

AGENTIC

AI



AI
AGENT

AI Agents vs. Agentic AI: A Conceptual Taxonomy, Applications and Challenges

Cholatid Ratanatharathorn

Department of Clinical Epidemiology and Biostatistics

Research objective

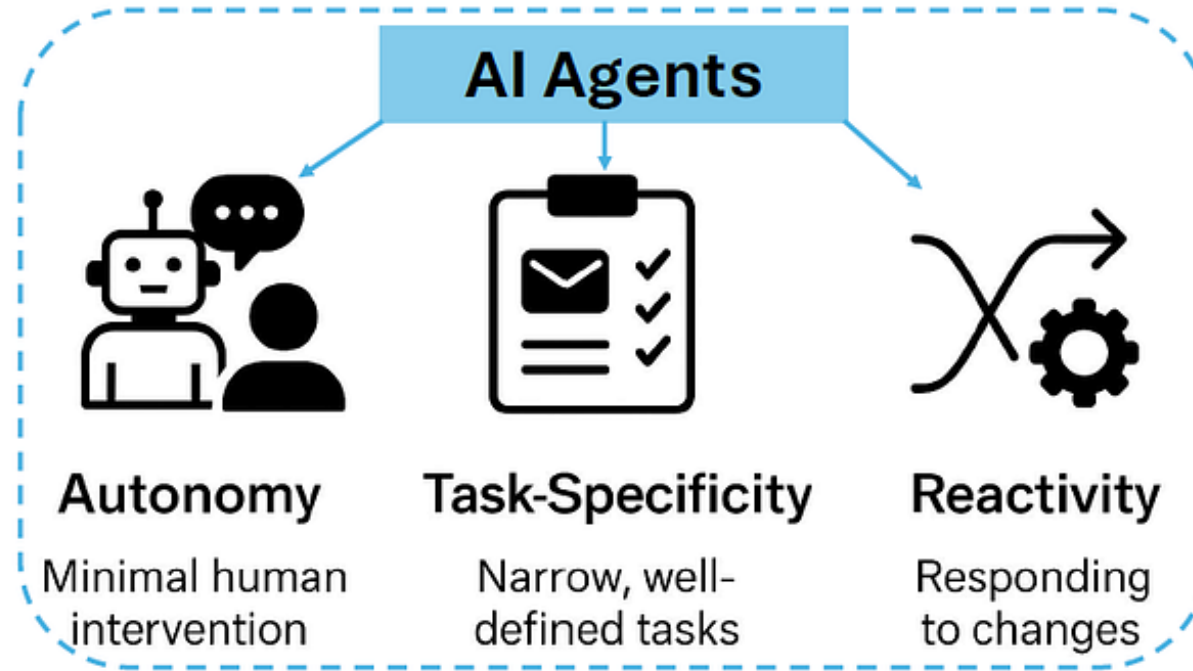


Problem: Ambiguity between "AI Agents" and "Agentic AI" causes confusion in system design and development.

Objective: Establish a clear taxonomy to differentiate these concepts, providing a roadmap for future intelligent systems.

AI Agent

An **AI Agent** is a foundational autonomous software entity engineered for narrow, goal-directed task execution within a bounded environment.



Autonomy

These systems **operate on their own** with minimal human input. They perceive, process, and act in real-time, allowing for scalable operations.

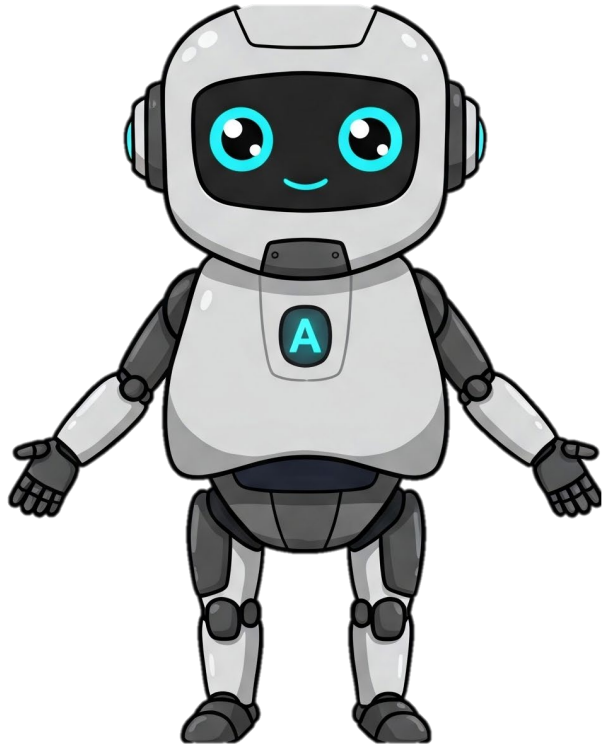
Reactivity

They **respond to environmental changes**, like user commands, and can even learn to improve their behavior over time.

Task-Specificity

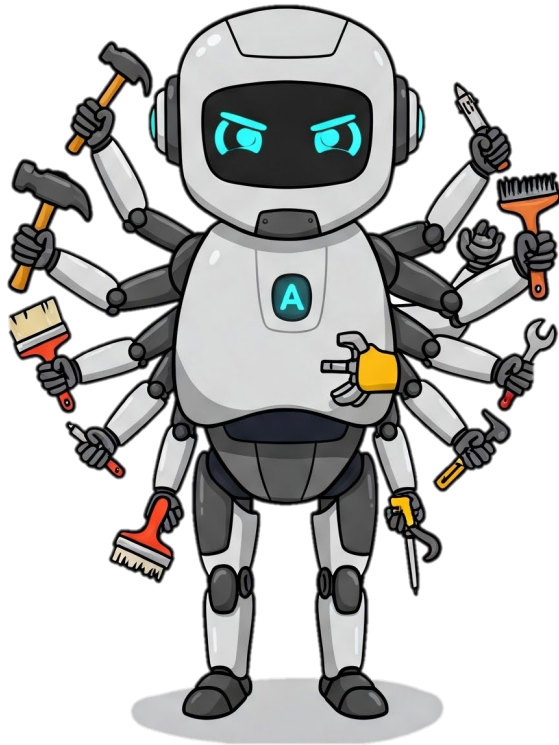
Designed for **specific tasks**, these systems excel in focused jobs, ensuring efficient, accurate, and understandable automation.

Generative AI (The Precursor):



- Purely reactive systems that produce novel content (text, images, code) based on a user prompt.
- They are **stateless** and **prompt-dependent**.
- They inherently lack the ability to interact with external tools, access real-time data, or pursue goals autonomously.

Tool-Augmented AI Agents



- The key innovation is integrating LLMs with external **tools** and **APIs** through **function calling**.
- This architecture allows the agent to break free from its static knowledge. It can now access real-time information (e.g., search the web), execute code, and interact with other software.
- Frameworks like **ReAct (Reason + Act)** exemplify this by creating a loop where the agent thinks, uses a tool, observes the result, and thinks again.

Agentic AI : Key concept

Agentic AI is not just a more advanced agent; it's a paradigm shift from an isolated entity to a **system of multiple, specialized agents** that collaborate to achieve complex, high-level goals.

Multi-Agent

Collaboration: A complex task is handled by a team of agents with unique, specialized roles

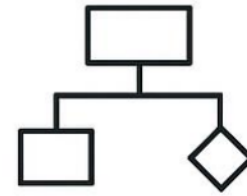
Persistent & Shared Memory:

Agents can access a shared memory to maintain context, share findings, and learn from past actions.



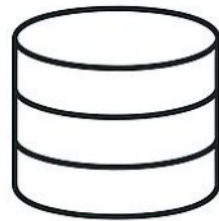
Specialized Agents

Multi-Agent Collaboration



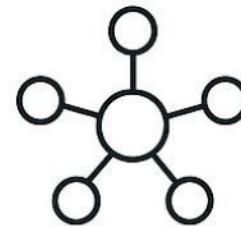
Advanced Reasoning & Planning

Task-Decomposition



Persistent Memory

Shared Context



Orchestration

System Coordination

Dynamic Task

Decomposition: A high-level, often ambiguous goal is automatically broken down into a series of concrete sub-tasks.

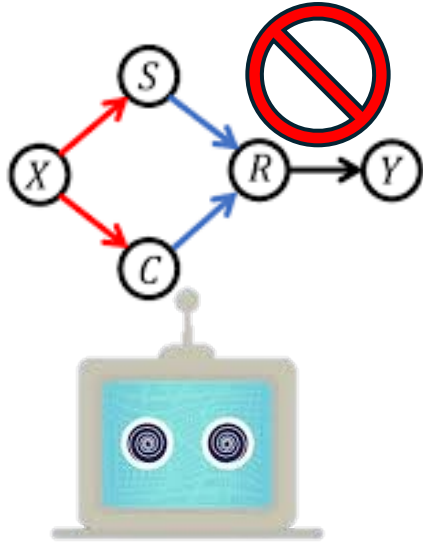
Orchestration:

A dedicated "**meta-agent**" or orchestration layer acts as a project manager.

AI Agent vs Agentic AI

Feature	AI Agents	Agentic AI
Definition	Autonomous software programs that perform specific tasks.	Systems of multiple AI agents collaborating to achieve complex goals.
Autonomy Level	High autonomy within specific tasks.	Higher autonomy with the ability to manage multi-step, complex tasks.
Task Complexity	Typically handle single, specific tasks.	Handle complex, multi-step tasks requiring coordination.
Collaboration	Operate independently.	Involve multi-agent collaboration and information sharing.
Learning and Adaptation	Learn and adapt within their specific domain.	Learn and adapt across a wider range of tasks and environments.
Applications	Customer service chatbots, virtual assistants, automated workflows.	Supply chain management, business process optimization, virtual project managers.

Challenges of AI Agent



Lack of Causal Reasoning: They excel at finding statistical correlations but don't understand true cause-and-effect, making them brittle when conditions change.

Inherited LLM Flaws: They are prone to hallucination, are sensitive to small changes in prompts, and can perpetuate biases from their training data.



Limited Long-Horizon Planning: They struggle with complex tasks requiring many sequential steps, often losing context or getting stuck in loops.

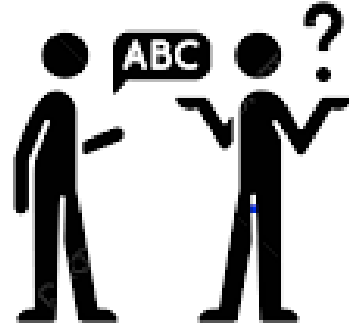


Reliability and Safety Concerns: AI Agents are not yet safe or verifiable enough for deployment in critical infrastructure

Challenges of Agentic AI

Amplified Causality Issues

Complex inter-agent interactions make cause-effect modeling difficult, hindering coordination.

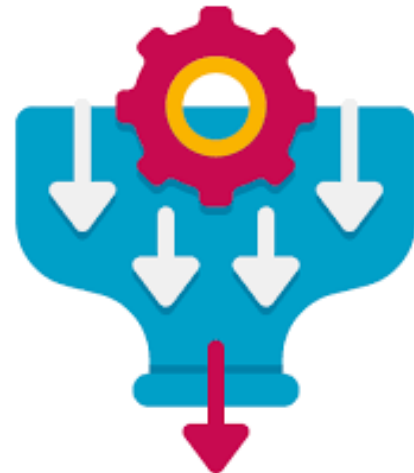


Unpredictable Emergent Behaviors

Complex interactions lead to unexpected outcomes, infinite loops, and deadlocks.

