**Challenge 3:** Optimising record viewing and easing entry of symptoms

**Objective:** Patients visit different doctors/ hospitals as per their need and convenience. Each facility has different ways of capturing patient visit information – EMR/ physical paper format. Even in case of digital information capture, sharing information and viewing in a format that is usable by physicians is important. Leveraging on OCR technology and Health Information Exchange concept, helps physicians view health information from disparate sources (including physical copies) and enables better clinical decision making

Providing user friendly methods for patients/ providers to capture symptoms experienced/ diagnosis will enable better adoption of technology amongst end users. Furthermore, having a structured way of capturing this information will enable the doctors to understand and interpret the information better and enable more efficient data sharing across systems. Technology innovations such as auto-completion/ auto-suggestion reduce time required for documentation and increase efficiency, resulting is better patient care.

**Benefits:**

* Ability to capture and convert information from disparate systems in disparate forms to structured data that can be consumed by health care professionals
* Help physicians with Clinical Decision making
* Enable better adoption of technology with user friendly features

**Features:**

* ***Implemented Features***
  + Health Information Exchange Application:
    - Ability to share structured information (allergies, medication, lab test) from across different digital systems (Cerner Sandbox – Open Source & JSON file from a sample facility)
    - Integration using FHIR protocol
    - Ability to check slot/schedules for the day
  + OCR technology: To enable reading information from pictures (PNG/ JPEG) post conversion from PDF and converting into structured data leveraging the TESSERACT library
  + Combining data structured using OCR and medication search with the text with data from secondary source (JSON file) to present aggregated structured information
  + Auto-suggestion
  + Diagnosis based on standard medical nomenclature such as ICD 11 and SNOMED CT
  + ICD 11 diagnosis is achieved over API call. However, as the key expires every hour due to security, it needs to be updated regularly. Also a new API was developed with all the ICD 11 codes to be used as Auto-suggestion for diagnosis documentation.
  + With the implementation of SNOMED CT which is a concept nomenclature we have been able to achieve the following: For example, if the provider wants to document for heart attack, he/ she can type heart attack and all relevant medical terminologies such as Myocardial infarction along with the code will get listed
  + Ability to upload documentation and photo captures from the phone to patient record
  + Application in local Indian language (Hindi)
  + Android (For mobility and to cater to smaller clinics) and Web based applications developed
* **Conversion of semi-structured data into structured data:**
  + This conversion uses an Open Source Text Recognition Engine called Tesseract which recognizes the machine readable text from an external source detector such as handwritten images, hosted imageURL’s and text Images.
  + It uses the Convolutional Neural Network for the image Recognition.

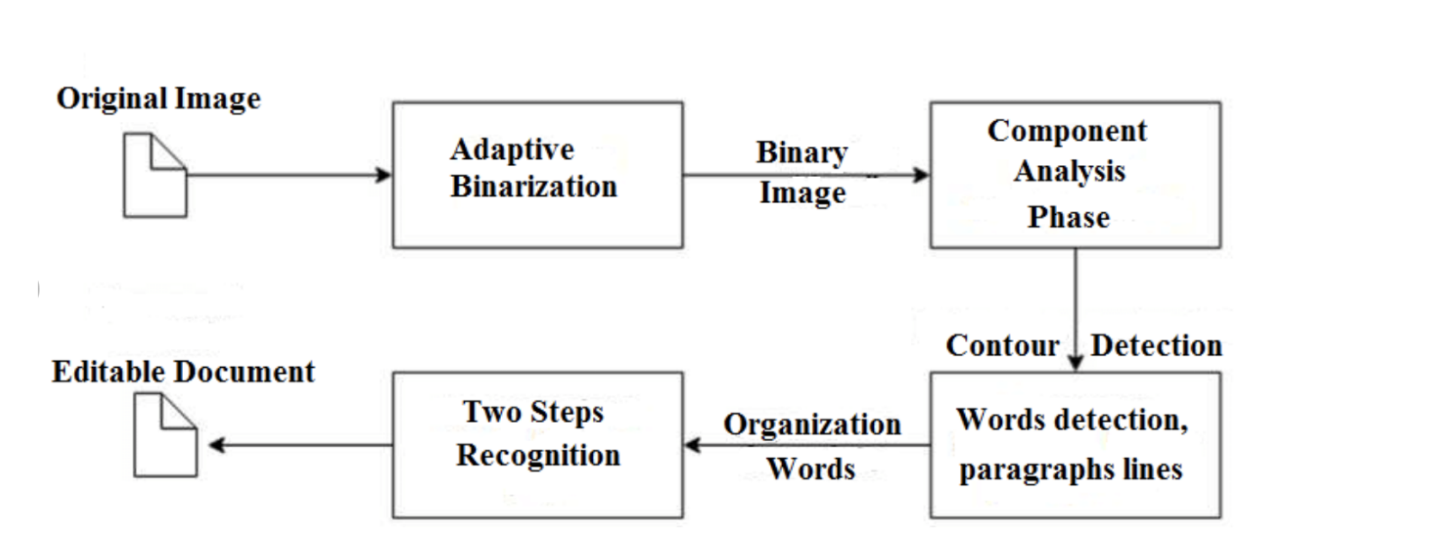
The Hosted Image URL’s are:

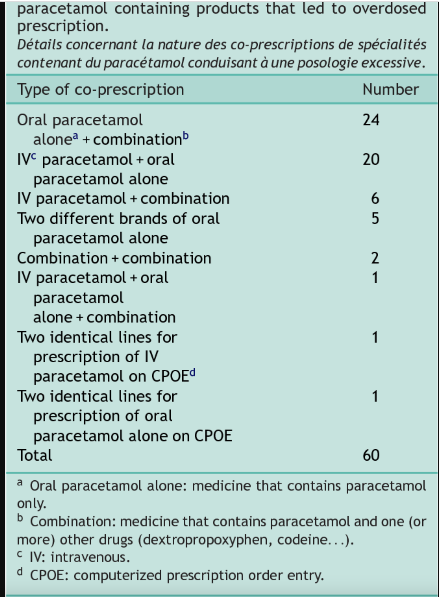
<https://hackathon888.s3.us-west2.amazonaws.com/Common+Drugs.png>

<https://hackathon888.s3.us-west-2.amazonaws.com/Med_para.png>

<https://hackathon888.s3.us-west-2.amazonaws.com/Aspirin_Omeprazole.png>

* + The OCR Text Recognition is analysed into a machine readable format coming from different source providers.



****

**Security**

* + Authorization-tokens for api calls - X-api-key.
  + Hypertext Transfer Protocol Secure.
  + Login functionality and password authentication.
  + Encryption Hashing - using Blowfish Block Cipher Cryptomatic Algorithm (Upto 128 Bits)

***Suggested Features (Future scope of enhancement)***

* + Predictive model for identifying probable diagnosis based on the symptoms recorded
  + Stronger models and algorithms for OCR and local copy document analysis

**Assumptions:**

* Complete ICD 11 Codes and Description was used for auto suggestion functionality of Diagnosis
* A small data set of symptoms was used for auto suggestion functionality

**Technical Design**

1. **Architecture Diagrams**

**Diagram

Description automatically generated**

**Diagram

Description automatically generated**

**Diagram

Description automatically generated**

1. **Model APIs**

The following APIs have been developed for this challenge track –

1. **Autopredict:** [**https://intelli-search-csh.herokuapp.com/autopredict**](https://intelli-search-csh.herokuapp.com/autopredict)
   1. The API can be used to autosuggest Symptoms and Diagnosis
   2. The list of symptoms is extracted from <https://www.kaggle.com/datasets/itachi9604/disease-symptom-description-dataset>
   3. Symptoms and Diagnosis are autosuggested based on similarity of the input text with the set of symptoms, diagnosis etc. (if the list are extended as new symptoms are entered – this model can also take into consideration the newer symptoms as well)
   4. Diagnosis is linked with complete set of ICD 11 Codes
   5. API takes two parameters – “text” and “text\_type”
   6. To predict symptoms – pass parameter text\_type:”symptoms”
   7. To predict diagnosis – pass parameter text\_type:”diagnosis”

**How to set-up**

Store the API in a folder structure as mentioned: API> Script folder> in .py format and then run the following code: python IntelligentSearch.py run on command prompt and then start the server

