Medical Expertise Style Transfer System for Layman Patients

Mohammad Sabik Irbaz
ID: 160041004
sabikirbaz@iut-dhaka.edu

Abir Azad
ID: 160041024
abirazad@iut-dhaka.edu

Anika Tasnim Preoty
ID: 160041044
anikatasnim@iut-dhaka.edu

Tani Barkat Shalanyuy
ID: 160041083
tanibarkat@iut-dhaka.edu

Abstract

Due to the huge cognitive bias and the curse of knowledge, there is a notable communication gap between experts and laymen. This communication gap creates a huge problem in the medical domain. The patients do not understand what the doctors (domain expert) are saying and the doctors also face some ambiguity issues since they are not used to the laymen style. Bridging the gap between laymen and experts is a challenging task as it requires the models to have expert intelligence in order to modify text with a deep understanding of domain knowledge and structures. To bridge the gap between doctors and patients, we designed a web application in aid for layman patients to get a better understanding of expert texts by medical experts. We used two transformer-based deep learning models for inference. The qualitative evaluation was performed in a cognitive walk through which mostly provided us with positive feedbacks and some constructive ideas. While evaluators were happy to test the working feature of our conversion system they also asked for some added feature like adding a summarize section and a function for providing the external links related to the topics for extra readings.

Introduction

When doctors communicate with patients one of the problems that arises most of the time is the curse of knowledge. Whatever the doctors are saying the patients cannot understand. The problem arises even more when the communication base is online. A simple dictionary is not enough for getting the whole meaning in case of medical terms. Because there are thousands of terms with multiple meanings. The only possible solution would be if someone describes the whole thing in laymen terms. In our project we tried to provide this simple solution with the help of deep learning and we are offering this service through our website. Here people can write their expert sentences and get the laymen conversion in an instance. This is more convenient than a traditional medical dictionary because it provides the direct sentence form in a laymen version. Our website will also include voice input and OCR. So people can write up the sentence or just use voice input or provide a image / scanned writing like prescription. Our system is end to end and it will give the output in full sentence on the screen. There will be an option to read out the sentence also.

Requirement Analysis

Our project is a user-centric project. It is based on the user needs. That's why the requirement analysis is also user based. But at first for getting a good idea about the problem we had to build a conceptual model. Conceptual model is a high-level description of how a system is organized and operates. It enables designers to straighten out their thinking before they start

laying out their widgets. Developing a conceptual model involves, envisioning the proposed product based on the users' needs and other requirements identified. Conceptual model based on activities - Interaction types are: Instructing, Conversing, Manipulating, Exploring. Then we thought the whole system as a goal-directed problem solving activity informed by intended use, target domain, materials, cost, and feasibility. We had to communicate with the user to set the requirements properly. Some important factors here are making sure that the users' views and expectation of the new product are realistic and ensuring that, there are no surprises, no disappointments for users when the product arrives.

A requirement is a statement about an intended product that specifies what it should do or how it should perform. One of the aims of requirement activity is to make the requirement as specific, unambiguous, and clear as possible. Some of the practical issues that arise in each requirement identification are- Who are the users? What are 'needs'? Where do alternatives come from? How do you choose among alternatives?

Prototyping

A prototype is a limited representation of a design that allows users to interact with it and to explore its suitability. Evaluation and feedback are central to interaction design. Prototypes answer questions, and support designers in choosing between alternatives. Prototype serves a variety of purposes: Test out technical feasibility of an idea; Clarify vague requirements; Do user testing and evaluation; check design compatibility.

There are mainly two types of prototyping: Low Fidelity and High Fidelity Prototyping.

