrual health before proceeding to the general questions. The patient's responses are then mapped to probable diagnoses and organs of origin, as described in the previous section. The system generates a report based on the patient's responses, which contains all the answers provided by the patient and the results for the physician's reference.

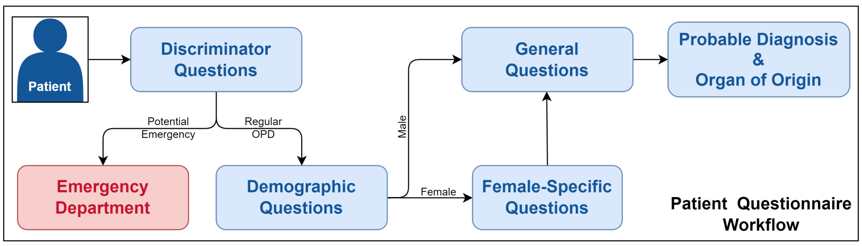

Figure 5.4: Workflow of the Patient-Questionnaire Interaction

5.3 Implementation of the Application

The application was implemented using a comprehensive technology stack. The aim was to create a system that is user-friendly, efficient, auditable, and easily accessible to physicians. The system was designed to be integrated into the hospital's existing infrastructure, allowing for seamless interaction with the application.

The implementation involved the creation of both an Android application and a web application for patient interaction. The initial version Android application was developed using Java and Kotlin in Android Studio, while the web application was built using Streamlit in Python. Streamlit is an open-source Python framework for data scientists and AI/ML engineers to deliver dynamic data apps [34].

5.3.1 Frontend Development and User Interface

The frontend development involved creating a user-friendly interface for patient interaction. The Android application was designed to be intuitive and easy to navigate, with a clean and simple layout. The web application was developed to provide a seamless experience for patients interacting with the system on a desktop or mobile device. The frontend was designed to be responsive, ensuring that the system is accessible across a wide range of devices.

The initial prototype of both the Android application was designed using Figma, a web-based design tool. Both applications were designed to follow the workflow defined in Figure 5.4. Both prototypes followed a click-through option for patient interaction, where the patient could navigate through the questionnaire by answering each question sequentially. The responses were stored in the system and used to generate the probable diagnosis and organ of origin. As seen in [29], even a simple interface with a well-defined workflow can be effective and have a significant impact on user-friendliness.

Android Application: Four screenshots of the Android application are shown in the Figure 5.5. The screenshots show the question and their respective options.

[Android Application] Four screenshots showing for questions:
A: Gender selection. **B**: Abdominal pain region selection. **C**: Radiation of pain. **D**: Character of pain.

Figure 5.5: Screenshots of the Android Application.

Web Application: The screenshots of the web application are shown in the Figure 5.6. The screenshots for the same questions and their respective options as in the Android application are shown.

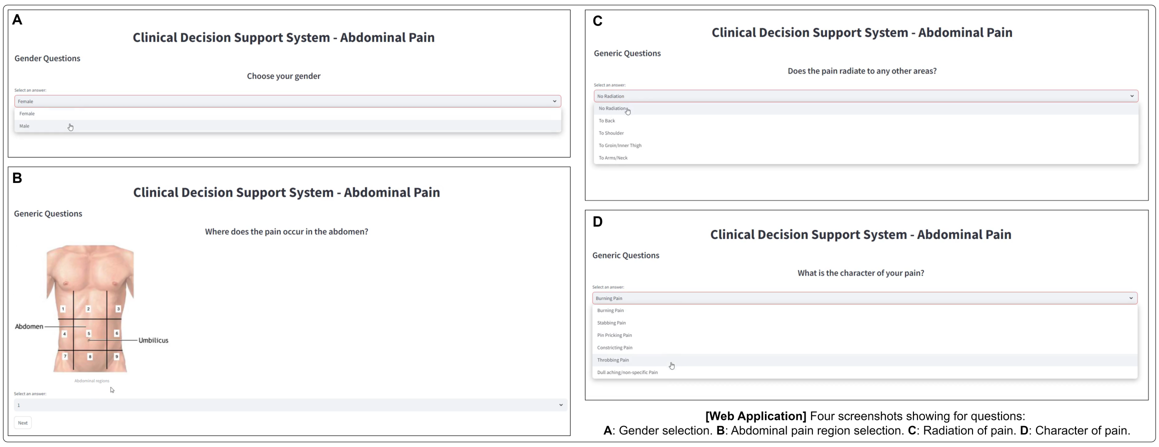

Figure 5.6: Screenshots of the Web Application.