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

**Workflow with UI**

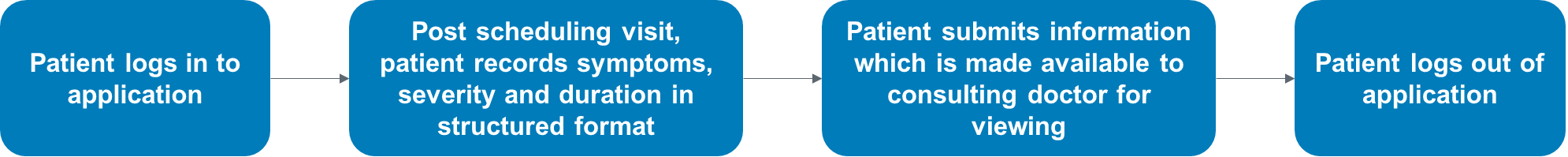
**Patient side**

Graphical user interface, website

Description automatically generatedGraphical user interface

Description automatically generatedGraphical user interface, text

Description automatically generatedGraphical user interface, website

Description automatically generated****

**Provider side:**

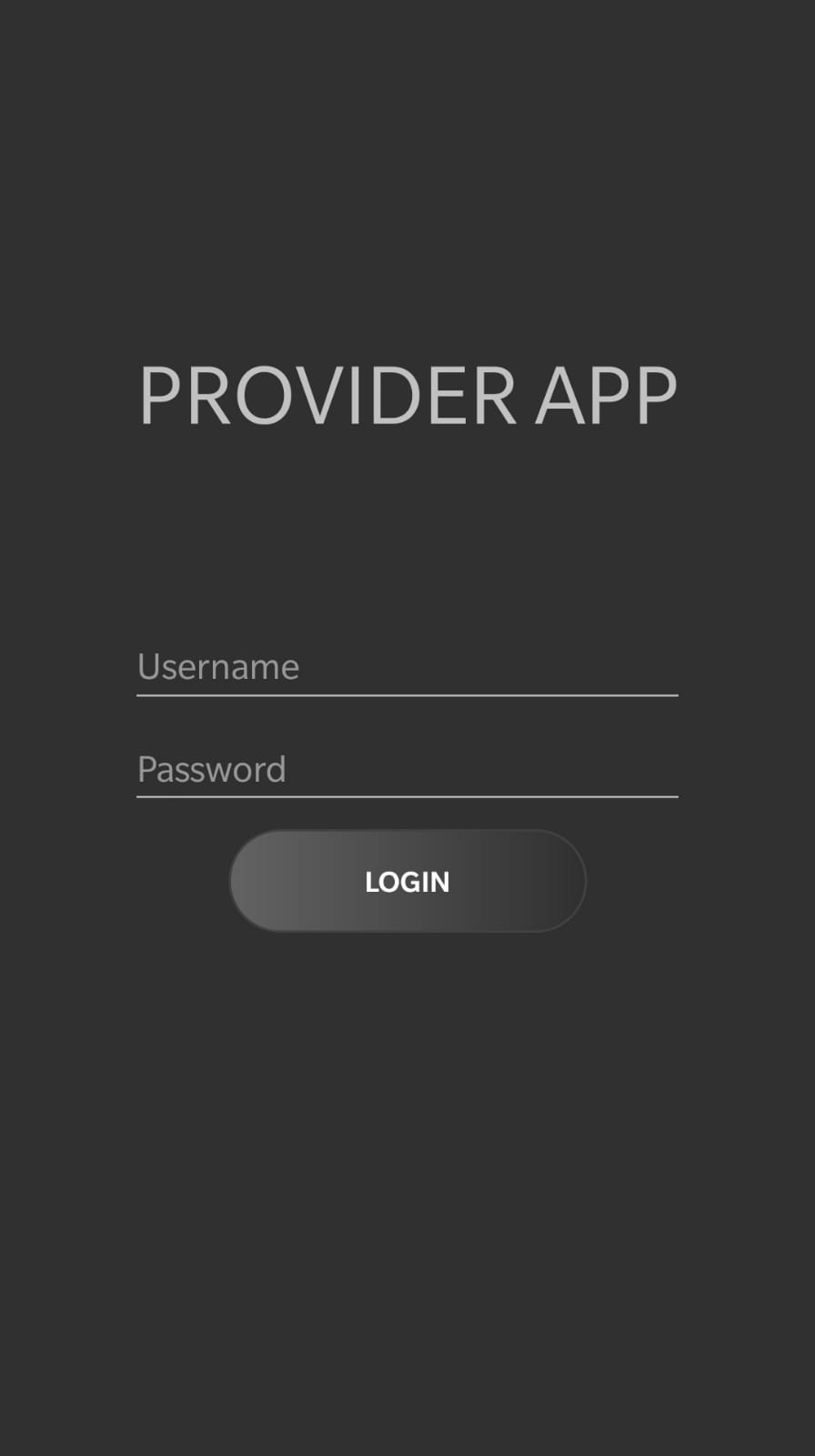
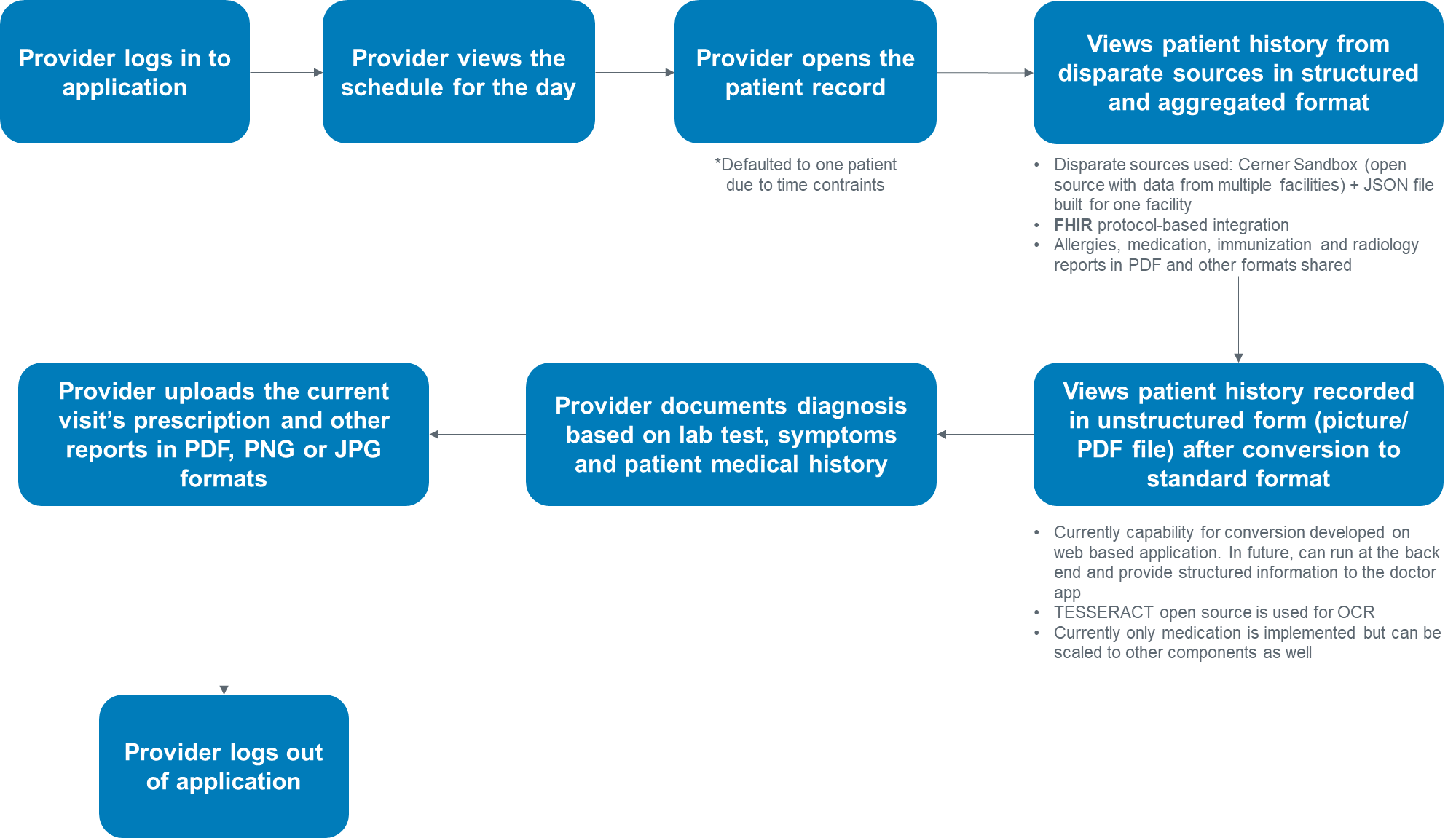
Graphical user interface

Description automatically generatedGraphical user interface, application, email

Description automatically generatedText

Description automatically generated with medium confidenceA screenshot of a computer

Description automatically generated with medium confidenceA screenshot of a computer

Description automatically generated with medium confidence ****  ****

Graphical user interface, website

Description automatically generated

**Sample data**

* **Autosuggest**



* **Structured and unstructured data**

**Graphical user interface, text

Description automatically generated**Unstructured Data

Structured Data