International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

# Healthsphere: A Proposed Web-Based Framework using Machine Learning and Generative Ai for Personalized Health Information and Wellbeing Support

**Pranav Prashanth Hegde[1], Nara Siddharth[2], Kusuma N[3]**

[1] Student, Department of BCA, St. Francis College, Koramangala, Bengaluru.

[2] Student, Department of BCA, St. Francis College, Koramangala, Bengaluru.

[3] Assistant Professor, Department of BCA, St. Francis College, Koramangala, Bengaluru.

**Abstract -** Information that pertains to health issues may be expressed in terms of specific terminology which is difficult for common people to understand. This paper describes HealthSphere, a web-based information supporting system which has the following characteristics: (i) Extraction and presentation in plain English of chosen values from a blood test report, (ii) fuzzy matching technique for finding a drug name from the reference set which consists of more than 10,531 items as described in the project, and (iii) a chatbot based on natural language processing (NLP) for managing common health issues and stress. The system architecture involves input validation, modular processing, reference matching, safety check of output, and outputs themselves. Fuzzy similarity measure is used to rank the candidate items for drug search, but not to find whether those drugs are equivalent. The chatbot is expected to provide supporting information and not to make any diagnosis or statements related to emergency treatment. This is an intended-for-use-in-real-world paper; no details about its accuracy, clinical importance, or user feedback are provided. Module-wise evaluation scheme is outlined, consisting of quality of extraction, detection of abnormalities, drug search, safety of the answers, usability and privacy.

**Keywords**—HealthSphere, proposed system, machine learning, generative AI, blood-report interpretation, fuzzy string matching, medicine-name recognition, natural language processing, digital wellbeing, Precision, Recall, F1-Score, F-measure, true negative rate, true positive rate.

## I. INTRODUCTION

Although there is plenty of information about health on the Internet, people still tend to face difficulty in understanding laboratory reports and other medical documents. Laboratory reports include lots of abbreviations, results of measurements, units, and reference ranges for laboratories. It may be impossible to find the medicine, if there is any minor mistake in entering its name to the search field. People under stress may want to get information on well-being.

HealthSphere concept is a web application which would complete all the three tasks stated above at once. Blood report interpreter module would need to interpret the data in supported fields in a plain language. The medicine recognizer module would need to rank the medicines based on some reference list. The well-being chatbot would make conversation easy. These are features related to gaining information and not to treatment or diagnostics.

The most important contributions of the paper include the proposed architecture, the flow within each component, the aspects concerning the safety and privacy, and the reproducible evaluation procedure. As the experiments are not performed, there are no statements about the performance of the paper.

### A. Objectives

1. Construct an input routine for handling blood-test inputs, where the test name, result, units, and reference intervals can be extracted from the input.

2. Compare the results against the reference intervals available in the report, while flagging the results that lie outside the reference intervals without diagnosing.



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

3. Apply fuzzy string matching algorithm to find matches for medicine names and verifying the matches before recording the data.