	Advantages	**Disadvantages**
Low Fidelity	- Lower development cost - Evaluate multiple designconcepts - Useful communication device - Useful for identifying marketrequirements	- Limited error checking - Poor detailed specification to code to - Facilitator driven - Limited utility after requirements established - Limited usefulness for usability tests - Navigation and flow limitations
High Fidelity	- Complete functionality - Fully interactive - User-driven - Clearly defines navigational scheme - Use for exploration and test - Look and feel of final productServes as a living specification	- More expensive to develop - Time consuming to create - Inefficient for proof-of-concept design - Not effective for requirement gathering

All prototypes involve compromises. For our case we chose High Fidelity prototyping. As it provide more realistic feeling. Even though it is more time consuming but it can provide the

most consistent feedback. And in our case the interaction model is not very complex. So it was in one had easier to build and on the other hand more efficient way to go.

Evaluation

Evaluation is the process of systematically collecting data that informs us about what it is like for a particular user or group of users to use a product for a particular task in a certain type of environment. The basic premise of user-centered design is that users' needs are taken into account throughout design and development. This is achieved by evaluating the design at various stages as it develops and by amending it to suit users' needs. The evaluation approaches are - Usability testing, Field studies, Analytical evaluation. Methods of evaluation can be – Observing, Asking users, Asking experts, Testing, Modeling. There can be Heuristic evaluation or a guided Walkthrough to perform the evaluation process properly.

In our evaluation process we have gained mostly positive feedback especially from the patient side. There was some mixed feedback also. Some questioned the procedure and some also said the product does not matter to them. But most of them were not our target audience. So within our target audience we could generate supporting feedbacks which eased our designing easier.

Methodology

In our workflow we started with the idea. The idea started centering on a common need. Breaking the barrier of expertization, removing the curse of knowledge. Then we worked on the backend software to build a deep learning model from the ground that can detect expert sentences and generate laymen conversion. Then for reaching the general customer we need to have a common front end. Which also demand certain evaluations. That's why we first created the prototype of the interface. Then we approached for the evaluation. We evaluated our front end and back end simultaneously side by side. Though they were not integrated together we tried to give the evaluators a complete scenario. Then we went through several designing, evaluation and redesigning circles. The development process is still going on. The recent edition on development is attached here.

Requirement Gathering

An important step to formulate an user-centered design is to gather data. There are many ways to gather data. We used the following techniques for data gathering.

- *Interviews:* We interviewed potential users of different ages and expertise to get a holistic view of what the requirements are. We collected some quantitative and qualitative data which seemed most appropriate for the task. Though it took a lot of time but by doing this we actually get some field overview.
- *Direct Observation:* This part is based on our research and observations. We collected some qualitative data here to justify our intuitions and findings.

We had to go through the current procedure in details and study the general approach in depth. As a client consultant we had to prepare certain questionnaires that can reflect the similar ideology. The requirement gathering process enriched as time went by and we continued it until we had enough evidence to start with the project.

