

# INNOVATING HEALTHCARE DATA ANALYTICS

---

Tsinghua University  
IEDE-Spring 2024

# THE TEAM



Name: Zunaira Irfan  
University: Binzhou  
Medical University  
Major: MBBS (Bachelors)  
IEDE-ID: 2024060  
Attendance: 10



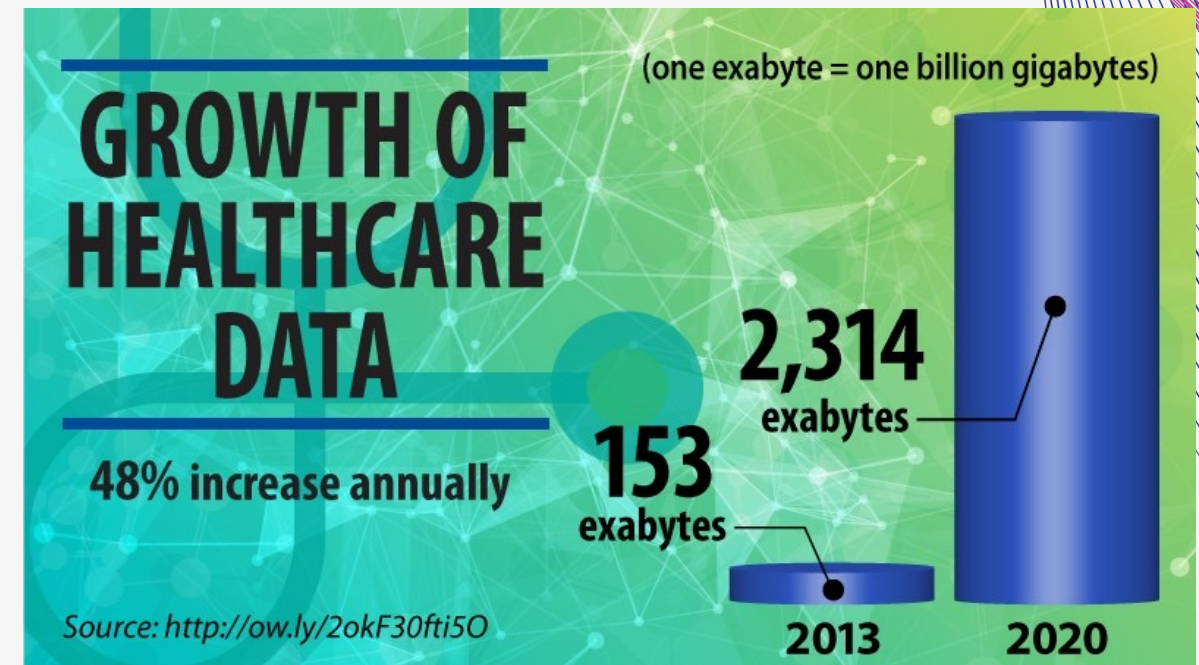
Name: Saim Liaqat Ali  
University: Binzhou Medical  
University  
Major: Medicine (MBBS)  
IEDE-ID: 2024163  
Attendance: 10



Name: Vachiraporn  
Puhwachanakarn  
University: Chulalongkorn  
University  
Major: Pharmaceutical (Master)  
IEDE-ID: 2024175  
Attendance: 10



Name: Araz  
University: Beijing  
University of Technology  
Major: Software  
Engineering (Masters)  
IEDE-ID: 2024233  
Attendance: 10



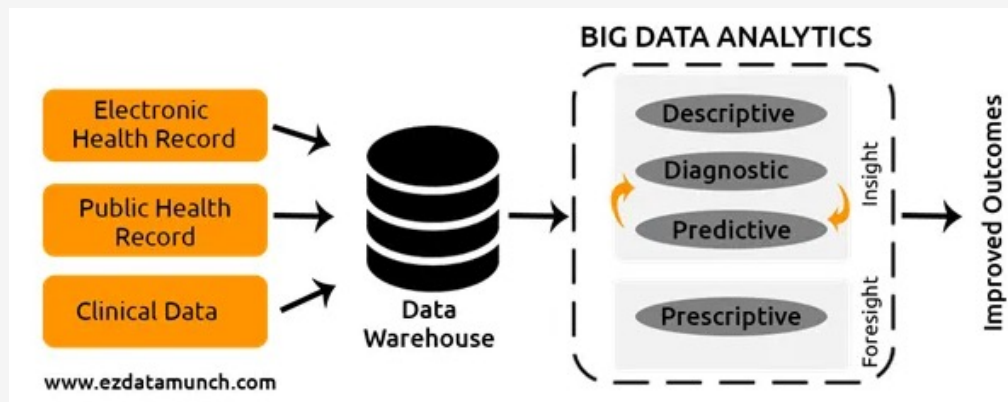
# Healthcare Data

The healthcare industry generates about 30% of the world's data. By 2020, the healthcare sector was estimated to generate 2.3 zettabytes of data, which is equivalent to 2.3 trillion DVDs of data.