Communication Obstacles

Goal misalignment, protocol limits, and resource competition impede effective agent collaboration.



Scalability & Debugging

Tracing errors in nested agent interactions becomes exponentially harder with more agents.

Challenges of Agentic AI

Lack of Trust & Explainability

Understanding system decisions is challenging due to compounded opacity.



Increased Security Risks

More components mean more attack vectors

Ethical & Governance Complexities

Accountability, fairness, and value alignment are harder in distributed systems



Immature Foundations

Lack of standards (architectures, causal models, benchmarks) hampers reliable implementation.

POTENTIAL SOLUTIONS AND FUTURE ROADMAP

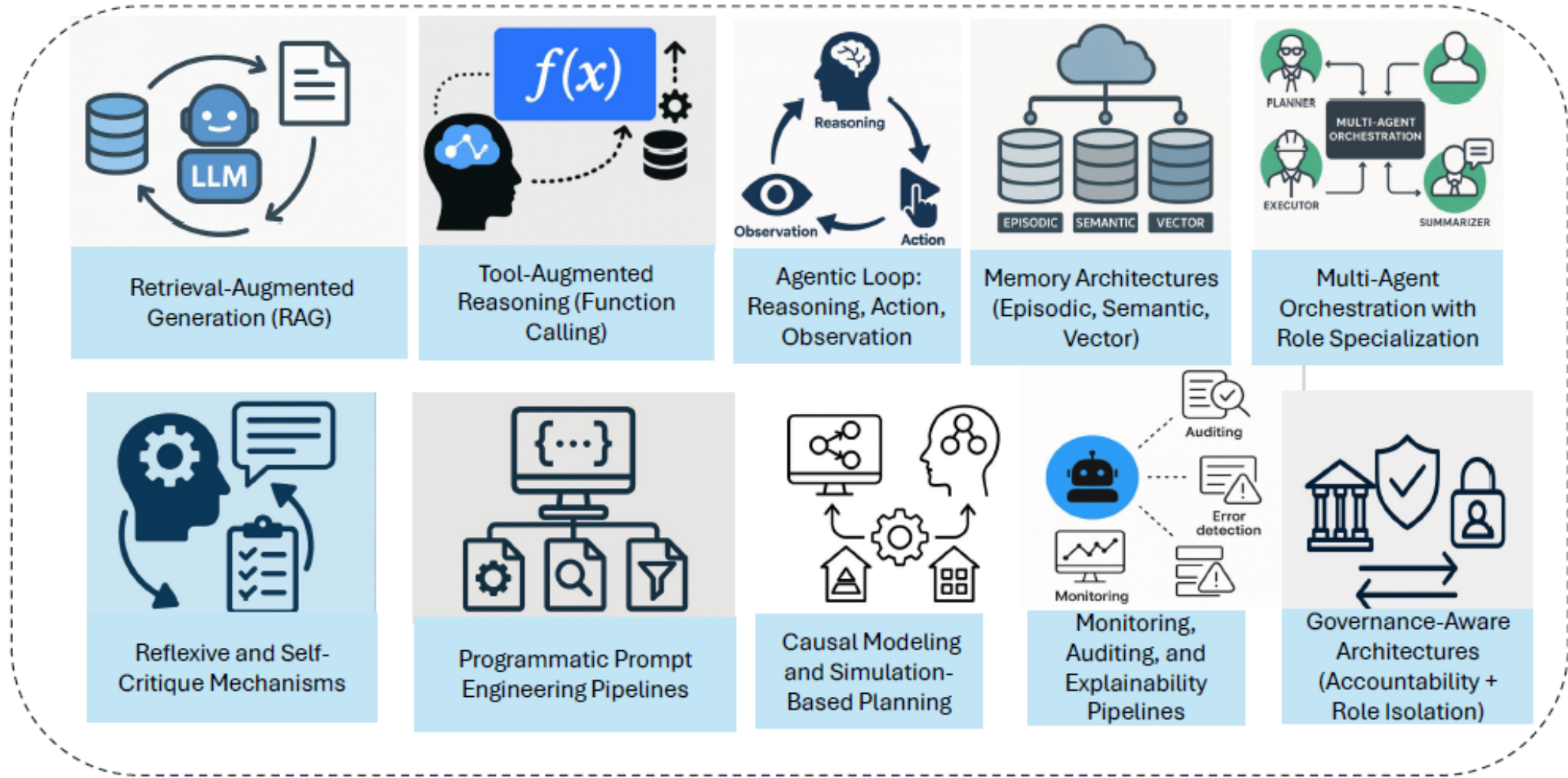


Fig. 13: Ten emerging architectural and algorithmic solutions such as RAG, tool use, memory, orchestration, and reflexive mechanisms addressing reliability, scalability, and explainability across both paradigms

Case I

Problem :

A researcher have no time to search and read about their interest everyday.

Solution :

Collect title and abstract of new publish papers of specific query term in Pubmed then summarize for a researcher daily.

API Pubmed search



National Library of Medicine
National Center for Biotechnology Information

<https://www.ncbi.nlm.nih.gov/books/NBK25499/>

Searching a Database

Go to:

Basic Searching

`esearch.fcgi?db=<database>&term=<query>`

Input: Entrez database (&db); Any Entrez text query (&term)

Output: List of UIDs matching the Entrez query

Example: Get the PubMed IDs (PMIDs) for articles about breast cancer published in Science in 2008

[https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?
db=pubmed&term=science\[journal\]+AND+breast+cancer+AND+2008\[pdat\]](https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=science[journal]+AND+breast+cancer+AND+2008[pdat])

API Pubmed search

Searching a Database

Go to:

Basic Searching

`esearch.fcgi?db=<database>&term=<query>`

Input: Entrez database (&db); Any Entrez text query (&term)

Output: List of UIDs matching the Entrez query

Example: Get the PubMed IDs (PMIDs) for articles about breast cancer published in Science in 2008

[https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=science\[journal\]+AND+breast+cancer+AND+2008\[pdat\]](https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=science[journal]+AND+breast+cancer+AND+2008[pdat])

```
<eSearchResult>
  <Count>6</Count>
  <RetMax>6</RetMax>
  <RetStart>0</RetStart>
  <IdList>
    <Id>19008416</Id>
    <Id>18927361</Id>
    <Id>18787170</Id>
    <Id>18487186</Id>
    <Id>18239126</Id>
    <Id>18239125</Id>
  </IdList>
  <TranslationSet>
    <Translation>
      <From>science[journal]</From>
      <To>"Science"[Journal: __jid0404511] OR "Science (1979)"[Journal: __jid8000640]</To>
    </Translation>
    <Translation>
      <From>breast cancer</From>
      <To>"breast neoplasms"[MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields])
      OR "breast neoplasms"[All Fields] OR ("breast"[All Fields] AND "cancer"[All Fields]) OR
      "breast cancer"[All Fields]</To>
    </Translation>
  </TranslationSet>
  <QueryTranslation>("science"[Journal] OR "science 1979"[Journal]) AND ("breast neoplasms"
  [MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields]) OR "breast neoplasms"[All
  Fields] OR ("breast"[All Fields] AND "cancer"[All Fields]) OR "breast cancer"[All Fields])
  AND 2008/01/01:2008/12/31[Date - Publication]</QueryTranslation>
</eSearchResult>
```

List of result's PUBMED ID

Query term

API Pubmed search

`https://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi`
`?db=pubmed` Pubmed
`&term=emergency medicine+AND+AI` Search term
`&reldate=10` within the last *n* days
`&datetype= pdat` Date by publication date
`&retmax=2` Max *n* results

```
▼ <eSearchResult>
  <Count>14</Count>
  <RetMax>2</RetMax>
  <RetStart>0</RetStart>
  <QueryKey>1</QueryKey>
  <WebEnv>MCID_685258f0e012ab15ae0f711e</WebEnv>
  ▼ <IdList>
    <Id>40522351</Id>
    <Id>40517148</Id>
  </IdList>
  <TranslationSet/>
  <QueryTranslation>"emergency medicine"[All Fields] AND "AI"[All
  Fields] AND 2025/06/08:2025/06/18[Date - Entry]
  </QueryTranslation>
</eSearchResult>
```

List of result's PUBMED ID

Query term

<https://pubmed.ncbi.nlm.nih.gov/40522351/>

Review > J Med Syst. 2025 Jun 16;49(1):81. doi: 10.1007/s10916-025-02210-2.

A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting Emergency Department Revisits

Kuang-Ming Kuo ^{# 1}, Wen-Shiann Wu ^{# 2}, Chao Sheng Chang ^{3 4}

Affiliations + expand

PMID: 40522351 DOI: 10.1007/s10916-025-02210-2

Abstract

The revisit of the emergency department (ED) is a key indicator of emergency care quality. Various strategies have been proposed to reduce ED revisits, including the use of artificial intelligence (AI) models for prediction. However, AI model performance varies significantly, and its true predictive capability remains unclear. To address these gaps, the primary purpose of this study is to evaluate the performance of AI in predicting ED revisits through a meta-analysis. Specifically, this study aims to (1) Quantitatively assess the predictive performance of AI in ED revisit prediction and (2) Identify covariates contributing to between-study heterogeneity. A systematic search was conducted on December 31, 2024, across multiple electronic databases, including Scopus, SpringerLink, ScienceDirect, PubMed, Wiley, Sage, and Google Scholar, to identify relevant studies meeting the following criteria: (1) Utilized machine learning, deep learning, or artificial intelligence techniques to predict patient return visits to the ED, (2) Written in English, and (3) Peer-reviewed. Diagnostic accuracy was assessed using pooled sensitivity, specificity, and area under receiver operating characteristic curve (AUROC), while subgroup analysis explored factors contributing to heterogeneity. This meta-analysis included 20 articles, comprising 27 AI models. The summary estimates for ED revisit prediction were as follows: (1) Sensitivity: 0.56 (95% Confidence Interval [CI]: 0.44-0.67), (2) Specificity: 0.92 (95% CI: 0.86-0.96), and (3) AUROC: 0.81 (95% CI: 0.71-0.88). Subgroup analysis identified nationality, missing value-handling strategies, and specific disease samples as potential contributors to between-study heterogeneity. Future research should focus on improving missing value processing and using specific disease samples to enhance model reliability.

List of result's PUBMED ID

<Id>40522351</Id>

<Id>40517148</Id>

<https://pubmed.ncbi.nlm.nih.gov/40517148/>

> NPJ Digit Med. 2025 Jun 14;8(1):360. doi: 10.1038/s41746-025-01667-2.