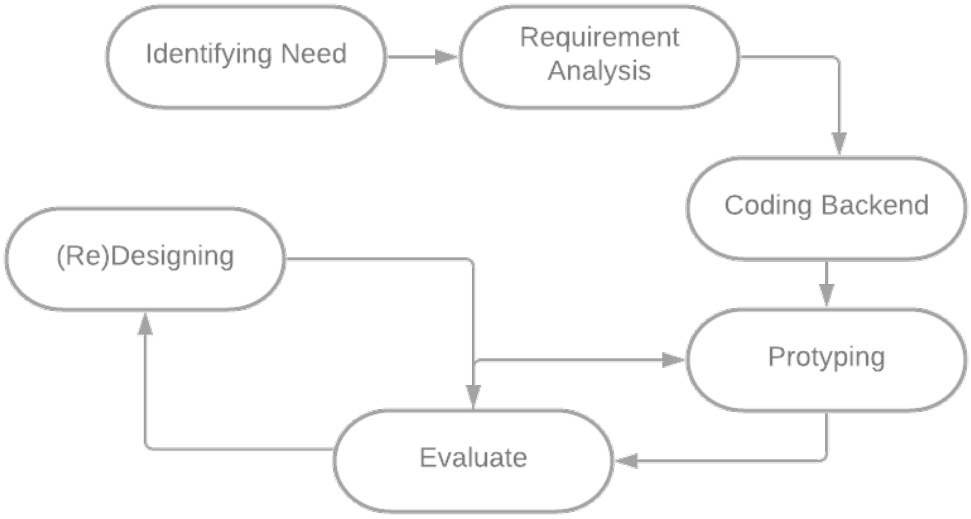

Figure 1: HCI Lifecycle Model

Affinity Diagram

The affinity diagram organizes the individual interpretation session, or affinity, notes into a wall-sized, hierarchical diagram grouping the data into key issues under labels that reveal the customer's needs. It shows in one place the common issues, themes, and scope of the customer problems and needs acting as the voice of the customers. We created the affinity diagram and used certain colors to differentiate the tasks and their importance regarding urgency. Green was for group label summarizing an area of concern, pink was for group label summarizing a set of group, blue for group label summarizing points below and white for individual point captured in interpretation.

Creating Persona

Personas are the synthesis people motivated from real users capturing a set of characteristics with their name, background and goals. For creating our personas we communicated with our customers who are actually any range of literate people. So we had to talk with people of different age ranges and our personas actually reflect this fact. For creating our persona we also thought about the background of our users. That's why we tried to add personas habituating in different classes.

Creating Scenarios

A scenario is an "informal narrative description". It describes human activities or tasks in a story that allows exploration and discussion of contexts, needs, and requirements. It does not explicitly describe the use of software or other technological support to achieve a task. The construction of scenarios by stakeholders is often the first step in establishing requirements. We created some scenarios to associate the personas and reflect them to the target population. For scenarios we thought of daily usual cases of the personas and tried to make them as relatable as possible.

Prototyping and sketching

A prototype is a limited representation of a design that allows users to interact with it and to explore its suitability. There are two types of prototyping, i.e, low-fidelity and high-fidelity. We did both for our project.

Low-Fidelity Prototype

A low-fidelity prototype is one that is quick, cheap, easily changable and does not look very much the final product. For our case we used card-based prototyping technique. We used this technique since our plan was to build an interactive web app. Card-based prototyping is very flexible for website development prototyping.

Figure 2 shows us the landing page of our web application. There we can see our goals and features that we integrated. Specifically, we have 2 goals.

1. Expertise Classifier : Given a text we can classify the expertise of that.
2. Expert to Layman : Given an expert style text we can transfer it to layman style text.