Between 2020 and 2025, the total amount of global healthcare data is projected to increase from 2,300 to 10,800 exabytes, representing an annual growth rate of 36%. This is 6% faster than manufacturing, 10% faster than financial services, and 11% faster than media and entertainment.

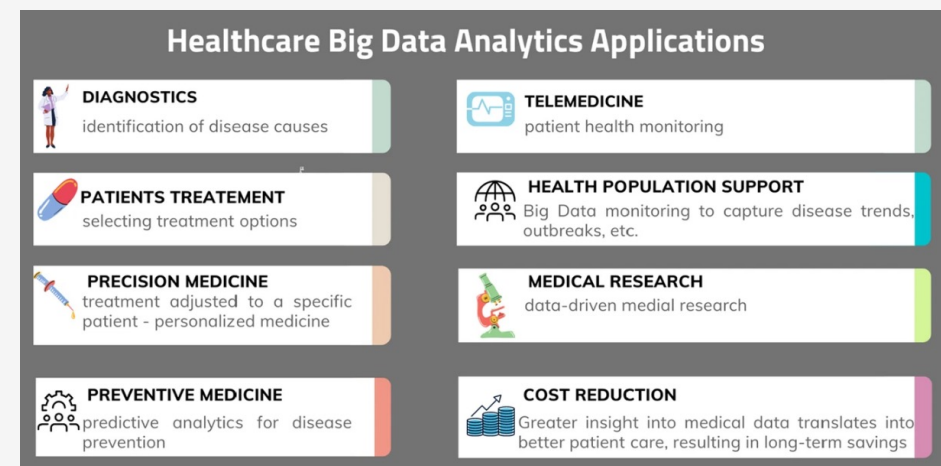
# Data Analysis in Healthcare

Healthcare data analytics, powered by AI and NLP technologies, is driving a shift towards data-driven healthcare for enhanced diagnostics and patient care, but challenges such as resource limitations impede widespread adoption despite significant data generation growth.



## WHY IS BIG DATA ANALYTICS & NLP IMPORTANT IN HEALTHCARE?

- 1 IMPROVED DIAGNOSTICS
- 2 BETTER PATIENT CARE
- 3 BETTER HEALTH OUTCOMES
- 4 IMPROVED PATIENT INTERACTIONS
- 5 INCREASED HEALTH INDICATORS
- 6 ENHANCED PATIENT EXPERIENCES
- 7 BETTER DECISION MAKING
- 8 PREVENTATIVE CARE
- 9 FACILITIES MANAGEMENT
- 10 REAL-TIME MONITORING & TELEMEDICINE



# Current **STATUS&TRENDS**

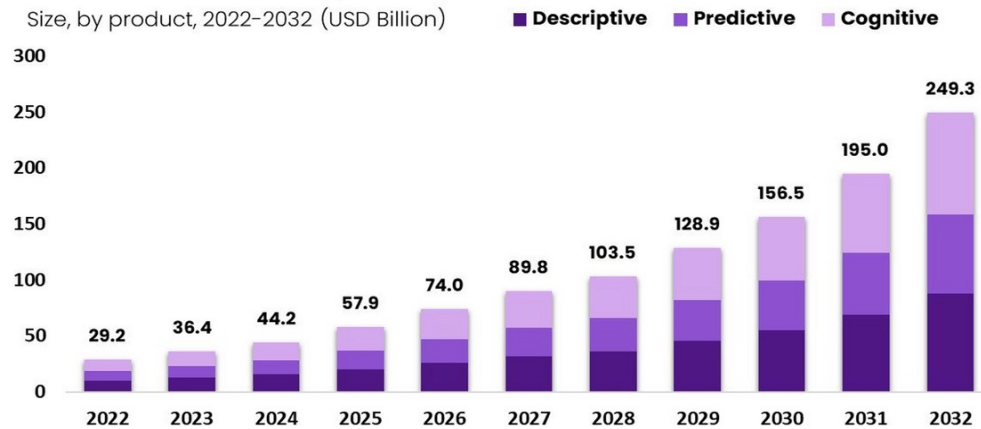
---

- Increasing Data Volume
- Advanced Analytics Techniques
- Artificial Intelligence and Machine Learning
- Focus on Real-Time Analytics
- Data Privacy and Ethics