A scoping review and evidence gap analysis of clinical AI fairness

Mingxuan Liu ^{# 1}, Yilin Ning ^{# 1}, Salinelat Teixayavong ¹, Xiaoxuan Liu ^{2 3 4}, Mayli Mertens ^{5 6}, Yuqing Shang ¹, Xin Li ¹, Di Miao ¹, Jingchi Liao ¹, Jie Xu ⁷, Daniel Shu Wei Ting ^{1 8}, Lionel Tim-Ee Cheng ⁹, Jasmine Chiat Ling Ong ¹⁰, Zhen Ling Teo ⁸, Ting Fang Tan ⁸, Narrendar RaviChandran ⁸, Fei Wang ¹¹, Leo Anthony Celi ^{12 13 14}, Marcus Eng Hock Ong ^{15 16}, Nan Liu ^{17 18 19}

Affiliations + expand

PMID: 40517148 PMCID: PMC12167363 DOI: 10.1038/s41746-025-01667-2

Abstract

The ethical integration of artificial intelligence (AI) in healthcare necessitates addressing fairness. AI fairness involves mitigating biases in AI and leveraging AI to promote equity. Despite advancements, significant disconnects persist between technical solutions and clinical applications. Through evidence gap analysis, this review systematically pinpoints the gaps at the intersection of healthcare contexts—including medical fields, healthcare datasets, and bias-relevant attributes (e.g., gender/sex)-and AI fairness techniques for bias detection, evaluation, and mitigation. We highlight the scarcity of AI fairness research in medical domains, the narrow focus on bias-relevant attributes, the dominance of group fairness centering on model performance equality, and the limited integration of clinician-in-the-loop to improve AI fairness. To bridge the gaps, we propose actionable strategies for future research to accelerate the development of AI fairness in healthcare, ultimately advancing equitable healthcare delivery.

© 2025. The Author(s).

List of result's PUBMED ID

<Id>40522351</Id>

<Id>40517148</Id>

A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting Emergency Department Revisits

The revisit of the emergency department (ED) is a key indicator of emergency care quality. Various strategies have been proposed to reduce ED revisits, including the use of artificial intelligence (AI) models for prediction. However, AI model performance varies significantly, and its true predictive capability remains unclear. To address these gaps, the primary purpose of this study is to evaluate the performance of AI in predicting ED revisits through a meta-analysis. Specifically, this study aims to (1) Quantitatively assess the predictive performance of AI in ED revisit prediction and (2) Identify covariates contributing to between-study heterogeneity. A systematic search was conducted on December 31, 2024, across multiple electronic databases, including Scopus, SpringerLink, ScienceDirect, PubMed, Wiley, Sage, and Google Scholar, to identify relevant studies meeting the following criteria: (1) Utilized machine learning, deep learning, or artificial intelligence techniques to predict patient return visits to the ED, (2) Written in English, and (3) Peer-reviewed. Diagnostic accuracy was assessed using pooled sensitivity, specificity, and area under receiver operating characteristic curve (AUROC), while subgroup analysis explored factors contributing to heterogeneity. This meta-analysis included 20 articles, comprising 27 AI models. The summary estimates for ED revisit prediction were as follows: (1) Sensitivity: 0.56 (95% Confidence Interval [CI]: 0.44-0.67), (2) Specificity: 0.92 (95% CI: 0.86-0.96), and (3) AUROC: 0.81 (95% CI: 0.71-0.88). Subgroup analysis identified nationality, missing value-handling strategies, and specific disease samples as potential contributors to between-study heterogeneity. Future research should focus on improving missing value processing and using specific disease samples to enhance model reliability.

Prompt

You are a medical researcher who summarizes new medical research papers daily in an understandable way. Write a short ,concludes the following papers.

Only paper in real world production

Input

A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting Emergency Department Revisits

The revisit of the emergency department (ED) is a key indicator of emergency care quality. Various strategies have been proposed to reduce ED revisits, including the use of artificial intelligence (AI) models for prediction. However, AI model performance varies significantly, and its true predictive capability remains unclear. To address these gaps, the primary purpose of this study is to evaluate the performance of AI in predicting ED revisits through a meta-analysis. Specifically, this study aims to (1) Quantitatively assess the predictive performance of AI in ED revisit prediction and (2) Identify covariates contributing to between-study heterogeneity. A systematic search was conducted on December 31, 2024, across multiple electronic databases, including Scopus, SpringerLink, ScienceDirect, PubMed, Wiley, Sage, and Google Scholar, to identify relevant studies meeting the following criteria: (1) Utilized machine learning, deep learning, or artificial intelligence techniques to predict patient return visits to the ED, (2) Written in English, and (3) Peer-reviewed. Diagnostic accuracy was assessed using pooled sensitivity, specificity, and area under receiver operating characteristic curve (AUROC), while subgroup analysis explored factors contributing to heterogeneity. This meta-analysis included 20 articles, comprising 27 AI models. The summary estimates for ED revisit prediction were as follows: (1) Sensitivity: 0.56 (95% Confidence Interval [CI]: 0.44-0.67), (2) Specificity: 0.92 (95% CI: 0.86-0.96), and (3) AUROC: 0.81 (95% CI: 0.71-0.88). Subgroup analysis identified nationality, missing value-handling strategies, and specific disease samples as potential contributors to between-study heterogeneity. Future research should focus on improving missing value processing and using specific disease samples to enhance model reliability.

A scoping review and evidence gap analysis of clinical AI fairness

The ethical integration of artificial intelligence (AI) in healthcare necessitates addressing fairness. AI fairness involves mitigating biases in AI and leveraging AI to promote equity. Despite advancements, significant disconnects persist between technical solutions and clinical applications. Through evidence gap analysis, this review systematically pinpoints the gaps at the intersection of healthcare contexts-including medical fields, healthcare datasets, and bias-relevant attributes (e.g., gender/sex)-and AI fairness techniques for bias detection, evaluation, and mitigation. We highlight the scarcity of AI fairness research in medical domains, the narrow focus on bias-relevant attributes, the dominance of group fairness centering on model performance equality, and the limited integration of clinician-in-the-loop to improve AI fairness. To bridge the gaps, we propose actionable strategies for future research to accelerate the development of AI fairness in healthcare, ultimately advancing equitable healthcare delivery.

Prompt

You're a podcast maker who summarizes new medical research papers daily in an understandable way. Write a short, text-to-speech-ready podcast script that concludes the following papers, without any names or side notes.

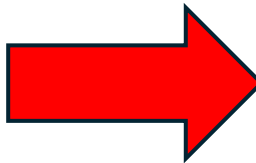
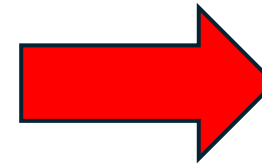
Input

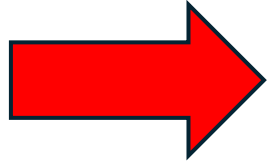
A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting Emergency Department Revisits

The revisit of the emergency department (ED) is a key indicator of emergency care quality. Various strategies have been proposed to reduce ED revisits, including the use of artificial intelligence (AI) models for prediction. However, AI model performance varies significantly, and its true predictive capability remains unclear. To address these gaps, the primary purpose of this study is to evaluate the performance of AI in predicting ED revisits through a meta-analysis. Specifically, this study aims to (1) Quantitatively assess the predictive performance of AI in ED revisit prediction and (2) Identify covariates contributing to between-study heterogeneity. A systematic search was conducted on December 31, 2024, across multiple electronic databases, including Scopus, SpringerLink, ScienceDirect, PubMed, Wiley, Sage, and Google Scholar, to identify relevant studies meeting the following criteria: (1) Utilized machine learning, deep learning, or artificial intelligence techniques to predict patient return visits to the ED, (2) Written in English, and (3) Peer-reviewed. Diagnostic accuracy was assessed using pooled sensitivity, specificity, and area under receiver operating characteristic curve (AUROC), while subgroup analysis explored factors contributing to heterogeneity. This meta-analysis included 20 articles, comprising 27 AI models. The summary estimates for ED revisit prediction were as follows: (1) Sensitivity: 0.56 (95% Confidence Interval [CI]: 0.44-0.67), (2) Specificity: 0.92 (95% CI: 0.86-0.96), and (3) AUROC: 0.81 (95% CI: 0.71-0.88). Subgroup analysis identified nationality, missing value-handling strategies, and specific disease samples as potential contributors to between-study heterogeneity. Future research should focus on improving missing value processing and using specific disease samples to enhance model reliability.

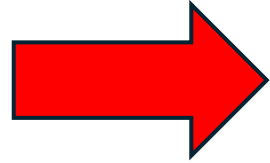
A scoping review and evidence gap analysis of clinical AI fairness

The ethical integration of artificial intelligence (AI) in healthcare necessitates addressing fairness. AI fairness involves mitigating biases in AI and leveraging AI to promote equity. Despite advancements, significant disconnects persist between technical solutions and clinical applications. Through evidence gap analysis, this review systematically pinpoints the gaps at the intersection of healthcare contexts-including medical fields, healthcare datasets, and bias-relevant attributes (e.g., gender/sex)-and AI fairness techniques for bias detection, evaluation, and mitigation. We highlight the scarcity of AI fairness research in medical domains, the narrow focus on bias-relevant attributes, the dominance of group fairness centering on model performance equality, and the limited integration of clinician-in-the-loop to improve AI fairness. To bridge the gaps, we propose actionable strategies for future research to accelerate the development of AI fairness in healthcare, ultimately advancing equitable healthcare delivery.





 Gemini



Here are the conclusions for the new medical papers, selected for their novelty or real-world impact:

****1. A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting Emergency Department Revisits****

****Conclusion:**** This meta-analysis consolidates the predictive performance of AI models for emergency department revisits, showing a pooled sensitivity of 0.56, specificity of 0.92, and AUROC of 0.81 across 20 articles. The findings suggest AI's potential in this area but highlight the need for improved missing value processing and disease-specific models to enhance reliability. This synthesis provides valuable insights into the current state and future directions for AI's role in optimizing emergency care resource management.

****2. A scoping review and evidence gap analysis of clinical AI fairness****

****Conclusion:**** This scoping ...

Text style

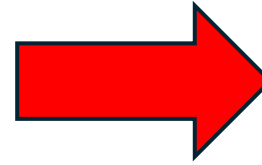
Here are the conclusions for the new medical papers, selected for their novelty or real-world impact:

****1. A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting Emergency Department Revisits****

****Conclusion:**** This meta-analysis consolidates the predictive performance of AI models for emergency department revisits, showing a pooled sensitivity of 0.56, specificity of 0.92, and AUROC of 0.81 across 20 articles. The findings suggest AI's potential in this area but highlight the need for improved missing value processing and disease-specific models to enhance reliability. This synthesis provides valuable insights into the current state and future directions for AI's role in optimizing emergency care resource management.

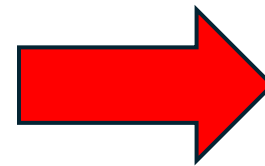
****2. A scoping review and evidence gap analysis of clinical AI fairness****

****Conclusion:**** This scoping ...



Podcast style

Welcome to your daily dose of medical research insights, where we break down the latest papers into understandable takeaways. Today, we're looking at two fascinating studies about artificial intelligence in healthcare. First, a meta-analysis explored how well AI can predict if patients will return to the emergency department. This is a crucial area for improving care quality. The analysis, which combined data from 20 articles and 27 different AI models, found that AI shows promising capability. Specifically, it could correctly identify about 56% of patients who would revisit, and accurately rule out revisits in 92% of cases. The overall predictive power was good, but researchers noted that factors.....