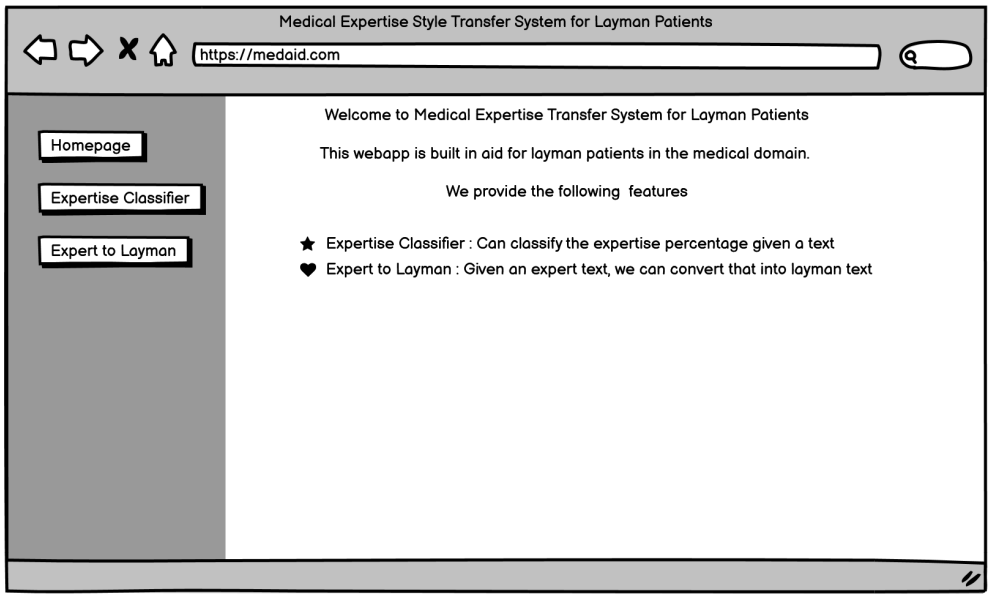

Figure 2: LoFi Prototype – Homepage

Figure 3 shows us the page where we will land if we click on 'Expertise Classifier' in Figure 2, 3, 4, 5 and 6. There we have a form to write something and a button to check the expertise style of whatever we write. When we click on 'Check Expertise Style' button, we land on Figure 4.

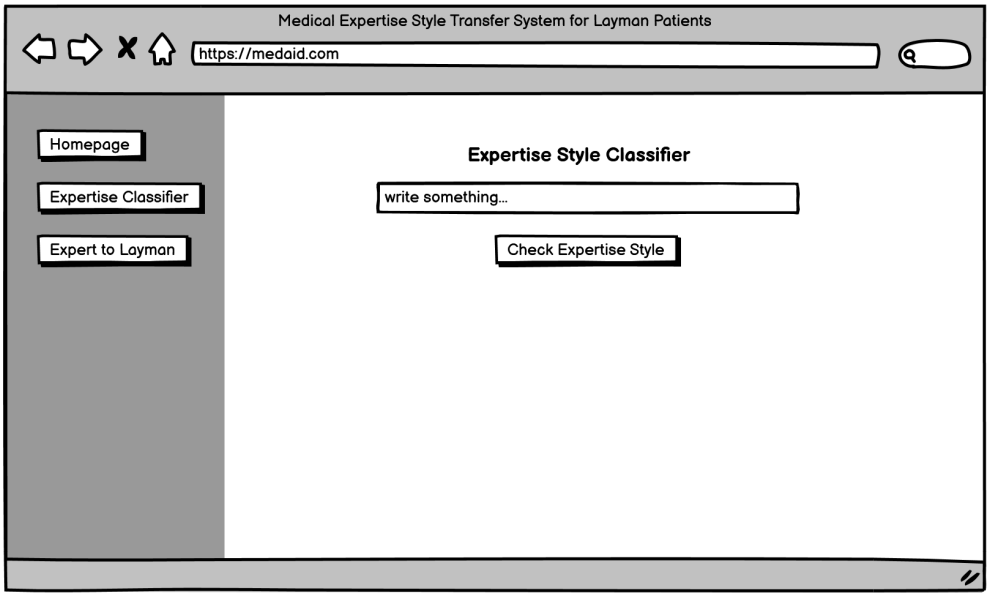

Figure 3: LoFi Prototype – Expertise Classifier

Figure 4 shows us the page where we will land if we click click on 'Check Expertise Style' button in Figure 3. This page shows us the results after style classification. We show the input text, expertise percentage (how much expert the text is) and the style of the text there.