# Healthcare Data Analytics Market

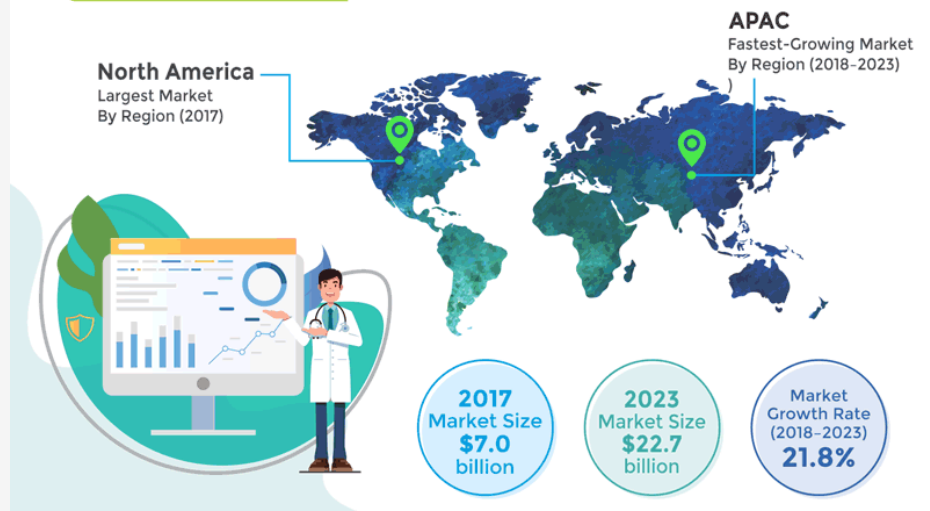
## Global Healthcare Analytics Market

Size, by product, 2022-2032 (USD Billion)



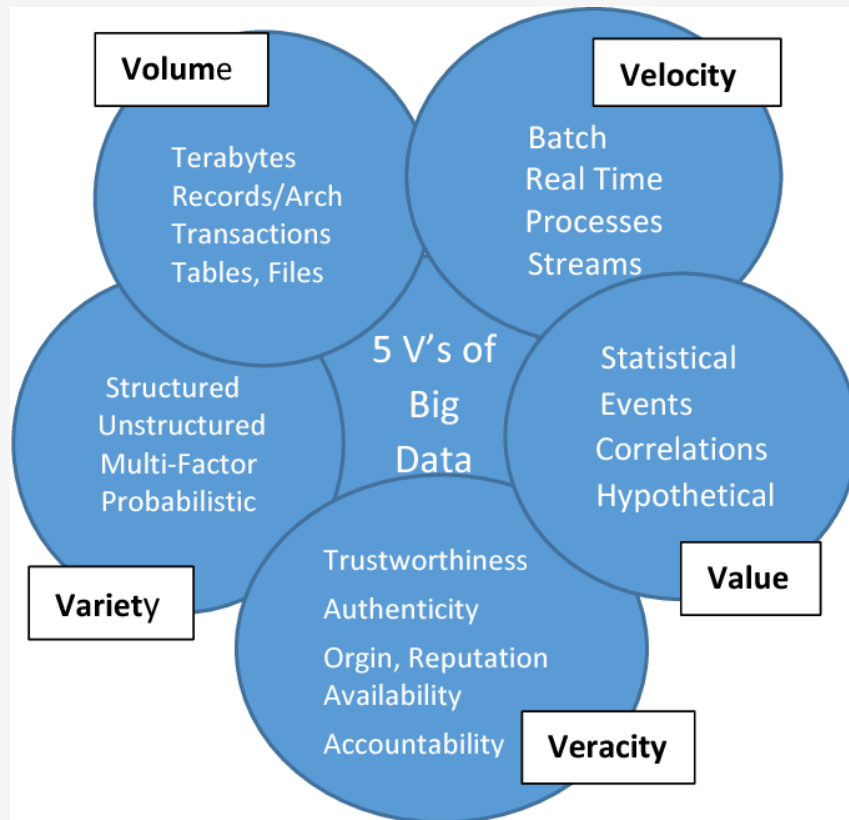
The Market will Grow At the CAGR of: **24.6%** The forecasted market size for 2032 in USD: **\$249.3B** market.us

## GLOBAL BIG DATA ANALYTICS IN HEALTHCARE MARKET



The Data Analytics market is projected to grow to \$132.9 billion by 2026, driven by AI, IoT, and real-time analytics. With rapid data growth and advanced analytics, the industry faces challenges like data quality and skills gaps but offers opportunities in business optimization and customer insights. AI advancements and privacy concerns will significantly influence future growth.