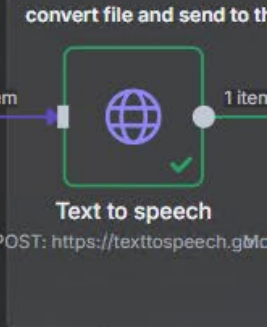
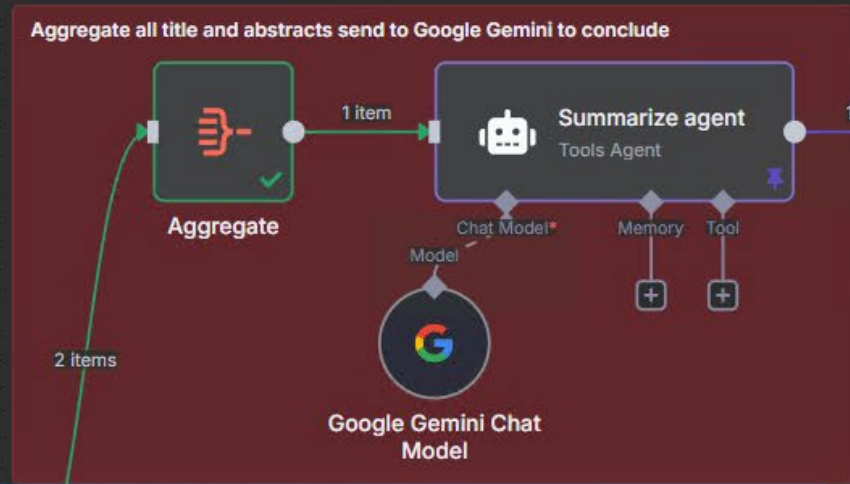
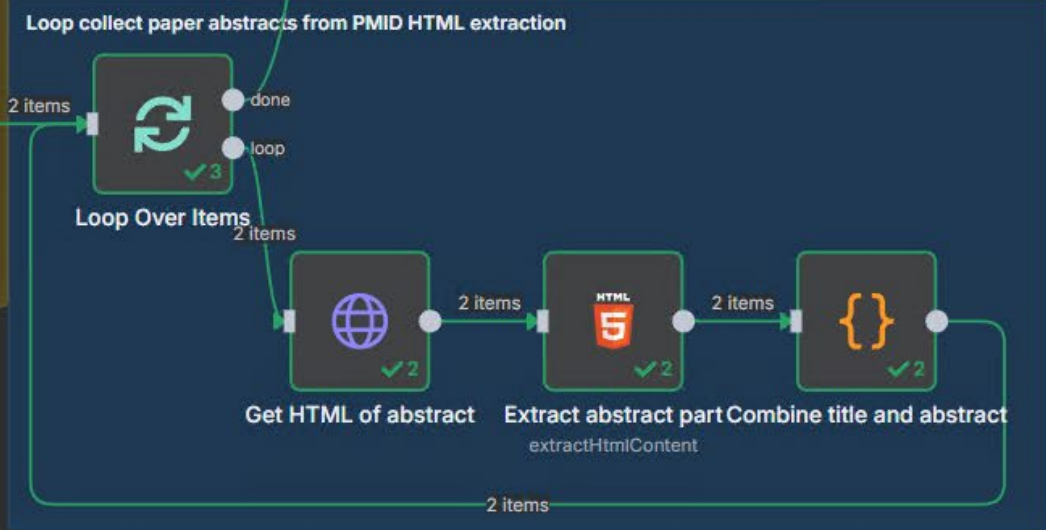
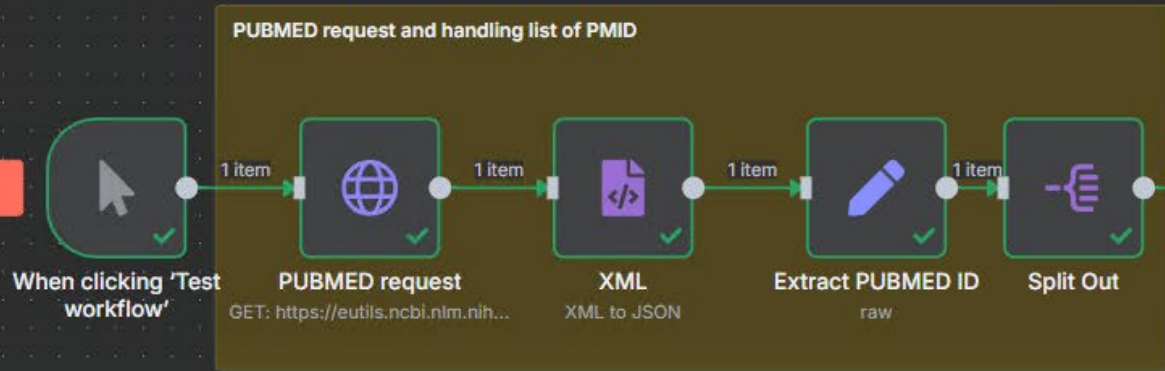


Google text-to-speech





n8n





<https://n8ndemoceb.a.pinggy.link/>

User : cholatid.rat@ceb-rama.org

Pass : [PROVIDED IN JC]

Why n8n ?

- It's free (self-host)
- Easy to use.
- Agentic workflow.
- Low-code to No-code.



n8n

<https://n8ndemoceb.a.pinggy.link/>



n8n

⚡ Starter

Perfect for getting started and seeing the power of n8n

~~€24~~ **€20** per month, billed annually

2.5k workflow executions with unlimited steps

5 active workflows and unlimited test ones

☁ Hosted by n8n

This plan includes:

- ✓ 1 shared project
- ✓ 5 concurrent executions

⚡ Pro

Optimal for a team collaborating to run more workflows

~~€60~~ **€50** per month, billed annually

10k workflow executions with unlimited steps

15 active workflows and unlimited test ones

☁ Hosted by n8n


Everything in Starter plan, plus:


- ✓ 3 shared projects
- ✓ 20 concurrent executions

⚡ Enterprise

Ideal for businesses with strict security and performance requirements

Contact us

up to  **workflow executions** with unlimited steps

up to  **active workflows** and unlimited test ones

📁 Self-hosted (or hosted by n8n)

Everything in Pro plan, plus:

- ✓ Unlimited shared projects
- ✓ 200+ concurrent executions
- ✓ 365 days of insights



n8n

<https://n8ndemoceb.a.pinggy.link/>

Generate, Get data

Trigger Node
(Start)



When clicking "Test Workflow"

Node
(Input > Function > Output)

Get fake sample data



Customer Datastore (1st)
getAllPeople

Extract data and prepare it for use in the next node



Edit Fields1
manual

1 Item

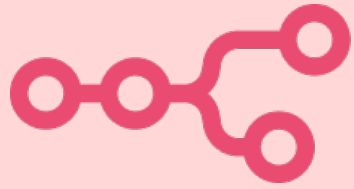
5 Items

Data Transfer
1/ Items Count
2/ Data Type

- Open
- Test step
- Rename
- Deactivate
- Pin
- Copy
- Duplicate
- Tidy up workflow
- Select all
- Clear selection
- Delete