Result Page: The result page of the web application is shown in the Figure 5.7. The result page shows two screenshots, one with a successful probable and organ of origin diagnosis, and the other with a prompt to visit emergency department, triggered by the presence of danger signs in discriminator section of the questionnaire.

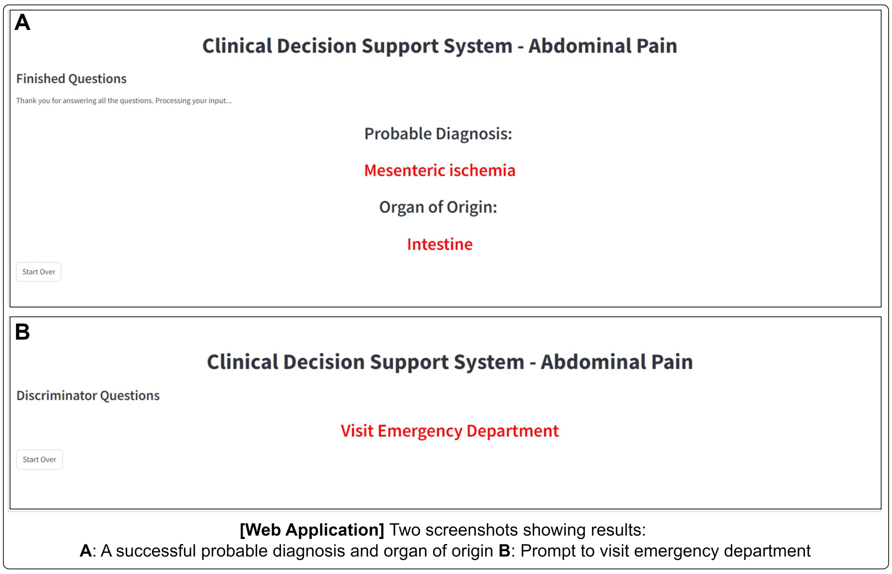

Figure 5.7: Result Page of the Web Application.

5.3.2 Backend Development

The backend of the application has two major components: the data dictionary and the conversational agent. The two components work together via a Python script. The entire backend is containerized using Docker for easy deployment and management.

5.3.2.1 Data Dictionary

Multiple data dictionaries contain the mapping between patient responses and probable diagnoses and organs of origin. These dictionaries are used to identify the probable diagnosis and organ of origin based on the patient's responses. Dictionaries are stored in JSON format for easy access and retrieval.

5.3.2.2 Conversational Agent

The conversational agent is responsible for interacting with the patient and guiding them through the questionnaire. The agent asks questions based on the patient's responses and stores the answers for further processing. The design of the conversational agent is shown in Figure 5.8.

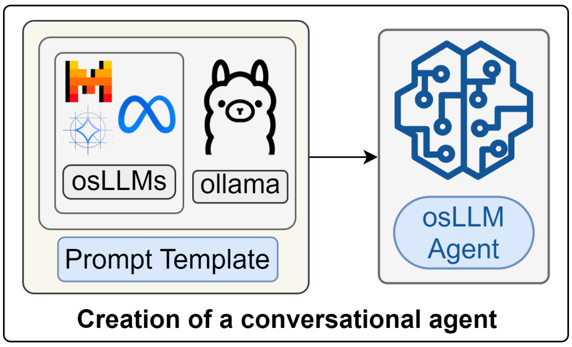

Figure 5.8: Creation of Conversational Agent

The conversational agent is created using the open source Large Language Models (osLLMs) like Mistral (`Mistral 7B`)[35], Llama (`Llama 3.1 8B`) [36], Gemma (`Gemma 2 9B`) [37], and others. The models were selected based on their performance and parameter size. From the above-mentioned models, `Gemma 2 9B` was selected for the conversational agent based on its performance and parameter size. Prompt engineering [38] is used to provide the conversational agent with the necessary context to extract the patient's responses as accurately as possible after asking each question with its respective answer choices.