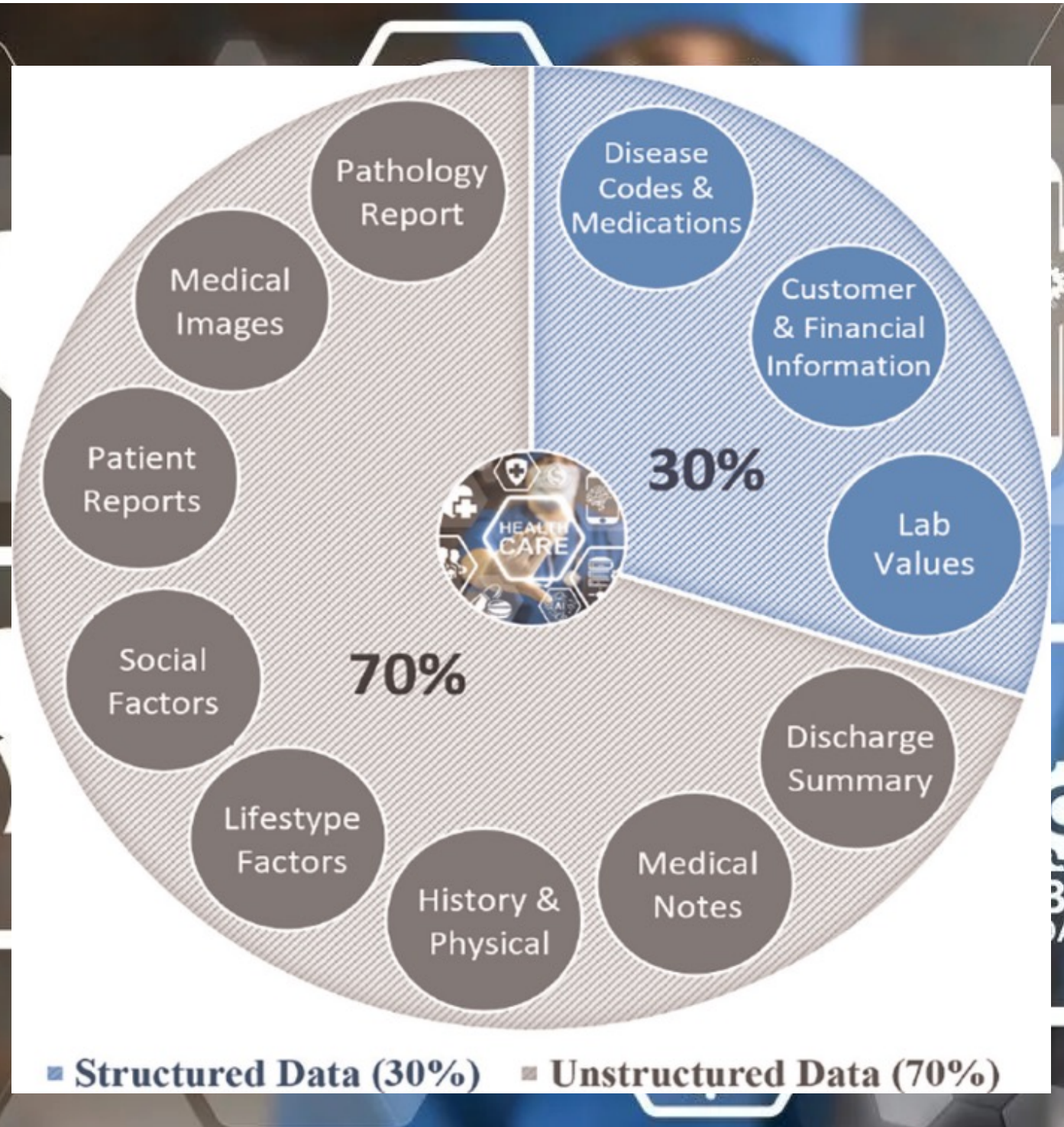
# DATA ANALYTICS CHALLENGES IN HEALTHCARE



The challenge of big data lies in effectively managing, analyzing, and deriving actionable insights from vast and diverse datasets while ensuring data quality, security, and compliance with regulatory standards.

# Problem

Most of Healthcare Data is available in unstructured format and there are different types of medical data available







# OUR IDEA

---

To make unstructured data structured and make that useful, our solution is to use LLM which can also be used to diagnose diseases and create summary of whole data at the end.