4. Develop a wellness chatbot for managing stress.

## II. LITERATURE SURVEY

### A. RELATED WORK

| Sl. No. | Author(s) and Year | Research Title / Source | Objective of the Study | Methodology / Approach | Major Findings | Limitations / Research Gap | Relevance to HealthSphere |
|---|---|---|---|---|---|---|---|
| 1 | WHO, 2021 [1] | Ethics and Governance of Artificial Intelligence for Health | To identify ethical challenges in healthcare AI. | Reviewed AI ethics, governance, privacy, and accountability. | Emphasized transparency, safety, and privacy. | Provides general guidance, not a specific implementation. | Supports privacy, transparency, and responsible AI use. |
| 2 | Ayers et al., 2023 [2] | Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions | To compare AI and physician responses. | Compared patient responses evaluated by healthcare professionals. | AI responses received higher quality and empathy ratings in the study. | Does not establish consistent clinical accuracy. | Supports clear AI-generated health explanations with appropriate safeguards. |
| 3 | MedlinePlus [3] | How to Understand Your Lab Results | To explain laboratory results and reference ranges. | Provides educational guidance on interpreting test results. | Results must be understood in their clinical context. | Does not automatically interpret uploaded reports. | Guides HealthSphere's blood-report interpretation module. |
| 4 | Li et al., 2023 [5] | AI-based Conversational Agents for Mental Health and Well-being | To evaluate AI conversational agents for mental health. | Systematic review and meta-analysis of clinical trials. | Found reductions in depression symptoms and distress in evaluated studies. | Long-term effectiveness and safety need further study. | Supports the wellbeing chatbot while highlighting the need for evaluation. |
| 5 | WHO, 2023 [6] | Regulatory Considerations on Artificial Intelligence for Health | To identify regulatory requirements for healthcare AI. | Reviewed risk assessment, performance, and monitoring. | Emphasized evaluation and responsible AI deployment. | Provides general rather than application-specific guidance. | Guides HealthSphere's testing, safety, and performance evaluation. |
| 6 | Integrated Analysis | Combined AI-based Health Support | To identify the need for integrated health-support | Synthesized findings from the above sources. | Existing research supports individual AI | HealthSphere's integrated performance remains to be | Establishes the rationale for combining report interpretation, medicine |



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | services. | | healthcare applications. | evaluated. | recognition, and wellbeing support. |

**Research Gap**

All these sources provide a foundation for certain individual functions which can be implemented together in order to develop a web-based application for HealthSphere. There are guidelines for ethical design, privacy, and accountability regarding the use of AI in healthcare [1], [6]. Also, there are works about AI chatbots in connection with their ability to respond to questions about health [2]. There are also educational materials about laboratory tests which describe the necessity to interpret the results of the tests in accordance with the correct reference range and with regard to the individuality of the patient [3]. Edit distance algorithms provide a technical base for approximate matching of medicine names [4]. Review papers on research studies suggest that conversational agents may help with some but not all well-being outcomes [5].

However, none of these sources provides any information about the performance of the application that uses all three aforementioned functions. The suggestion of HealthSphere is a single web-based application for these purposes.

## B. TECHNICAL BACKGROUND

### A. Digital health communication and laboratory results

Digital health solutions may help in providing information; nevertheless, clarity and security of the provided information are equally vital. Laboratory results have to be interpreted, since the units of the measurement, reference range, performed laboratory test, and patient's medical history all affect the result. Therefore, the suggested interface should display not only the result itself but also the range in which it is presented and any simplification of the result.

### B. Machine learning and generative AI

The former technology can assist in recognizing patterns and performing classification; the latter may be able to create explanations based on structured input data in a comprehensible way. Nevertheless, the explanations may be erroneous and incomplete as well. In this regard, HealthSphere recommends implementing a more restricted approach in which the explanations are created based on the extracted data and reference documents [2]–[4].

### C. Fuzzy string matching for medicine names

The exact match query can suffer from the mistakes in spelling, partial match, and typos. In contrast, fuzzy matching allows comparing strings based on a degree of similarity, e.g., the edit distance, and ranking candidates based on that measure. Ranking process assists in finding what is needed; nevertheless, similar strings may point to some other ingredient, strength, or preparation of the substance. Recognition system proposed by me will provide an ability to present candidate strings for validation along with the source [5].

### D. NLP-based wellbeing support

An NLP chatbot will be able to comprehend the language of the user and reply appropriately providing the necessary information to resolve the issue. Nonetheless, this cannot guarantee the level of competency of the person. The proposed chatbot will not be diagnosing or recommending treatment but only assisting, which implies that there should be a protocol for extremely anxious messages.

### E. Research motivation

The development of HealthSphere is oriented toward the display of diverse tasks to be executed using the same interface. However, their integration alone does not guarantee that the applications are correct, useful, and safe because all these qualities should be assessed independently. Thus, in this particular case, the evaluation will relate to HealthSphere as a system design.