The conversational agent is hosted and managed by using Ollama [39], an open-source platform that allows for the deployment and management of large language models. Ollama provides an efficient environment for running these models locally, ensuring privacy and faster processing. The model is queried with the patient's responses using a REST API. Nvidia RTX A5000 GPU were used for loading the model and processing the patient's responses.

5.3.3 System Design

The system design of the application is shown in Figure 5.9. The system consists of two main components: the frontend and the backend. The frontend includes the Android and web applications, while the backend includes the data dictionary and the conversational agent. The containerized backend can be deployed on a local server or cloud platform such as AWS or Google Cloud.

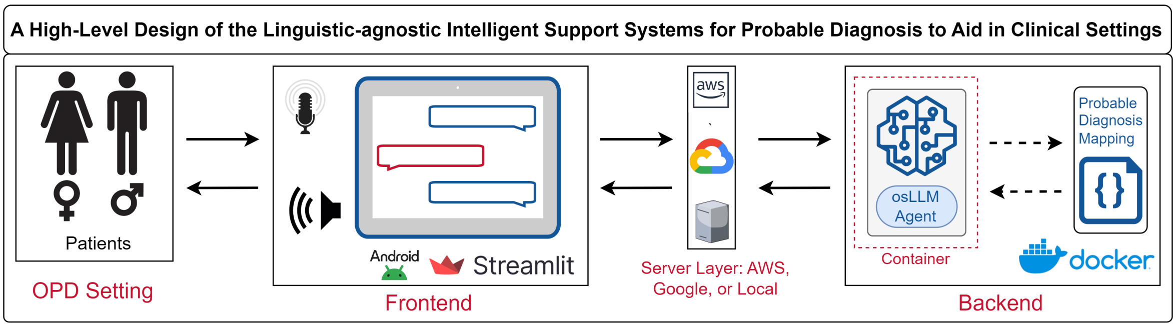

Figure 5.9: High-Level Design of the application

5.3.4 Augmented Workflow

The high-level augmentation of application into the existing workflow of evaluation of abdominal pain is shown in Figure 2.1 (B), depicted by the dotted box. A more detailed view of the augmented workflow is shown in the Figure 5.10.

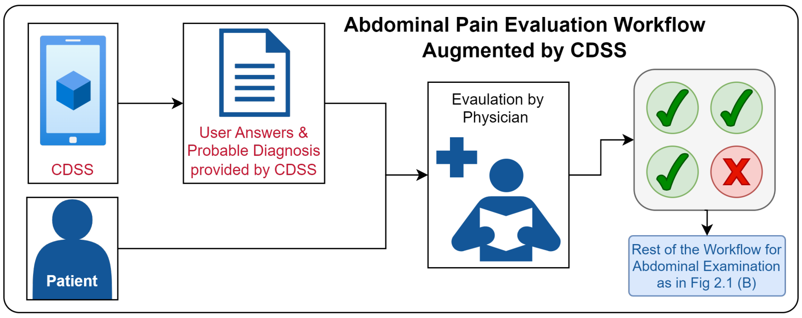

Figure 5.10: Application Augmented Workflow for Abdominal Pain Evaluation

The patient interacts with the application through the Android or web application, answering a set of structured questions. The responses are processed by the conversational agent, which identifies the probable diagnosis and organ of origin based on the patient's responses. The system generates a report containing the patient's responses, probable diagnosis, and organ of origin. The report is then reviewed by the physician, who uses the information to further the diagnostic process. The auditing of responses and the probable diagnosis is done by the physician to ensure the accuracy of the system. In case of any discrepancies, the physician can ask the patient for additional information to refine the diagnosis.

5.3.5 Challenges & Mitigations

One of the challenges of any AI-based application is the issue of explainability. The compartmentalization of the system ensures that the components can be updated independently without affecting the overall functionality of the system. It also ensures that the system can be easily audited and the output of each component can be tracked providing transparency, explainability, and interpretability. In case of any conflicting diagnosis, the output of the system can be easily traced back to the patient's responses and the mapping between the responses and the probable diagnosis.

5.4 Evaluation

Currently, the application is in the final stages of development and is yet to be deployed for clinical evaluation. The evaluation process will involve a beta testing with a small group of physicians. The physicians will interact with the system and provide feedback on the system's usability, experience, and overall performance. The feedback will be used to refine the system further before a full-scale deployment. A comprehensive evaluation will be conducted to assess the impact of the application on the diagnostic process for abdominal pain. The evaluation of impact will be performed by comparing no system (baseline), option and click based system with help of a healthcare staff, and a fully conversational system in OPD setting.