# Flow Diagram

Coversation with Patient asking questions about disease condition



History Taking



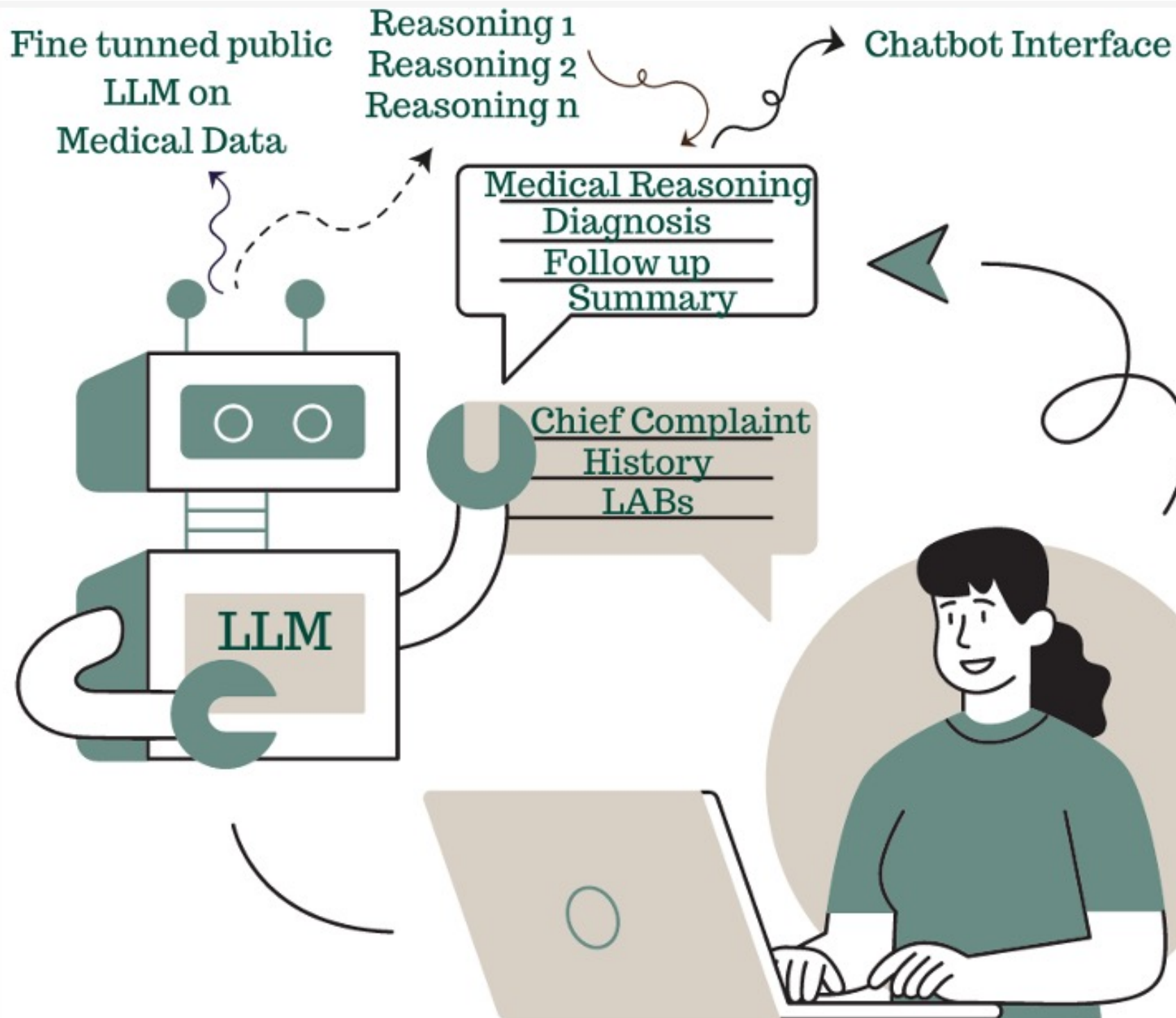
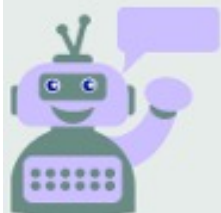
Differential Diagnosis