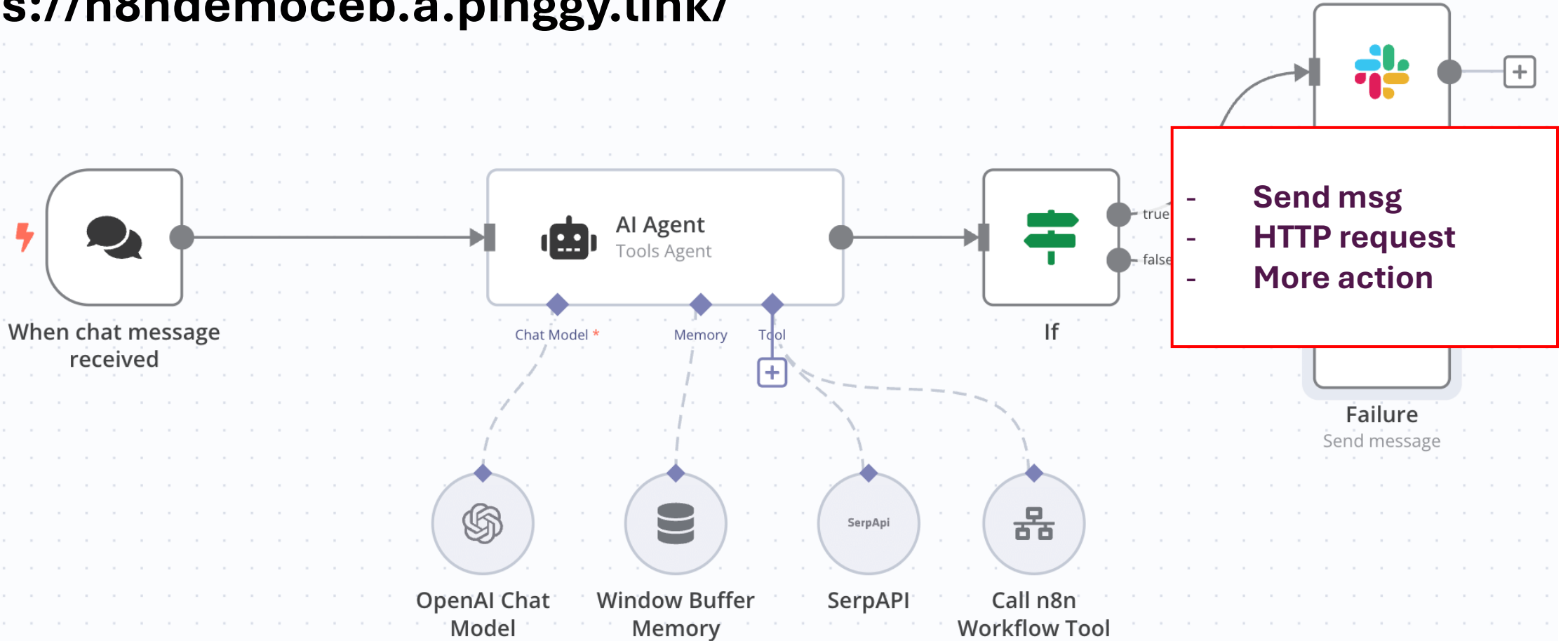
Chat, Time, Click

Extract, edit



n8n

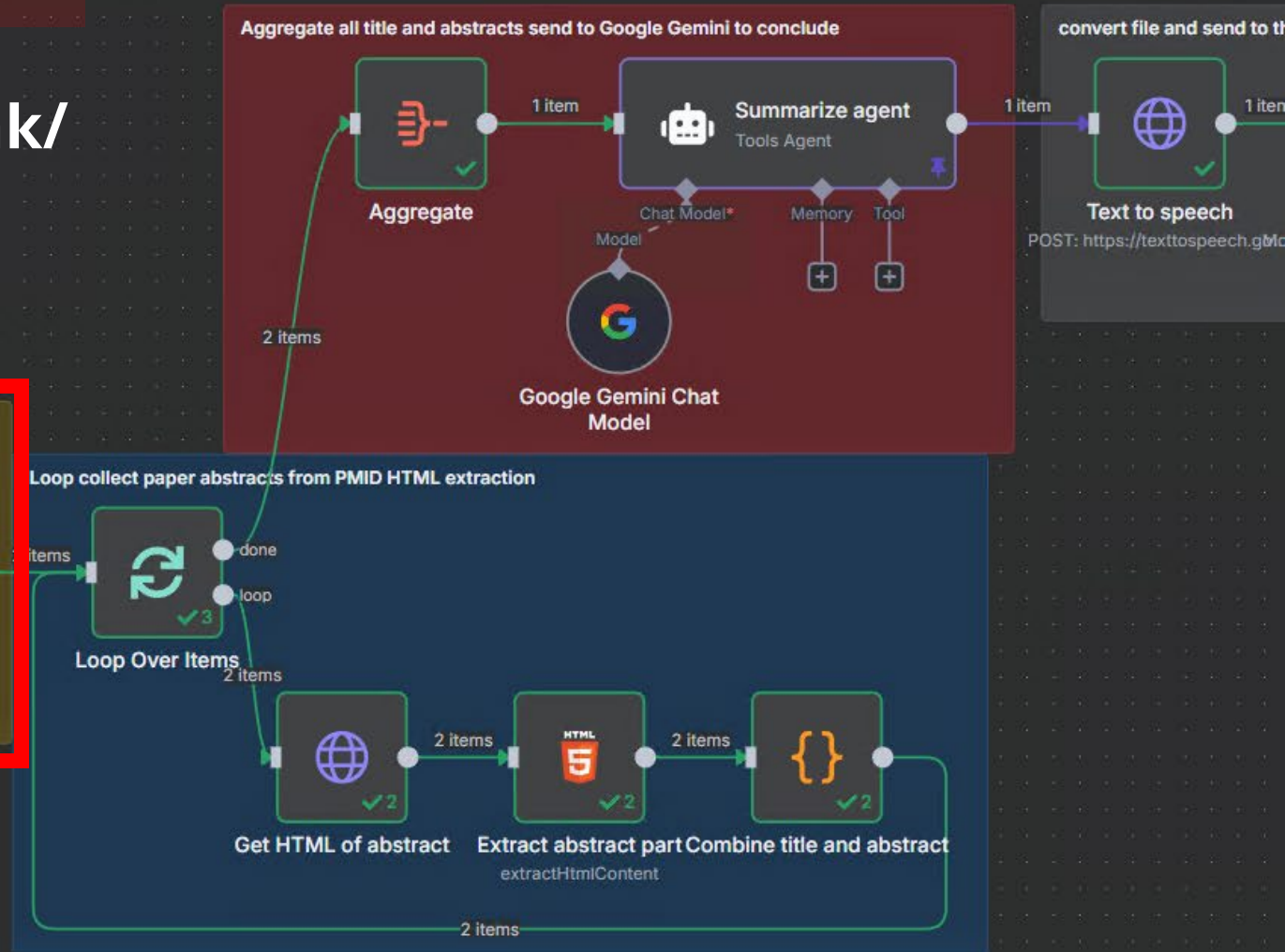
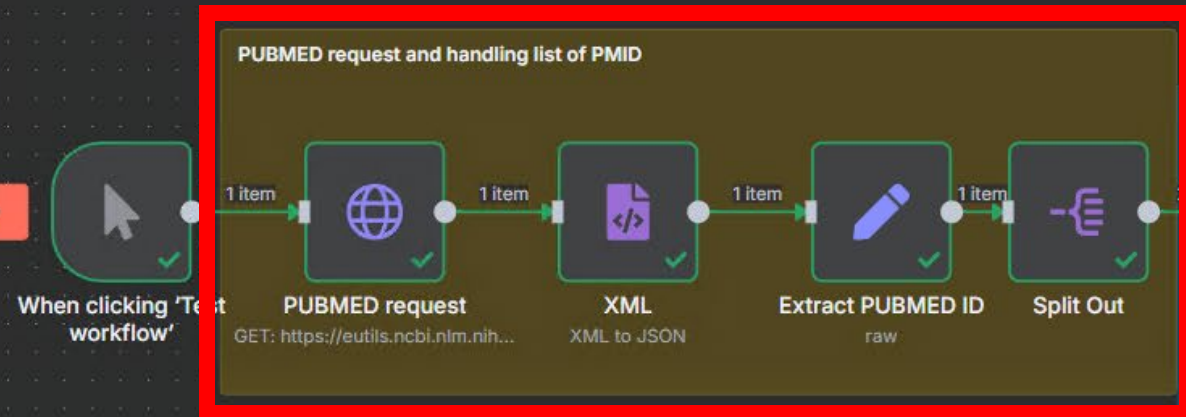
<https://n8ndemoceb.a.pinggy.link/>





n8n

<https://n8ndemoceb.a.pinggy.link/>



NODE1 : HTTP request

Method - GET

<https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=pubmed&term=%22emergency%20medicine%22+AND+%22AI%22&reldate=10&datetype=dat&retmax=2&usehistory=y>

NODE2 : XML to JSON

NODE3: Edit fields(set)

JSON - `{{ $json.eSearchResult.IdList}}`

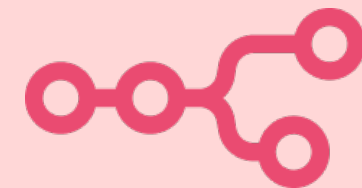
NODE 4: Split Out

NODE 5: Loop Over Items

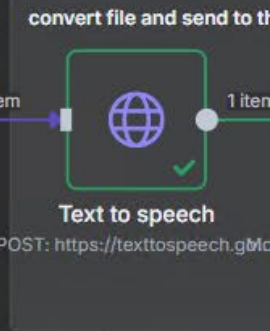
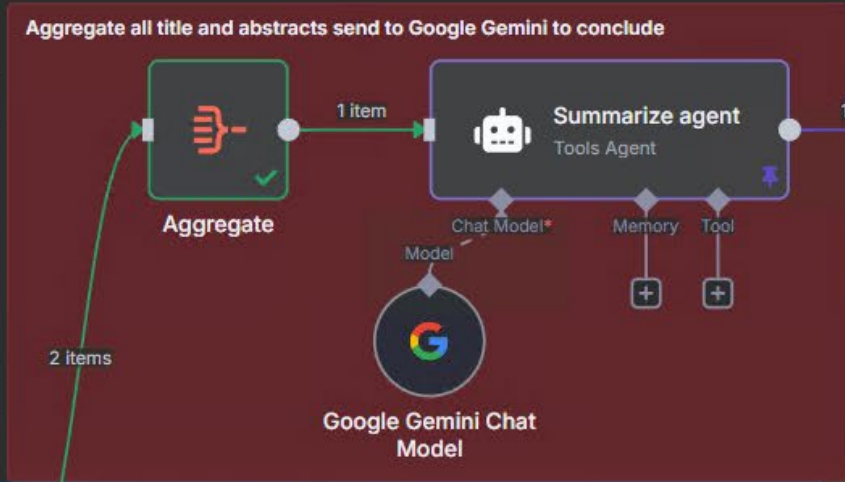
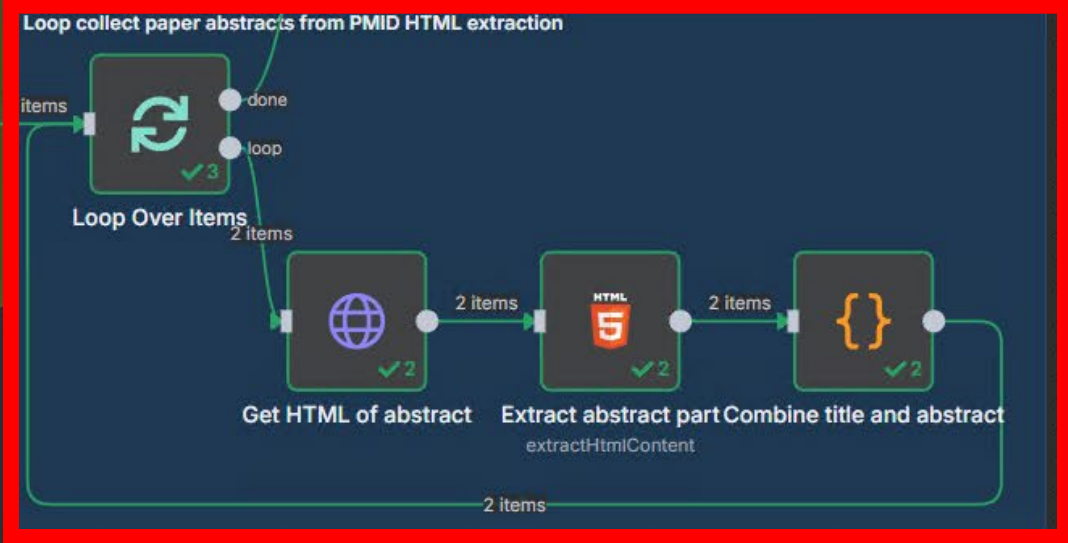
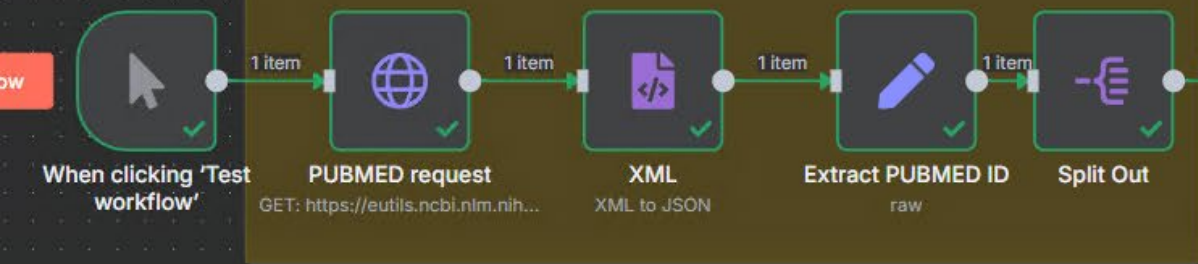
NODE 6: HTTP request


Method – GET

URL - `https://pubmed.ncbi.nlm.nih.gov/{{ $json.Id }}`



n8n



 **Extract abstract part**

Parameters Settings Docs [↗](#)

Operation
Extract HTML Content ▼

Source Data
JSON ▼

JSON Property
data

Extraction Values

Key
abstract-content selected

CSS Selector
div[class="abstract-content selected"]

Return Value
Text ▼

Skip Selectors
e.g. img, .className, #itemId

Return Array

Key
title

CSS Selector
h1[class="heading-title"]

Return Value
Text ▼

Extraction Values

Key : abstract-content selected
div[class="abstract-content selected"]

Key : Title
h1[class="heading-title"]

<https://pubmed.ncbi.nlm.nih.gov/40522351/>

Review > J Med Syst. 2025 Jun 16;49(1):81. doi: 10.1007/s10916-025-02210-2.

A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting Emergency Department Revisits

Kuang-Ming Kuo ^{# 1}, Wen-Shiann Wu ^{# 2}, Chao Sheng Chang ^{3 4}

Affiliations + expand

PMID: 40522351 DOI: 10.1007/s10916-025-02210-2

Abstract

The revisit of the emergency department (ED) is a key indicator of emergency care quality. Various strategies have been proposed to reduce ED revisits, including the use of artificial intelligence (AI) models for prediction. However, AI model performance varies significantly, and its true predictive capability remains unclear. To address these gaps, the primary purpose of this study is to evaluate the performance of AI in predicting ED revisits through a meta-analysis. Specifically, this study aims to (1) Quantitatively assess the predictive performance of AI in ED revisit prediction and (2) Identify covariates contributing to between-study heterogeneity. A systematic search was conducted on December 31, 2024, across multiple electronic databases, including Scopus, SpringerLink, ScienceDirect, PubMed, Wiley, Sage, and Google Scholar, to identify relevant studies meeting the following criteria: (1) Utilized machine learning, deep learning, or artificial intelligence techniques to predict patient return visits to the ED, (2) Written in English, and (3) Peer-reviewed. Diagnostic accuracy was assessed using pooled sensitivity, specificity, and area under receiver operating characteristic curve (AUROC), while subgroup analysis explored factors contributing to heterogeneity. This meta-analysis included 20 articles, comprising 27 AI models. The summary estimates for ED revisit prediction were as follows: (1) Sensitivity: 0.56 (95% Confidence Interval [CI]: 0.44-0.67), (2) Specificity: 0.92 (95% CI: 0.86-0.96), and (3) AUROC: 0.81 (95% CI: 0.71-0.88). Subgroup analysis identified nationality, missing value-handling strategies, and specific disease samples as potential contributors to between-study heterogeneity. Future research should focus on improving missing value processing and using specific disease samples to enhance model reliability.