6 Conclusion & Future Work

6.1 Conclusion

The evaluation of abdominal pain remains one of the most diagnostically challenging tasks in clinical practice, requiring systematic assessment across multiple dimensions. In high-volume settings like the Department of Gastroenterology and Human Nutrition at AIIMS, New Delhi, where physicians handle over 1,35,000 OPD cases in 2022-23 alone, the need for automated system to assist physician in initial patient evaluation is essential. While fully AI-driven CDSS offer potential solutions, their black-box nature and autonomous decision-making raise concerns about explainability and accountability in healthcare settings.

This application presents an approach that combines the benefits of structured diagnostic protocols with modern AI capabilities. The pplication uses a deterministic mapping approach to link patient responses from a 17-question diagnostic protocol to one of 29 probable diagnoses and 19 organs of origin for abdominal pain in OPD settings. The system is designed to assist physicians in initial physician-patient interaction phase of the clinical workflow. The system's compartmentalized architecture, featuring both Android and web-based interfaces backed by a containerized backend with open source large language models (osLLMs) powered conversational agent, ensures transparency and traceability in the results generated.

6.2 Future Work

Future development of the system will focus on several key areas. First, the osLLM agents will be added to the current prototypes of click and option-based input system to provide a conversational interface for patients. This will involve integrating the osLLM agents with the existing system to enable conversations with patients. Additionally,the Android and web applications will be enhanced with additional features including voice-based input, and final reports generation.

Current plan is to encorporate the support of openCHA framework [40] for empowering the conversational agent with voice-based functionality and external knowledge support with built-in translation tools as a part of the system. This starting point will be used to develop a more comprehensive linguistic-agnostic conversational agent capable of handling diverse patient populations.

The final phase will involve comprehensive evaluation at AIIMS. This will include rigorous testing of the system's impact on patient flow, and diagnostic efficiency. The evaluation will also assess the impact of the system on the clinical workflow for abdominal pain evaluation.

A further future direction is the design and development of the linguistic-agnostic conversational models. This will involve design and development of models leveraging the diverse patient population at AIIMS to create robust models capable of handling various Indian languages, accents, and dialects.

Bibliography

[1] T. A. Cohen, V. L. Patel, and E. H. Shortliffe. Introducing ai in medicine. In *Intelligent Systems in Medicine and Health: The Role of AI*, pages 3–20. Springer, 2022.

[2] R. S. Ledley and L. B. Lusted. Reasoning foundations of medical diagnosis: symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science*, 130(3366):9–21, 1959.

[3] A. Rajkomar, J. Dean, and I. Kohane. Machine learning in medicine. *New England Journal of Medicine*, 380(14):1347–1358, 2019.

[4] S. L. Gans, M. A. Pols, J. Stoker, M. A. Boermeester, and E. S. Group. Guideline for the diagnostic pathway in patients with acute abdominal pain. *Digestive surgery*, 32(1):23–31, 2015.

[5] C. A. Mealie, R. Ali, and D. E. Manthey. Abdominal Examination [online]. 2024. Last Updated: 2024-05-25. URL: https://www.ncbi.nlm.nih.gov/books/NBK459220/.

[6] S. L. Cartwright and M. P. Knudson. Evaluation of acute abdominal pain in adults. *American family physician*, 77(7):971–978, 2008.

[7] Differential Diagnosis [online]. Last Updated: 2022-01-27. URL: https://my.clevelandclinic.org/health/diagnostics/22327-differential-diagnosis.

[8] S. Visweswaran, A. J. King, and G. F. Cooper. Integration of ai for clinical decision support. In *Intelligent Systems in Medicine and Health: The Role of AI*, pages 285–308. Springer, 2022.

[9] T. Panch, P. Szolovits, and R. Atun. Artificial intelligence, machine learning and health systems. *Journal of global health*, 8(2), 2018.

[10] J. A. Osheroff, J. M. Teich, B. Middleton, E. B. Steen, A. Wright, and D. E. Detmer. A roadmap for national action on clinical decision support. *Journal of the American medical informatics association*, 14(2):141–145, 2007.