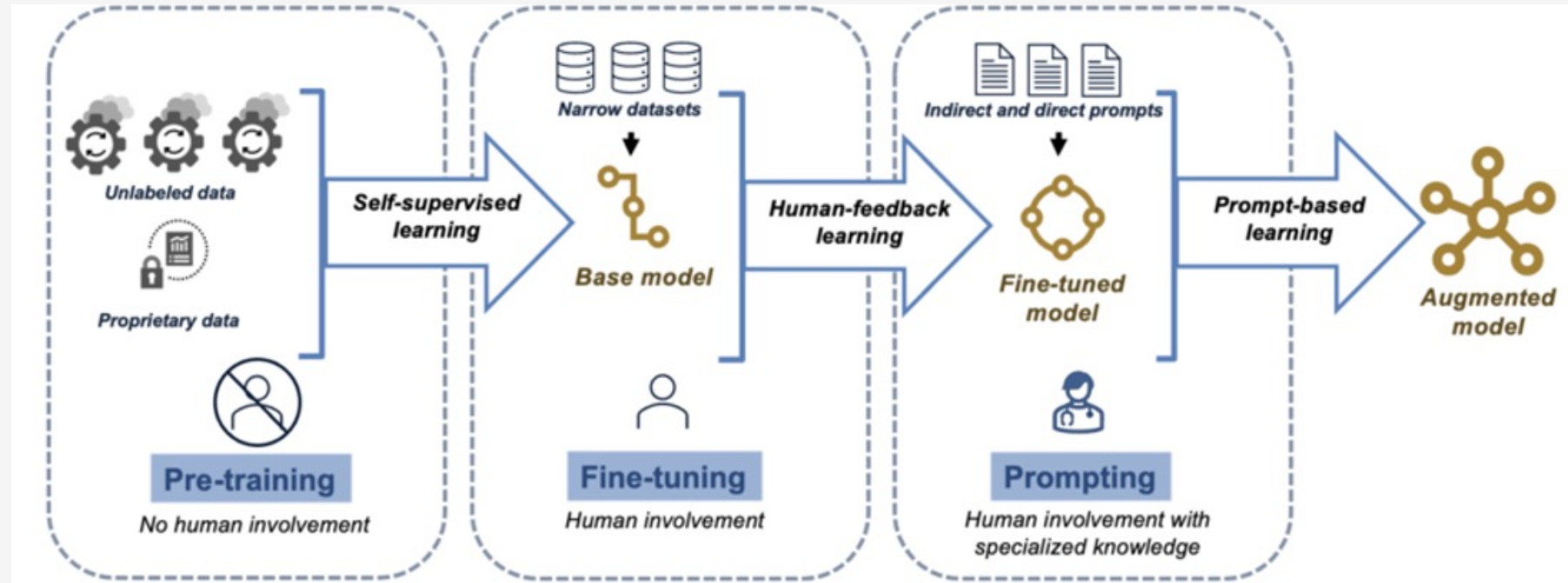


Diagnosis



Summary of Medical Records





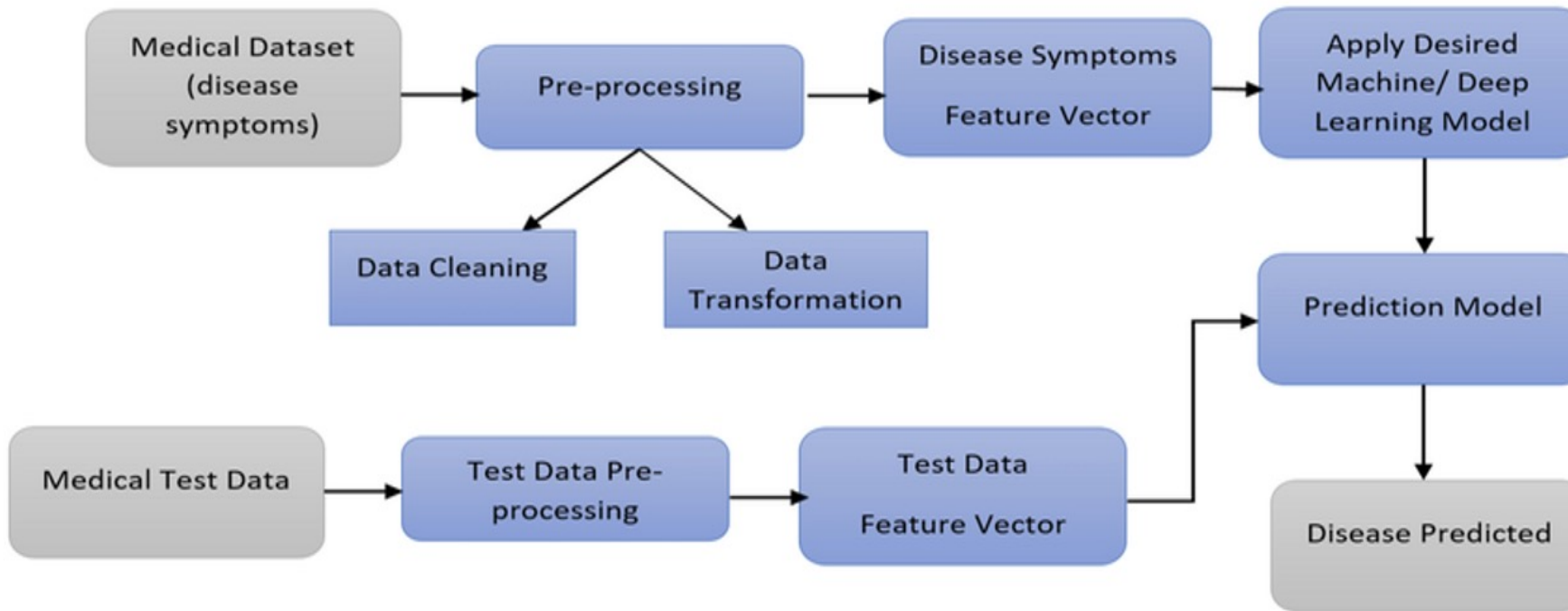
# Solutions

## 1. Fine-tuning a LLM on Medical Data

- Data Acquisition
- Data Preprocessing
- Model Training
- Model Validation

## 2. Creating a Chatbot powered by LLM

- Interface Development
- Model Integration
- Interactive Questioning
- Natural Language Understanding (NLU)