List of result's PUBMED ID

<Id>40522351</Id>

<Id>40517148</Id>


```
▼ <h1 class="heading-title">
  " A Meta-Analysis of the Diagnostic Test Accuracy of Artificial Intelligence for Predicting
</h1>
▶ <div class="inline-authors">⋮</div>
▶ <div class="short-article-details">⋮</div>
▶ <div class="extended-article-details" id="expanded-authors" aria-hidden="true">⋮</div>
▶ <ul class="identifiers" id="full-view-identifiers">⋮</ul>
  <div class="in-clipboard-label " hidden> Item in Clipboard </div>
</div>
▶ <div class="short-view" id="short-view-heading">⋮</div>
</header>
```

Extraction Values


Key: abstract-content selected -
div[class="abstract-content selected"]

Key: Title - h1[class="heading-title"]

```
▶ <div class="actions-buttons inline">⋮</div>
▼ <div class="abstract" id="abstract">
  <h2 class="title"> Abstract </h2>
▼ <div class="abstract-content selected" id="eng-abstract">
  ▼ <p>
    " The revisit of the emergency department (ED) is a key indicator of emergency care quality.
    including the use of artificial intelligence (AI) models for prediction. However, AI model p
    capability remains unclear. To address these gaps, the primary purpose of this study is to e
    a meta-analysis. Specifically, this study aims to (1) Quantitatively assess the predictive p
    covariates contributing to between-study heterogeneity. A systematic search was conducted or
    including Scopus, SpringerLink, ScienceDirect, PubMed, Wiley, Sage, and Google Scholar, to i
    Utilized machine learning, deep learning, or artificial intelligence techniques to predict p
    Peer-reviewed. Diagnostic accuracy was assessed using pooled sensitivity, specificity, and a
    while subgroup analysis explored factors contributing to heterogeneity. This meta-analysis i
    estimates for ED revisit prediction were as follows: (1) Sensitivity: 0.56 (95% Confidence I
    0.96), and (3) AUROC: 0.81 (95% CI: 0.71-0.88). Subgroup analysis identified nationality, mi
    as potential contributors to between-study heterogeneity. Future research should focus on in
    samples to enhance model reliability. "
  </p>
</div>
```

- **NODE 8 : Code – Python COMBINE TITLE AND ABSTRACT**

```
_input.item.json["text"] = _input.item.json["title"].replace("\n","") + "\n\n"+  
_input.item.json["abstract-content selected"].replace("\n","")  
return _input.item.json
```

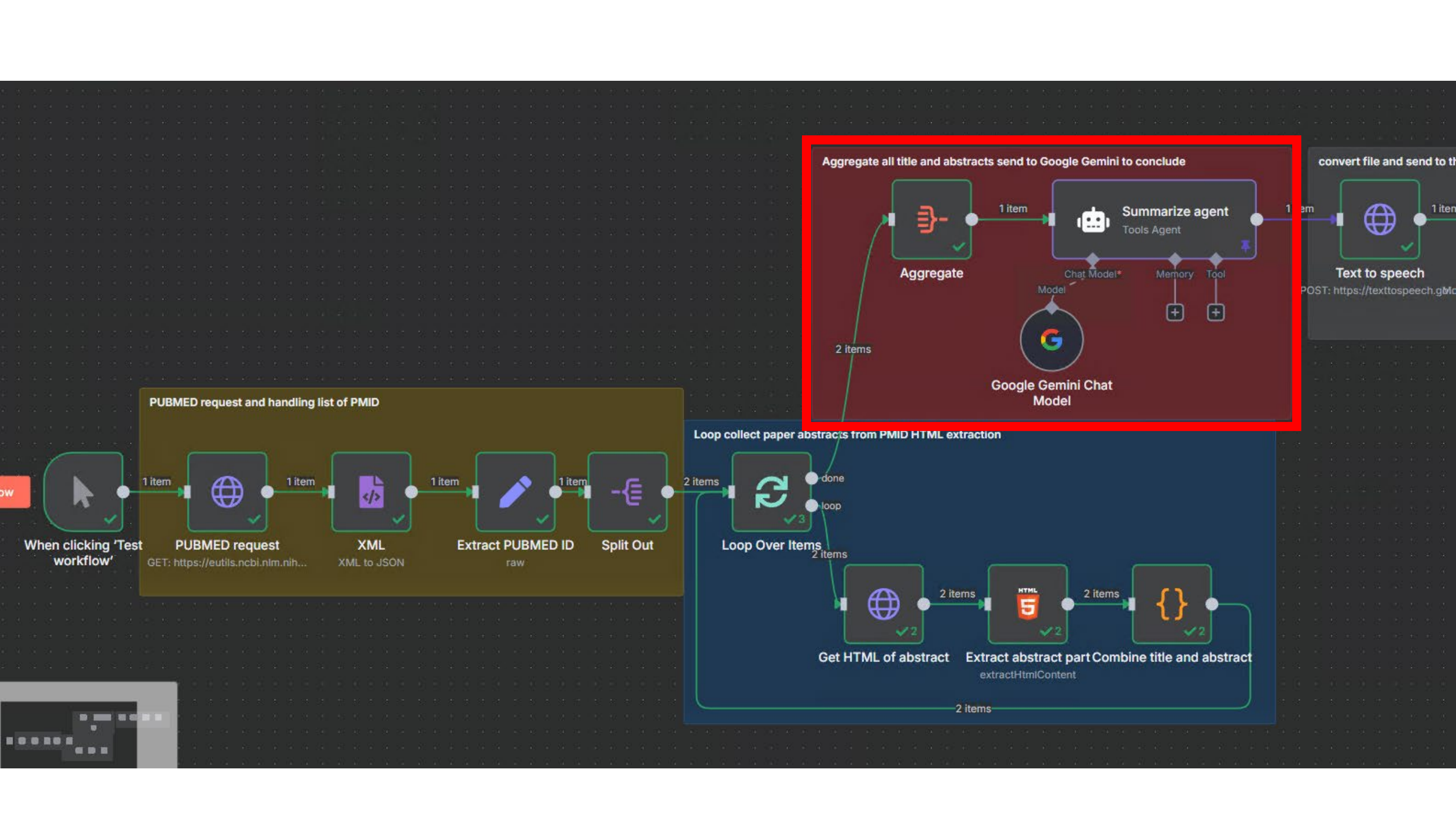


The screenshot shows a workflow editor interface for a Python node. The node is titled "Combine title and abstract" and has a "Test step" button. The "Parameters" tab is selected, showing the following settings:

- Mode:** Run Once for Each Item
- Language:** Python (Beta)

The Python code editor contains the following code:

```
1 # Add a new field called 'myNewField' to the JSON of the item  
2 _input.item.json["text"] = _input.item.json["title"].replace("\n","") + "\n\n"+  
  _input.item.json["abstract-content selected"].replace("\n","")  
3 return _input.item.json
```



NODE 9 : Aggregate

Aggregate Test step

Parameters Settings Docs

Aggregate

Individual Fields

Fields To Aggregate

Input Field Name

text

Enter the field name as text

Rename Field

Add Field To Aggregate

Options

Merge Lists

Add Field

NODE 10 : AI Agent

Summarize agent Test step

Parameters Settings Docs

Tip: Get a feel for agents with our quick [tutorial](#) or see an [example](#) of how this node works

Source for Prompt (User Message)

Define below

Prompt (User Message)

`fx {{ $json.text }}`

[Array: ["A Meta-Analysis of the Diagnostic Test Accuracy ...

Require Specific Output Format

Options

System Message

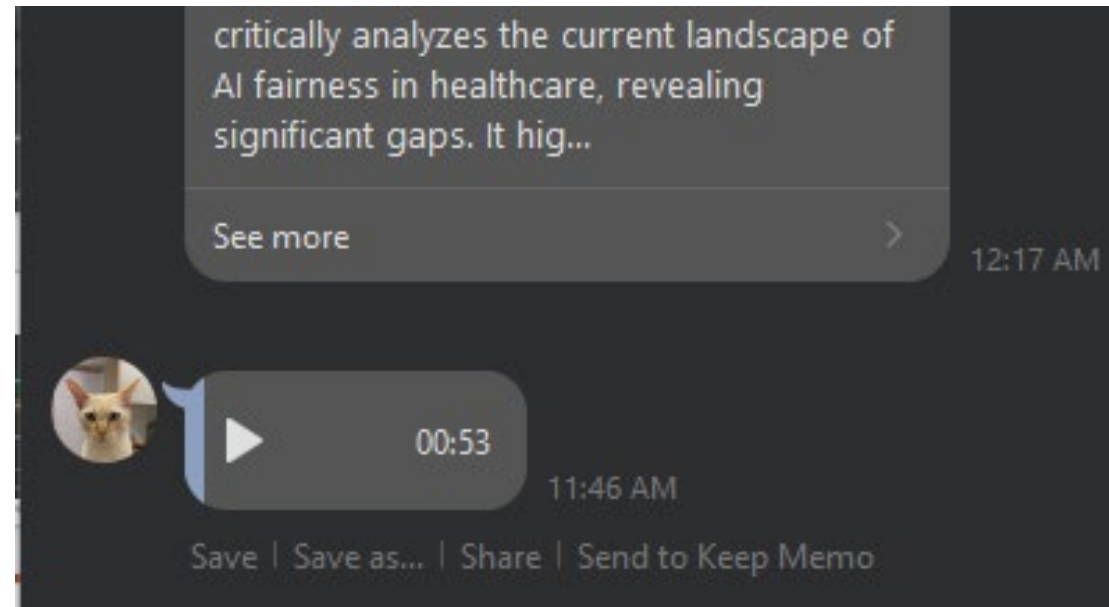
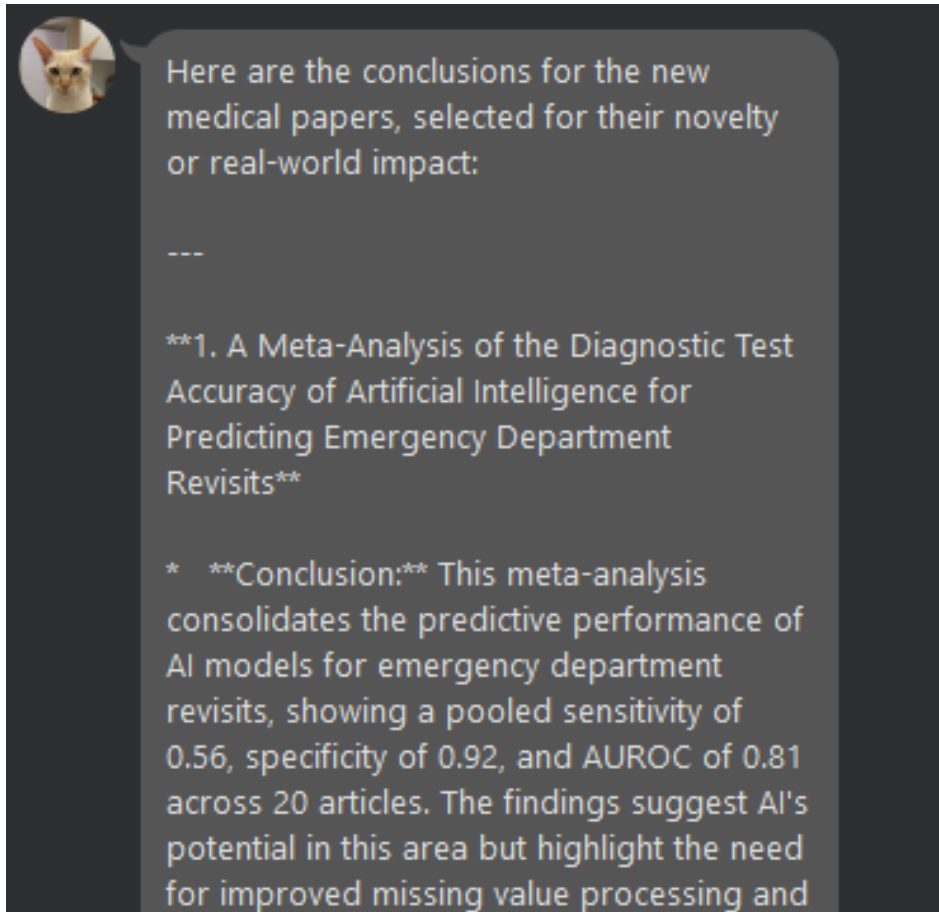
You are the podcast maker about daily medical research who conclude new medical paper every day in understandable way. write the script to be a short podcast to conclude all paper below. this script is ready to text-to-speech. should not