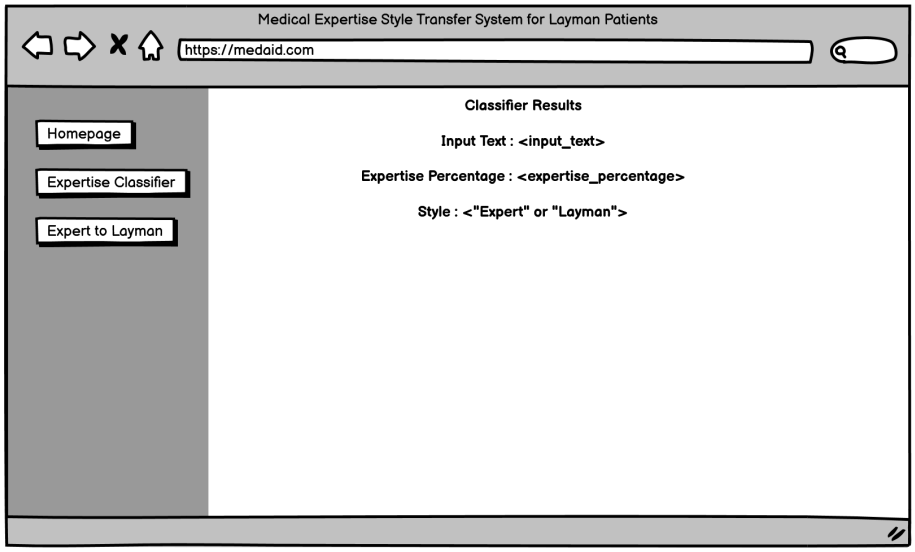

Figure 4: LoFi Prototype – Check Expertise Style

Figure 5 shows us the page where we will land if we click click on 'Expert to Layman' button in Figure 2, 3, 4, 5 and 6. We have three ways of taking inputs here.

- Writing Something
- Provide Audio Input [Converts Audio to Text first]
- Provide Image Input for OCR [[Converts Image to Text first]]

After giving the inputs, when we click on 'Transfer to Layman' button, we land on Figure 6.

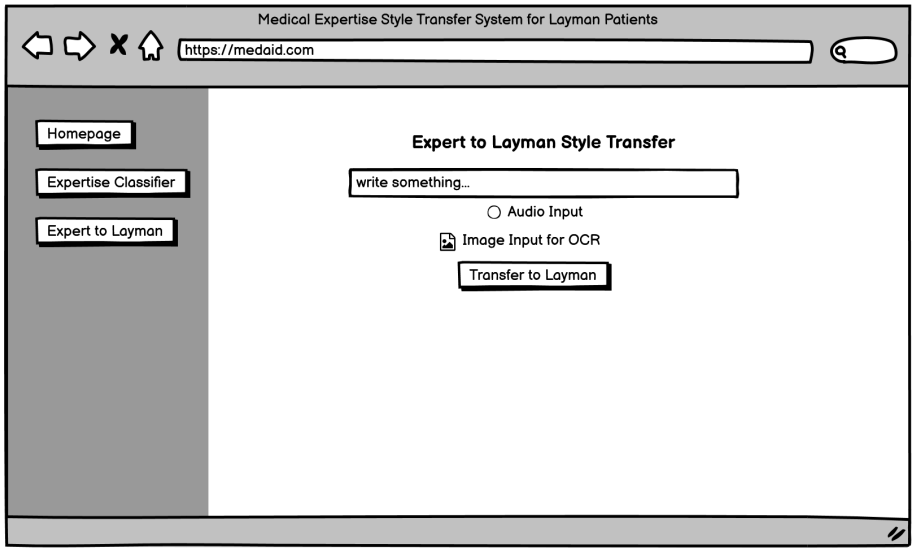

Figure 5: LoFi Prototype – Expert to Layman

Figure 6 shows us the page where we will land if we click click on 'Transfer to Layman' button in Figure 5. This page shows us the results after style transfer. We show the input text and the generated layman text. We can also get them as audio output if need by clicking the audio button.