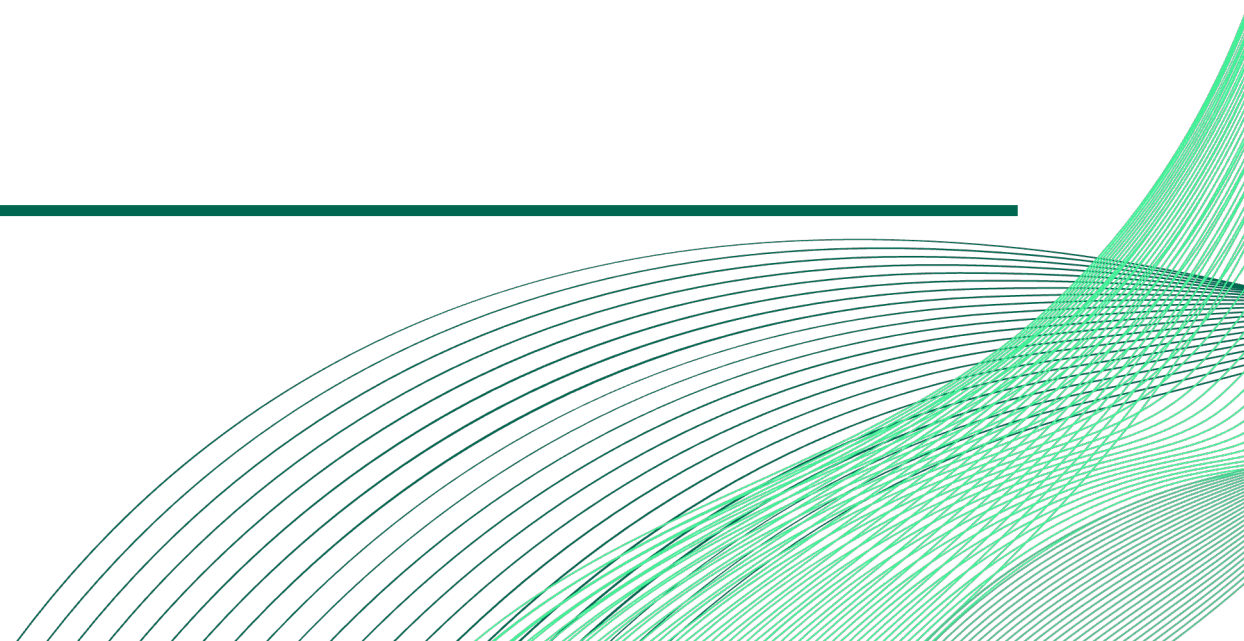


# Solutions

---

## 3. Diagnosing Diseases

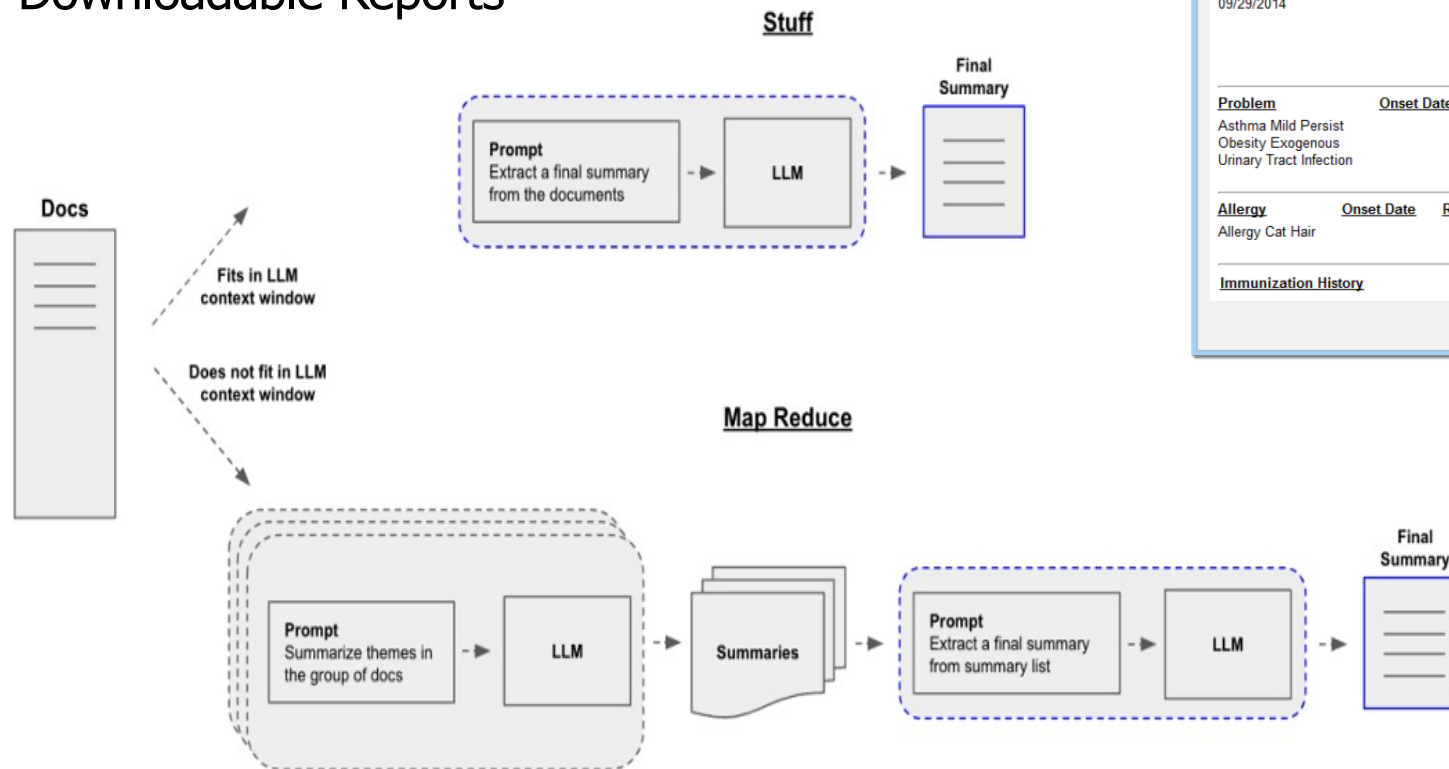
- Data Analysis for Diagnosis
- Algorithm Development
- Confidence Scoring



# Solutions

## 4. Summarizing Medical Data

- Integrate Summarization
- Extract Key Details
- Customize Format
- Downloadable Reports



Health Information Summary Report

### Health Information Summary

PCC Pediatric Test Associates  
20 Winooski Falls Way, Suite 7  
Winooski, VT 05404-2228  
(800) 722-7708

**Pebbles Flintstone**  
DOB: 09/21/05  
Sex: Female  
PCC #: 3336

**Address:**  
Wilma Flintstone  
15 Quarry Lane  
Winooski, VT 05404

**Date of Last Physical:**  
09/29/2014

**Phone:**  
Home Phone: 802-555-0194  
Work Phone: 802-555-0197  
Cell Phone: 802-555-0161  
Emg Phone: 802-555-0168

Problem	Onset Date	Status
Asthma Mild Persist		Active
Obesity Exogenous		Active
Urinary Tract Infection		Resolved

Allergy	Onset Date	Reaction	Status
Allergy Cat Hair			Active

**Immunization History**

Save as PDF Close Print

## 5. Testing and Iteration

- Conduct Testing
- Gather Feedback
- Iterate and Enhance

# FUTURE DIRECTION



Develop and test our proposed idea into mobile application and add functionalities