Add Option

The image shows a configuration window for a Google Gemini Chat Model. The window has a title bar with the Google logo and the text "Google Gemini Chat Model". Below the title bar are two tabs: "Parameters" (selected) and "Settings", along with a "Docs" link. The main content area is divided into sections: "Credential to connect with" with a dropdown menu showing "Google Gemini(PaLM) Api account"; "Model" with a dropdown menu showing "models/gemini-2.5-flash-preview-05-20"; and "Options" with the text "No properties" and an "Add Option" button. A green banner at the top right of the main area says "Connection tested successfully" with a checkmark icon. Below this, a yellow banner says "Need help filling out these fields". The bottom section contains a "Host" field with the URL "https://generativelanguage.googleapis.com" and an "API Key" field with a masked key represented by dots. A footer note states: "Enterprise plan users can pull in credentials from external vaults. More info".

API Key : AlzaSyBOK8Tj46GDAQxxxxxxxxxxxxc

Send to the designated application



Integrate LINE with N8N



- Create LINE official account and msg API.

<https://developers.line.biz/en/docs/messaging-api/getting-started/#create-oa-entry-form>

- Get LINE channel secret token.

<https://docs.klink.cloud/getting-started/setup-channels/line-oa/get-channel-secret-and-channel-access-token>

- Integrate with n8n

<https://n8gen.com/workflow/mastering-line-messaging-api-automation-with-n8n-instant-replies-and-push-messages>

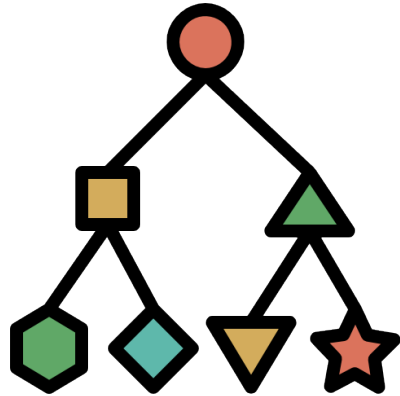
Case II

Problem :

Emergency medical services call taker workload is overwhelming. They cannot triage every call, and we have high call drop rate.

Proposed solution :

Agentic AI and agentic workflow might have potential as a concept. Triage and history taking are big challenges in this agentic workflow.



Symptom classifier Agent

- Detect symptom
- Classify symptom
- Patient/bystander



History taking and Triage Agent

- Turn based history taking
- Guideline retrieval
- Red flag (critical) detection
- History summarization
- Triage

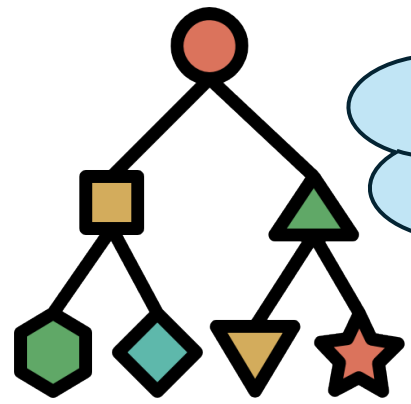


Making appointment Agent

- Find available slot in the google calendar
- Request understanding
- Book in the google calendar

I have headache.

Select one of list of symptoms in guideline :
[Abdominal pain, Headache, Rash ,...]



Symptom classifier Agent

Symptom : Headache
Who's talking with : Patient



History ta
Triaging

Symptom : Headache
Who's talking with : Patient

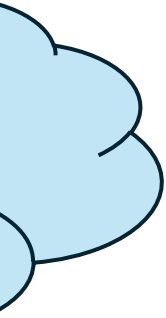


**History taking and
Triage Agent**

Reasoning : Patient have headache
I should ask about ...
Next question : How long have you
had headache?
Is red flag : False
Is finish : False

Retrieve triage guideline : Headache

```
"12. Headache/Neck Pain": {  
  "guide_number": 12,  
  "criteria": {  
    "Red (Critical)": [  
      "12 Red 1: The patient is unconscious or not breathing. ",  
      "12 Red 2: The patient has at least one sign of abnormal breathing. ",  
      "12 Red 3: The patient has at least two signs of shock. ",  
      "12 Red 4: The patient has a decreased level of consciousness, is uncooperative, or  
      "12 Red 5: The patient has a sudden, severe headache combined with slurred speech  
    ],  
    "Yellow (Urgent)": [  
      "12 Yellow 1: The patient is disoriented but can still talk and walk. ",  
      "12 Yellow 2: The patient has shortness of breath. ",  
      "12 Yellow 4: The patient has difficulty seeing. ",  
      "12 Yellow 5: The patient is dizzy. ",  
      "12 Yellow 9: The caller cannot confirm the patient's detailed symptoms. "    ]  
  }  
}
```



History taking and
Triaging Agent

RAG FLAG : True

History taking
complete

Stop history taking and tell patient
to call an ambulance

Give a triage code
If Critical, tell patient to call an ambulance.
If not critical, send to make an appointment

: Headache

```
not breathing. ",
sign of abnormal breathing. ",
signs of shock. ",
level of consciousness, is uncooperative,
are headache combined with slurred speech

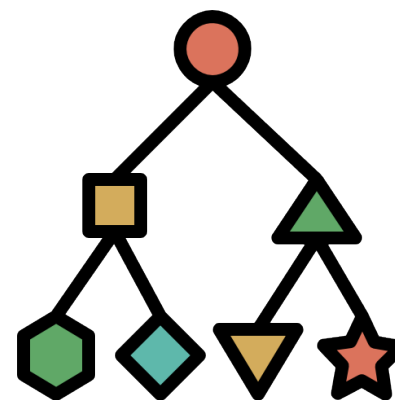
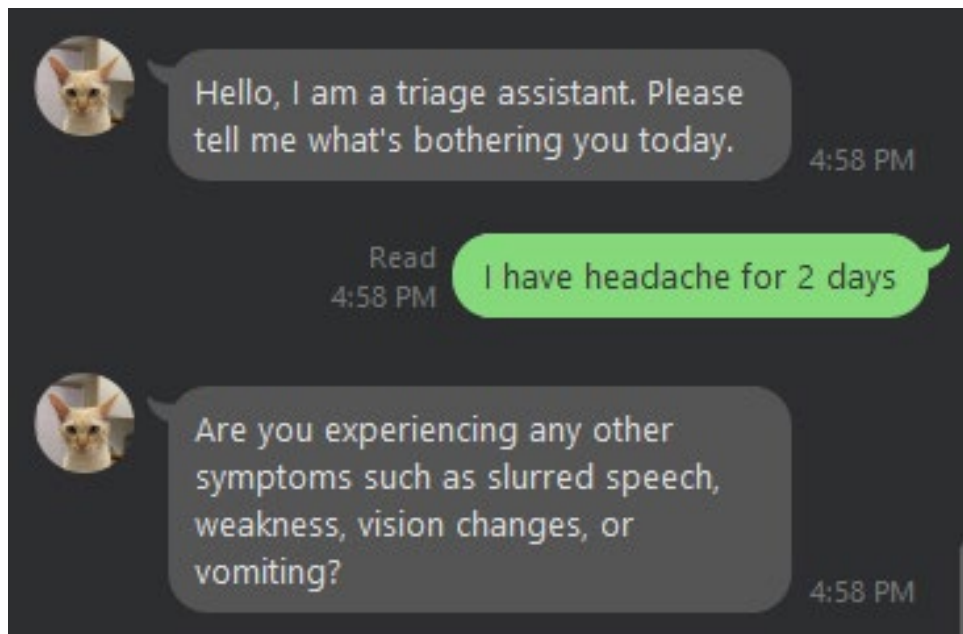
but can still talk and walk. ",
of breath. ",
seeing. ",

the patient's detailed symptoms. "
```



**Making appointment
Agent**

- Find available slot in the google calendar
- Request understanding e.g., tomorrow morning
- Book in the google calendar