## III. PROPOSED SYSTEM

### A. Proposed system architecture

The design consists of four different layers, which include the user interface layer, the validation and routing layer, three different functional layers, and the safety and output layer. The three functional layers use their own logic in executing their tasks.



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

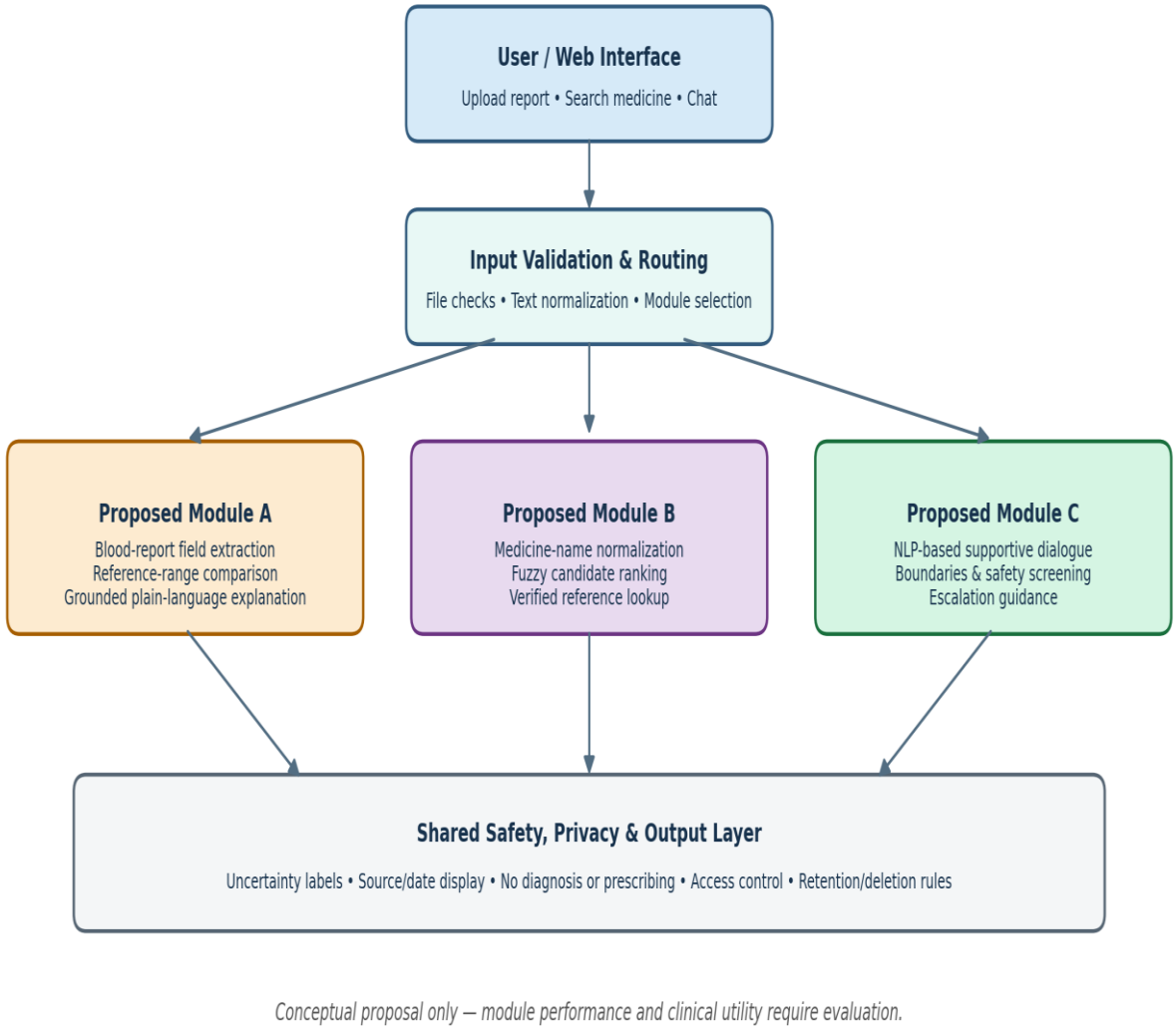



*Fig. 1. Proposed HealthSphere system architecture (conceptual design; not an experimental result).*

This can be seen from Fig. 1 where it is evident that the software sends every verified input to the right module. The output layer, which is common for both modules, will take care of uncertainty, sources, constraints, and future actions. This is how it should be done and does not mean that all controls have been implemented yet.