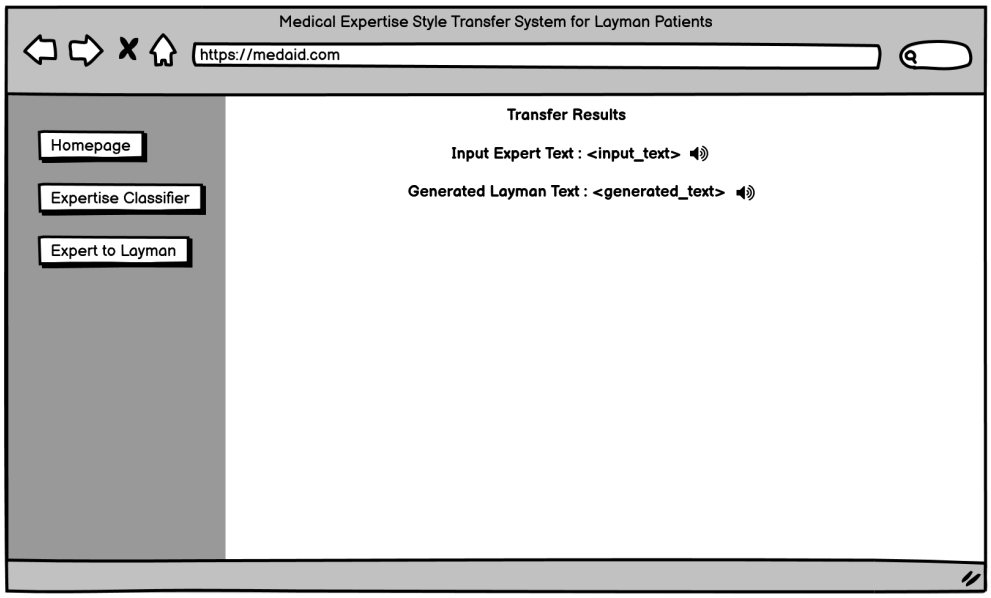

Figure 6: LoFi Prototype – Transfer to Layman

High-Fidelty Prototype

We built a high-fidelity working prototype which covers most of the basic features. We did not implement Audio to Text, Image to Text and Text to Audio conversion in the working prototype. The full overview of the implementation of the web application is provided in the attachment and also uploaded in GitHub [1]. We also provided a video walkthrough [2] on our web applcation.

[1] https://github.com/msi1427/Medical-Expertise-Style-Transfer-System-for-Layman-Patients
[2] https://youtu.be/1viSX3BSXQ4

Experimental Evaluation

Experiment Design

Experiment design is the process of deciding what variables to use, what tasks and procedures to use, how many participants to use and how to solicit them, and so on. There is two term related to experiment - The signal and the noise. The signal is a variable of interest and the noise is everything else (random influences). Experiment design seeks to enhance the signal, while minimizing the noise. One thing that precedes experiment design is the ethics approval, where we have to respect the safety, welfare, dignity and privacy of the participants.

Experimental Variables

The experimental design starts with deciding experimental variable. Experimental variable forces us to craft narrow and testable questions. There are different types of variables.

- Independent variables
- Dependent variables
- Control variables
- Random variables
- Confounding variables

Experiments designed with independent variables are called factorial experiment. In our project the independent variables were – Gender, height, weight, visual acuity etc. A dependent variable is a measured human behavior (related to an aspect of the interaction involving an independent variable). It depends on what the participant does. In our project the dependent variable was literacy, age. A control variable is a circumstance (not under investigation) that is kept constant while testing the effect of an independent variable. In our whole evaluation process we kept the prototypes same and controlled the outlook. A random variable is a circumstance that is allowed to vary randomly. In our evaluation process we did not keep track whether the participants are eager to know the details of their medical treatment or not. We evaluated on them regardless. It was a random variable in our case. A confounding variable is a circumstance that varies systematically with an independent variable. One of the confounding variable in our case was the gender. Though we took it as an independent variable, even then we had this conversation, does the gender makes a difference on how they perceive our system.

The test condition was natural in the evaluation process. We did not manipulate any external states while taking the feedback. One of the requirements was to know the natural feedback of the people using our system.

While evaluating we took both mainly qualitative evaluation. In this short period of time we could not reach enough participants to take quantitative evaluation. The very people we could reach were asked specifically about each function of the system and how they would like it to interact. They gave mostly constructive feedbacks.