[11] D. L. Hunt, R. B. Haynes, S. E. Hanna, and K. Smith. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *Jama*, 280(15):1339–1346, 1998.

[12] D. Bobrow. Natural language input for a computer problem solving system. 01 1964.

[13] D. Jurafsky. Speech and language processing, 2000.

[14] C. J. Haug and J. M. Drazen. Artificial intelligence and machine learning in clinical medicine, 2023. *New England Journal of Medicine*, 388(13):1201–1208, 2023.

[15] J. Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45, 1966.

[16] L. Clark, N. Pantidi, O. Cooney, P. Doyle, D. Garaialde, J. Edwards, B. Spillane, E. Gilmartin, C. Murad, C. Munteanu, et al. What makes a good conversation? challenges in designing truly conversational agents. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–12, 2019.

[17] J. Clusmann, F. R. Kolbinger, H. S. Muti, Z. I. Carrero, J.-N. Eckardt, N. G. Laleh, C. M. L. Löffler, S.-C. Schwarzkopf, M. Unger, G. P. Veldhuizen, et al. The future landscape of large language models in medicine. *Communications medicine*, 3(1):141, 2023.

[18] T. A. Cohen, V. L. Patel, and E. H. Shortliffe. *Intelligent systems in medicine and health: The role of AI.* Springer, 2022.

[19] Ministry of Education, Government of India. Indian Languages [online]. URL: https://www.education.gov.in/sites/upload_files/mhrd/files/upload_document/languagebr.pdf.

[20] National Language Translation Mission, Ministry of Electronics and Information technology, Government of India. Bhashini [online]. URL: https://bhashini.gov.in/.

[21] Indian Institute of Technology, Madras. AI4Bharat [online]. URL: https://ai4bharat.iitm.ac.in/.

[22] Folk Computing [online]. URL: https://kutumlab.github.io/folk-comp/.

[23] All India Institute of Medical Sciences (AIIMS). 67th annual report 2022-2023, February 2024. Last updated on: 02 Feb 2024. URL: https://www.aiims.edu/images/pdf/annual_reports/english1.pdf.

[24] All India Institute of Medical Sciences (AIIMS). AIIMS OUT PATIENT DEPARTMENT (OPD) SERVICES [online]. Last Updated: 2024-05-25. URL: https://www.aiims.edu/aiims/hosp-serv/citizen-charter/opd-services.htm.

[25] E. Shortliffe. *Computer-based medical consultations: MYCIN*, volume 2. Elsevier, 2012.

[26] E. H. Shortliffe and B. G. Buchanan. A model of inexact reasoning in medicine. *Mathematical biosciences*, 23(3-4):351–379, 1975.

[27] M. A. Musen, B. Middleton, and R. A. Greenes. Clinical decision-support systems. In *Biomedical informatics: computer applications in health care and biomedicine*, pages 795–840. Springer, 2021.

[28] D. Demner-Fushman, W. W. Chapman, and C. J. McDonald. What can natural language processing do for clinical decision support? *Journal of biomedical informatics*, 42(5):760–772, 2009.

[29] H.-P. Eich and C. Ohmann. Internet-based decision-support server for acute abdominal pain. In *Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, pages 103–112. Springer, 1999.

[30] P. Lee, S. Bubeck, and J. Petro. Benefits, limits, and risks of gpt-4 as an ai chatbot for medicine. *New England Journal of Medicine*, 388(13):1233–1239, 2023.

[31] A. Karthikesalingam and P. Natarajan. Amie: A research ai system for diagnostic medical reasoning and conversations. *Google Research*, 2024.

[32] R. Johnson, M. M. Li, A. Noori, O. Queen, and M. Zitnik. Graph artificial intelligence in medicine. *Annual Review of Biomedical Data Science*, 7, 2024.

[33] K. Miller. Should AI Models Be Explainable? That depends. [online]. 2021. Last Updated: 2021-05-16. URL: https://hai.stanford.edu/news/should-ai-models-be-explainable-depends.

[34] Streamlit [online]. URL: https://streamlit.io/.

[35] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. d. l. Casas, F. Bressand, G. Lengyel, G. Lample, L. Saulnier, et al. Mistral 7b. *arXiv preprint arXiv:2310.06825*, 2023.