**B. Proposed end-to-end workflow**

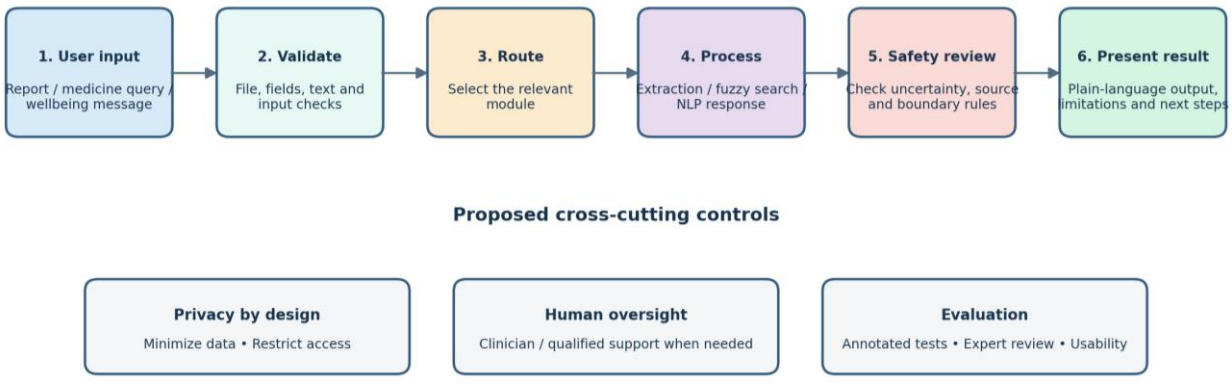



*Fig. 2. Proposed workflow and cross-cutting controls.*



(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

The work flow starts from the point where data entry and validation are carried out by the users. This request is sent to the module that will deal with the request either by extracting it, performing fuzzy look up, or having a conversation. The request passes through a lot of safety checks before any output can be generated. Where the data provided is not satisfactory and there is any ambiguity on the request, more input is needed.

### C. Proposed Module 1: Blood-report interpretation

A report with supporting documentation is provided by the user. The processing module proposed detects names of tests, measurement of value, units and reference intervals from the report. There will be validation for any missing fields, ambiguous characters, and error in parsing. If there is any reference interval found in the report, then the module compares the value with the reference interval and classifies it as below, within and above the range specified. The generative-AI module will provide a description for the test in few words based on fields identified and references available.

The value, units and reference interval must be available to the users. There should not be any replacement of the reference interval with a universal interval. Any poor quality scans and unusual format along with ambiguous extraction should trigger a warning or manual verification process. An out-of-range result cannot be used as a diagnosis.

### D. Proposed Module 2: A–Z medicine-name recognizer

The project description states that the medicine reference library would have more than 10,531 references, and in the final implementation, the source, license, scope, update rate, and the format of the individual entries should be specified. There will be a recommendation about the fact that the search algorithm would convert all cases, spacing, and some special symbols of the input query to the lowercase prior to the comparison of the names of the possible references. The fuzzy similarity algorithm will assess the possible references, after which the user will choose the needed medicine.

If the valid reference already exists, the application will provide indications, dosages, adverse effects, contraindications, source, and the date of validation of the medicine. They shall not be generated if they do not exist in the reference library. The similarity rates will be used for the ranking of the candidate medicines, but not to validate them as well as to give recommendations about initiation, discontinuation, and modification of medicines.

### E. Proposed Module 3: Wellbeing chatbot

The suggested chatbot will process the natural language input from the users and return the positive responses which do not include any medical advice about how to cope with the common stress. They may offer some useful tips on coping with the problems and reflecting upon them, but cannot claim anything concerning any medicines. It is necessary to be very careful with any safety messages and encourage the human assistance. It should be noted that the chatbot is not an emergency contact tool.