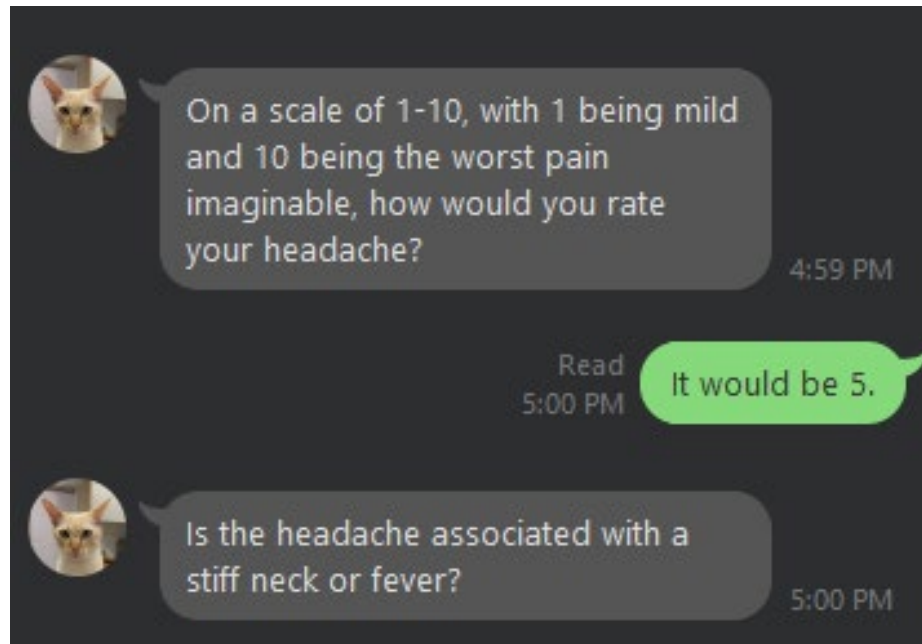


**Symptom classifier
Agent**



**History taking and
Triage Agent**

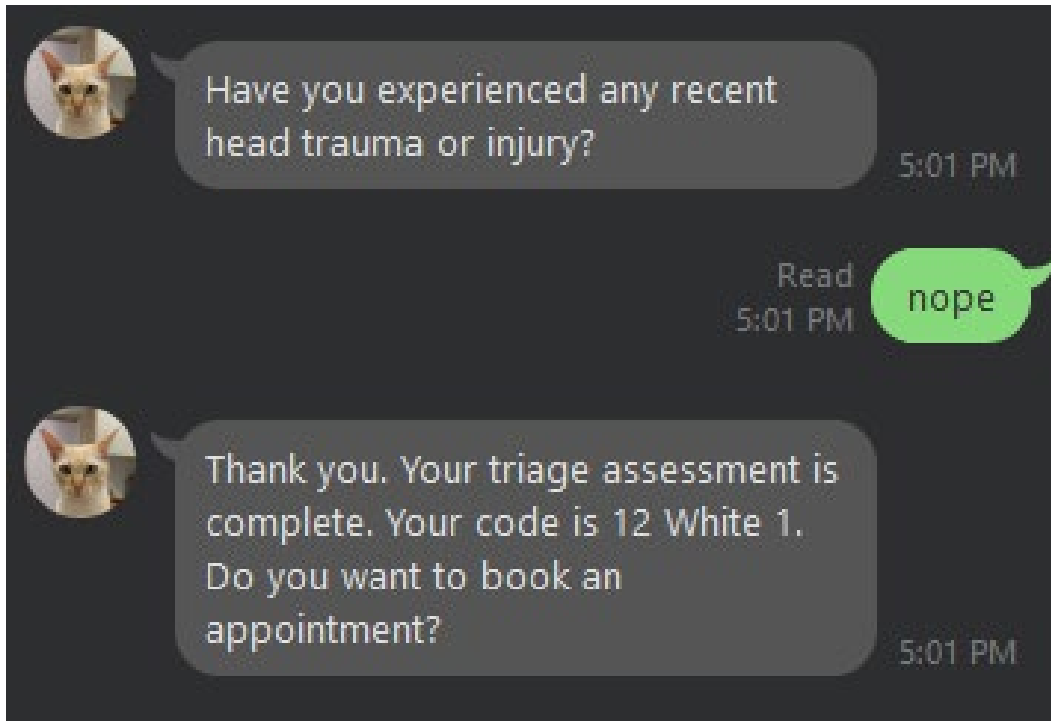
```
[MSG] First clinical message. Complaint: I have headache for 2 days
[Symptom] Complaint classified as: Headache/Neck Pain
[TOOL USE] GuidelineRetriever searching for: 'Headache/Neck Pain'
[TOOL USE] GuidelineRetriever found guide for 'Headache/Neck Pain'.
[REASONING] LLM Reasoning:
The patient reports a headache for two days. I need to gather more information to determine the severity of the head
ache and rule out any red flag symptoms. I will ask about associated symptoms to assess for red flags.
c
--> AI Response to U52de707b8f4a793ba554cad85d59be2a: Are you experiencing any other symptoms such as slurred speech
, weakness, vision changes, or vomiting?
```

**History taking and
Triage Agent**

[REASONING] LLM Reasoning:
The patient reports a headache for two days. The patient denies slurred speech, weakness, vision changes, or vomiting. Based on the guidelines, I need to rule out red flag symptoms. Since the patient denies the red flag symptoms, I will ask about the severity of the headache.
c
--> AI Response to U52de707b8f4a793ba554cad85d59be2a: On a scale of 1-10, with 1 being mild and 10 being the worst pain imaginable, how would you rate your headache?

[REASONING] LLM Reasoning:
The patient reports a headache for two days, rated as a 5/10. The patient denies any red flag symptoms like slurred speech, weakness, vision changes, or vomiting. The guidelines do not specify any need for an image. I will continue to gather information to rule out red flags and determine the appropriate triage code.
c
--> AI Response to U52de707b8f4a793ba554cad85d59be2a: Is the headache associated with a stiff neck or fever?



History taking and Triage Agent

[REASONING] LLM Reasoning:

The patient reports a headache for two days, rated as a 5/10. They deny associated symptoms like slurred speech, weakness, vision changes, vomiting, stiff neck, fever, or recent head trauma. Based on the guidelines, this presentation does not meet any red flag criteria. The information points towards a general headache.

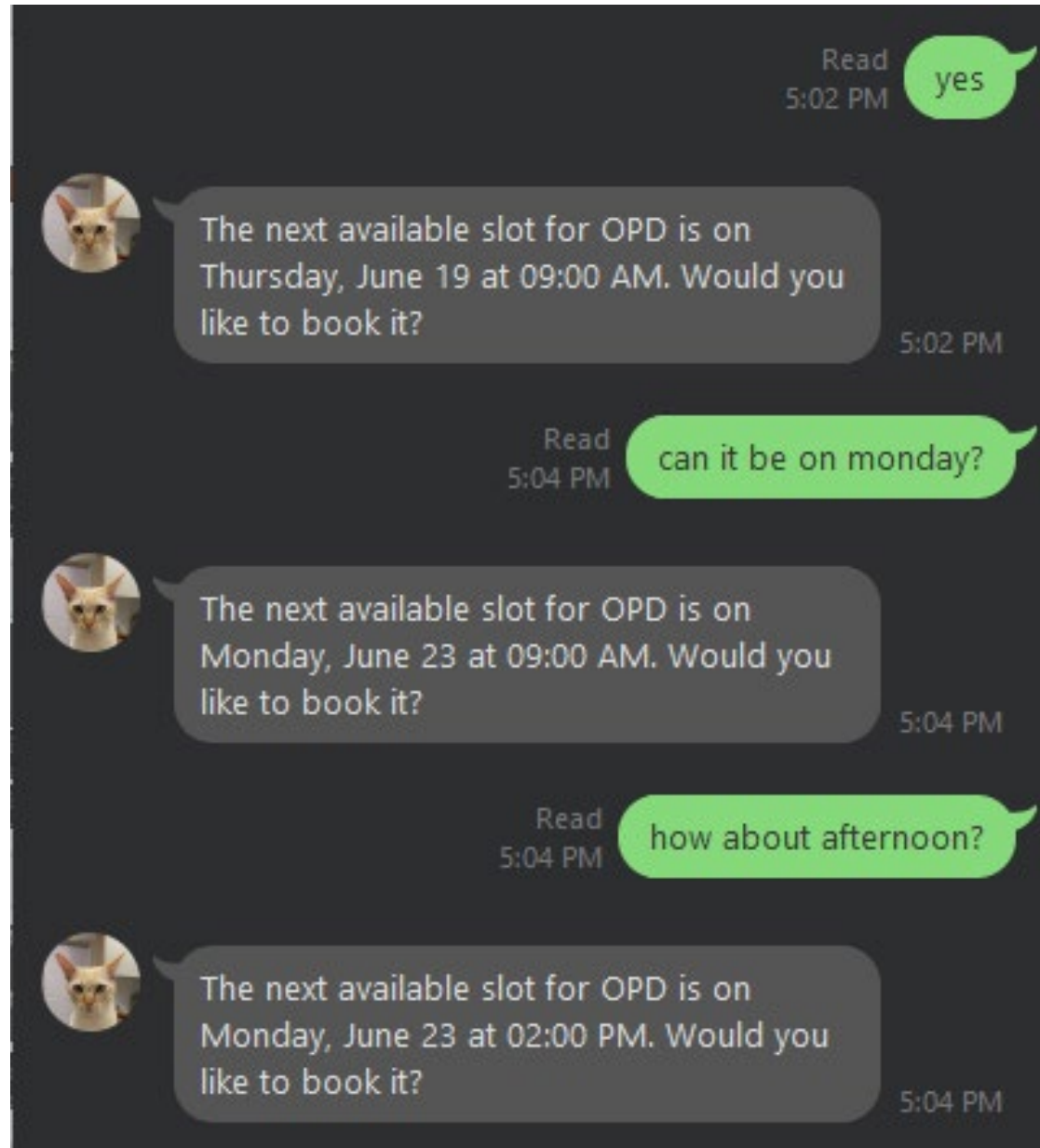
a

--> AI Response to U52de707b8f4a793ba554cad85d59be2a: Thank you. Your triage assessment is complete. Your code is 12 White 1. Do you want to book an appointment?

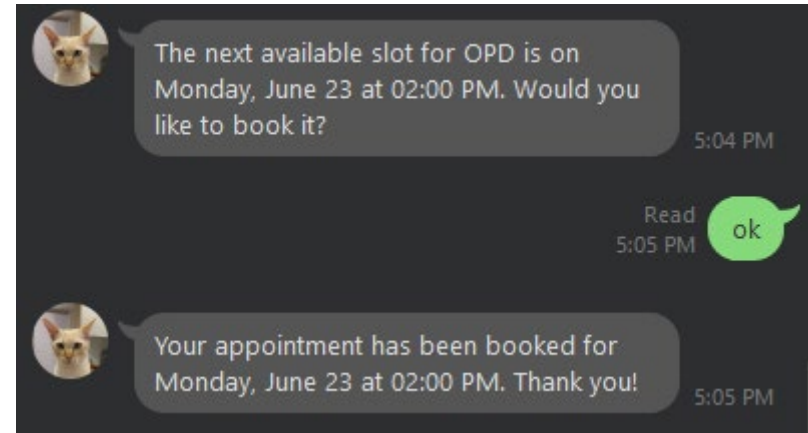
Found free slot!
--> AI Response to U52de707b8f4a793ba554cad85d59
be2a: The next available slot for OPD is on Thursday,
June 19 at 09:00 AM. Would you like to book it?



**Making appointment
Agent**



```
-- Enter event create --
2025-06-23T14:00:00+07:00
[PROGRAM AUTOMATION] Creating Google Calendar ev
ent...
--> AI Response to U52de707b8f4a793ba554cad85d59
be2a: Your appointment has been booked for Monda
y, June 23 at 02:00 PM. Thank you!
```



A screenshot of a mobile medical application interface. The main screen shows a calendar with a highlighted date '23' and a 'สำนักงาน' (Office) label. A side panel is open, displaying appointment details for 'OPD:Headache/Neck Pain(Triage code: 12 White 1)' on June 23, 2:00 - 3:00pm. The side panel includes a document icon for 'เพิ่มบันทึกการประชุม' (Add meeting notes), a 'Patient History' section with text 'Patient reports a headache for 2 days, rated 5/10, with no other concerning symptoms.', and a 'Hosp' section with 'สร้างโดย: Cholatid Ratanatharathorn'.



Making appointment Agent

AGENTIC

AI



AI

AGENT

AI Agents vs. Agentic AI: A Conceptual Taxonomy, Applications and Challenges

Cholatid Ratanatharathorn

Department of Clinical Epidemiology and Biostatistics