[36] A. Dubey, A. Jauhri, A. Pandey, A. Kadian, A. Al-Dahle, A. Letman, A. Mathur, A. Schelten, A. Yang, A. Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

[37] G. Team, M. Riviere, S. Pathak, P. G. Sessa, C. Hardin, S. Bhupatiraju, L. Hussenot, T. Mesnard, B. Shahriari, A. Ramé, et al. Gemma 2: Improving open language models at a practical size. *arXiv preprint arXiv:2408.00118*, 2024.

[38] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1–35, 2023.

[39] Ollama [online]. URL: https://ollama.com/.

[40] M. Abbasian, I. Azimi, A. M. Rahmani, and R. Jain. Conversational health agents: A personalized llm-powered agent framework. *arXiv preprint arXiv:2310.02374*, 2023.

Appendix

1 Probable Diagnosis and Organ of Origin

Given below is a table mapping the 19 organs of origin to the 29 probable diagnoses. The Table 1 is structured such that each row corresponds to a organ of origin, and each value in column `Probable Diagnoses` corresponds to a (list of) probable diagnosis(es) that can originate from that organ.

No.	Organ of Origin	Probable Diagnoses
1.	Gall Bladder, Biliary Ducts	Biliary Pain
2.	Liver & Gall Bladder	Rule Out Hepatobiliary Malignancy
3.	Biliary System	Acute Cholangitis
4.	Liver	1. Liver Abscess 2. Acute Hepatitis
5.	Pancreas	1. Acute Pancreatitis 2. Chronic Pancreatitis 3. Rule Out Pancreatic Malignancy
6.	Stomach	Gastritis, Dyspepsia
7.	Heart	Cardiac Pain
8.	Spleen	Spleen Related
9.	Kidney	1. Renal/Ureteric Calculi 2. Pyelonephritis
10.	Appendix	Appendicitis
11.	Intestine	1. Diverticulitis 2. Intestinal Obstruction 3. Mesenteric Ischemia
12.	Disorder of Gut Brain Interaction	Disorder of Gut Brain Interaction
13.	Uterus & Ovary	Dysmenorrhoea
14.	Uterus	1. Pelvic Inflammatory Disease 2. Fibroid 3. Adenomyosis
15.	Adnexa	1. Ectopic Pregnancy 2. Adnexal Lesion
16.	Urinary bladder	1. Cystitis 2. Bladder stones
17.	Metabolic	Look for Metabolic Causes: Diabetes, Hyperparathyroidism, Lead Intoxication, Porphyria
18.	Abdominal Wall	Neuropathic Pain (Herpes, etc), Abdominal Wall Hematoma
19.	Pleural Disorders	Pleural Effusion/Pleurodynia

Table 1: Mapping of Organs of Origin to Probable Diagnoses.

2 Questionnaire

The 17 questions used in the diagnostic questionnaire are listed below along with the possible answers. The Table 2 is structured such that each row corresponds to a question, and each value in column `Possible Answers` corresponds to a (list of) possible answer(s) to that question.
The Figure 1 below shows the 9 regions of the abdomen and 4 patterns of pain that are referred to in the questionnaire.