### F. Proposed integration, privacy, and output design

Both modules will have the navigation and the appearance parts, but will differ in their processing parts. The output needs to be delivered in the clear language, and the facts and explanations should be distinguished. Also, the issue of uncertainty needs to be addressed. The messages and reports about wellbeing may contain personal information, and, therefore, the proposed design of the outputs will involve encrypted data delivery, access control, data minimization, retention and deletion policy, and lack of sensitive data in logs.

## IV. PROPOSED IMPLEMENTATION

The architecture of the system will consist of the following blocks: (1) interface for uploading and queries; (2) analyzing and extracting fields from reports; (3) comparing the information with reference intervals; (4) generating a grounded explanation; (5) normalizing mentions of medicines and fuzzy search; (6) searching for references about medicines; (7) NLP-based conversation about well-being; and (8) verification, safety, privacy and fault-tolerance features. It is crucial to mention the used programming languages, libraries, models' version, cloud and data in the documentation of implementation.

### A. Proposed input and output handling

For each request it should be provided the schema of input data for the particular module. In case of request for reports, it should include the extracted name of the test, value, unit, time interval, extraction confidence (if there is any) and the source record's id. In case of requests to search for medicines, it should contain the original query, the list of candidates ordered by similarity, the matching algorithm and the source record id. In case of the request for a chatbot, it should include the restrictions of input, safety check and policy of response.


(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

### B. Proposed data governance

The project team should ensure that there is an inventory of the provenance, license, date of updates, and limitations of the medication database. All report samples that have been developed or tested should be anonymized as much as possible and processed accordingly. Before conducting tests on users, access permission, storage duration, deletion policy, and incident management should be defined.

## V. PROPOSED EVALUATION METHODOLOGY

Testing of the application has to be done on a per-module basis before usability tests can be conducted. The dataset, criteria for selection, annotation, software version, and testing conditions have to be recorded. Reporting of results should only be done after conducting experiments; the following are only proposed measures.

### A. Blood-report module

Create manually annotated data for all possible report formats. The precision, recall and F1 scores must be computed for the field extraction task. The below/within/above range category must be compared to the interval stated in the report. Reviewers expert in this task must ascertain whether the explanation provided for the extracted value is correct.

### B. Medicine-name recognizer

Create a test set containing exact names, spelling errors, part of string, variations in spacing, and confusing similarities. Measure top-1 and top-k retrieval rates, as well as mean reciprocal rank. Count false matches, particularly where similar medicines correspond to different products. Ensure that the displayed medicine information is consistent with a reliable and dated source. Performance testing must be stratified by query type.

### C. Wellbeing chatbot

The empirical assessment needs to be carried out based on the following criteria: empathy, relevance, clarity, unsubstantiated claims, boundaries, privacy, and escalations. The human evaluators will be expected to assess normal stress situations as well as the safety-critical situations. User studies require that informed consent is sought; proper safeguards and distress procedures are put in place.

### D. Integrated usability evaluation

Make the participants do typical tasks including uploading the report, understanding the explanation, finding the medicinal record and working with the wellbeing chatbot. Collect the performance information for the tasks done, the time taken and the errors committed. Include participant profile information, sample size and limitations in your description. Usability test should never be conducted in order to prove clinical efficacy.

### E. Proposed evaluation metrics

| Component | Proposed measures |
|---|---|
| **Field extraction** | Precision, Recall, F1-Score of Test Names, Values, Units, and Intervals |
| **Range flagging** | Sensitivity, Specificity, and Agreement of Report-Specific Intervals<br>Medicine Retrieval |
| **Medicine retrieval** | Precision, recall, F-measure, true negative rate, true positive rate |
| **Explanation quality** | Factual consistency, uncertainty management, unvalidated claims ratio |
| **Chatbot safety** | Rubric score, unsafe-response rate, success of escalation in test situations |
| **Usability** | Task completion rate, completion time, task clarity |


(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 11, CONFLUENCE - 2026

## VI. DISCUSSION

Possible benefits of the HealthSphere concerning design are convenience in using the platform because all three of its features such as report explanation, medicine name recognition, and well-being guidance can be accessed from one interface. Fuzzy matching may help to avoid negative consequences of typos, and grounding of the information in simple language would assist in explaining the terminology used. Output layer consistency will enable consistent warnings and privacy notifications.

All of the above are possible benefits. No labelled datasets or performance measures, expert opinions or results of usability studies have been provided in order to prove these benefits. Hence, no performance improvements have been reported in this paper numerically. The reliability of the platform depends on the quality of extraction, reference data, generated texts, disambiguation of the medicine names, and sensitive messages.

Some of the key limitations include poor quality of report scans, laboratory-specific timings, non-updated medicine information, similarity conflicts among names and wrong interpretations generated. In addition, a chatbot might interpret the message incorrectly and fail to detect any real distress. Here are some of the proposed mitigation strategies: uncertainties, source traceability, safe default options, peer reviews, safety testing and separation of information systems from practice.

## VII. ETHICAL, SAFETY, AND PRIVACY CONSIDERATIONS

HealthSphere is to be presented to the user community as a prototype of information till such time it is properly validated and governed. It should also be made clear to users that HealthSphere does not diagnose the disease or treat the same. In addition, it will not recommend any change to the patient's medications. Abnormal test results need to be explained carefully to the users.

Minimal data should be recorded in HealthSphere and they should be secure during transfer and storage. Permissions should be clearly set. Policies for retention and deletion of data should be made and should not record any sensitive information without need. Real testing of the system should be performed with proper permissions. Safety policies should be reviewed regularly by experts.

## VIII. CONCLUSION AND FUTURE WORK

In this paper, a web based system, called HealthSphere, is proposed that would provide an integration of blood reports' analysis, medicine identification with fuzzy logic techniques, and wellness assistance using NLP. This paper has discussed design of the architecture of the system, workflow of the modules, implementation details, evaluation methodology, and issues regarding safety and privacy.

Future work on this topic would be to document how the system was actually implemented and from where did the data come. The extraction of report information would have to be evaluated in various formats. The evaluation of medicine retrieval should be done in hard negative cases and explanation generated by the system should be evaluated by some experts.

## REFERENCES

[1] K. L. Ayers et al., “Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum,” JAMA Internal Medicine, vol. 183, no. 6, pp. 589–596, 2023. https://doi.org/10.1001/jamainternmed.2023.1838
[2] U.S. National Library of Medicine, “Understanding Laboratory Test Results,” MedlinePlus. https://medlineplus.gov/lab-tests/understanding-laboratory-test-results/
[3] H. Li et al., “Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being,” npj Digital Medicine, vol. 6, article 236, 2023. https://doi.org/10.1038/s41746-023-00979-5
[4] World Health Organization, Regulatory Considerations on Artificial Intelligence for Health. Geneva: WHO, 2023.
[5] World Health Organization, Global Strategy on Digital Health 2020–2025. Geneva, Switzerland: WHO, 2021.
[6] U.S. National Library of Medicine, “Understanding Laboratory Test Results,” MedlinePlus. Available: https://medlineplus.gov/lab-tests/understanding-laboratory-test-results/. Accessed: Oct. 9, 2026.
[7] U.S. National Library of Medicine, “DailyMed.” Available: https://dailymed.nlm.nih.gov/. Accessed: Oct. 9, 2026.
[8] World Health Organization, Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva: WHO, 2021.
[9] K. L. Ayers et al., “Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum,” JAMA Internal Medicine, vol. 183, no. 6, pp. 589–596, 2023, doi: 10.1001/jamainternmed.2023.1838.
[10] World Health Organization, Regulatory Considerations on Artificial Intelligence for Health. Geneva, Switzerland: WHO, 2023.


(This work is licensed under a Creative Commons Attribution 4.0 International License.)