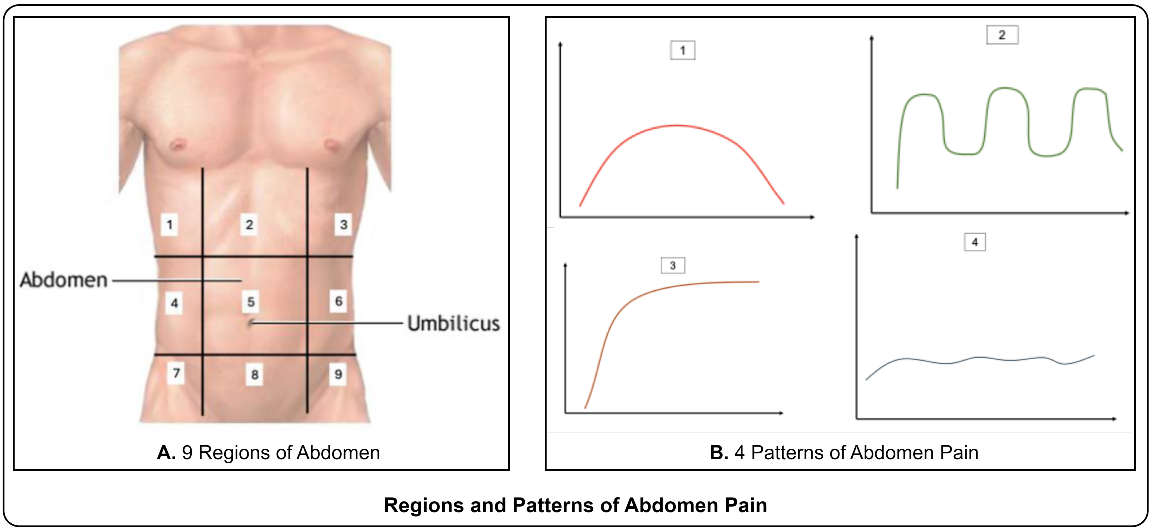

Figure 1: Regions of the Abdomen and Patterns of Pain.

No.	Question	Possible Answers (seperated by semi-colon)
1	Please rate the severity of the pain on a scale of 1 to 10	1-3; 4-7; 8-10
2	Have you experienced any Trauma?	Yes; No
3	Are there any danger signs present?	Light Headedness; Altered Sensorium; Respiratory Distress; None
4	Choose your gender	Male; Female
5	Choose your age group	0-15; 15-25; 25-45; 45-60; 60+
6	Where does the pain occur in the abdomen?	<Image of 9 regions of Abdomen>
7	How did the pain start?	Over Minutes to Hours (Acute); Over Hours to Days (Insidious)
8	What is the character of your pain?	Burning Pain; Stabbing Pain; Pin Pricking Pain; Constricting Pain; Throbbing Pain; Dull aching/non-specific Pain
9	What's the pattern of your pain?	<Image of 4 patterns of pain>
10	How long have you been experiencing the pain?	Less than 3 months; More than 3 months
11	Does the pain radiate to any other areas?	No Radiation; To Back; To Shoulder; To Groin/Inner Thigh; To Arms/Neck
12	Does the pain increase or decrease with the following aggravating/relieving factors?	No Aggravating/Relieving Factors; With Food Intake; Bending Forward; Passing Stool; Passing Urine; Menstruation; Bending Sideways; Deep Inspiration; Walking and Exercise
13	Are there any associated symptoms with your pain? Please specify if you experience none, or if you have any of the following symptoms.	None; Lump; Fever or Chills; Nausea and/or Vomiting; Abdominal Bloating; Constipation; Diarrhoea; Blood in Stools/Black Stools; Jaundice; Burning Micturition; Blood in Urine; Weight Loss; Loss of Appetite; Stress, Anxiety, Depression, Palpitation; Shortness of Breath; Swelling in Neck, Axilla; Skin Changes Over Abdominal Wall
14	Do you have any comorbidities ?	None; Diabetes; Heart Disease; Kidney Disease; Gall Stones
15	Do you have a history of any previous surgeries?	None; Gall Bladder; Intestine; Kidney; Uterus
16	Have you experienced any recent changes in your menstrual cycle?	Changes in Periods; Absence of Periods; None
17	Have you noticed any of these abnormalites?	Abnormal Vaginal Bleeding; Foul Smelling Discharge; None

Table 2: Questionnaire for the Diagnostic Protocol.

3 Demonstration for the Implemented Application

The demonstration videos of the implemented application are available here.

4 Additional Sources

The report was compiled using the TUM dissertation/PhD thesis LaTeX template as the base template. Various additional changes were made to the template to suit the requirements of the project report. The template is available under the CC-BY-4.0 license.

The Figure 2.1 (A) was source from Wikimedia Commons and is available here. The image of nine abdomen regions in Figure 5.6 (B) was sourced from A.D.A.M. Consumer Health.

5 Links

The following is the list of all the links mentioned in the report:

- Folk Computing Website: `https://kutumlab.github.io/folk-comp/`
- Demonstration Videos: `https://kutumlab.github.io/abdominal-pain-cdss/`
- Thesis Template: `https://github.com/TUM-LIS/tum-dissertation-latex`
- Wikimedia Commons: `https://commons.wikimedia.org/wiki/File:Gray1220-es.svg`
- A.D.A.M. Consumer Health: `https://www.adam.com/`