Evaluation Method

The evaluation process was done in Cognitive Walkthrough. Our idea wast to focus on ease of learning. We presented the aspect of design and usage scenarios. Participants were told the assumptions about the user population, context of use, task details. They walked though the design prototype with the scenario. They were mainly driven through three questions. The questions are - Will the correct action be sufficiently evident to the user? Will the user notice that the correct action is available? Will the user associate and interpret the response from the action correctly?
We prepared the prototypes of our interface. We also prepared a working version of the back end system that can generate simple version of expert sentences. So that the evaluators have a nice idea about both the front end and the back end.

Evaluation results

Our backend software that can effectively generate laymen sentences made a huge difference in the evaluation process. The evaluators were mainly interested in how well the software can perform rather than the usual interface. One of the reasons for that can be our qualitative approach. Quantitative approach might give different outcome. As our front end and back end were evaluated differently the backend evaluation actually made the front end evaluation slower than expected. Every evaluation step in the whole process only required single input actions for the users to make the process even smoother. The software was very much self explanatory and most of our evaluators were expert enough. We asked them to look through the eyes of general people. They did not face any confusion while evaluating. User performance mainly varied on basis of their taste. The method was actually very much easy to lean and operate. While checking the backend software, the participants were sometimes tired of writing expert sentences again and again as there was no voice input or OCR system running. They heavily insisted on implementing the OCR system because one of their main attractions was being able to know the readings on the medicine scripts which are mainly very much expert literature. Writing the whole thing would be very much tiresome and frustrating. Another thing they asked for was a summarized laymen version for long paragraphs which can also be useful as well as time efficient for the users. Some of the users asked to add some more features, like adding external reading links along with the laymen conversion.

Conclusion

We built a web application in aid for layman patients to get a better understanding of expert texts by medical experts. We used two transformer-based deep learning models for inference. Initially, we did low-fidelity prototyping using Balsamiq. Later, we filled up most important requirements with the web application we built using Flask. We found out that the layman users are satisfied with the generated outputs but they wanted a more summarized output. Some users preferred having related resources for further reading. In future, we plan to generate more sensible layman text. We also plan to add all the features we have shown in our updated low-fidelity prototyping and try to develop incrementally.

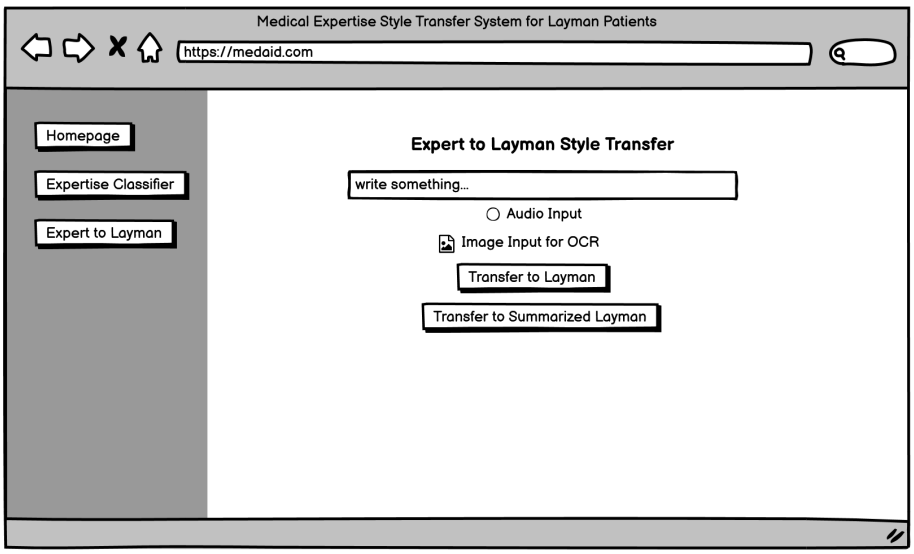

Figure 7: Updated LoFi Prototype – Expert to Layman

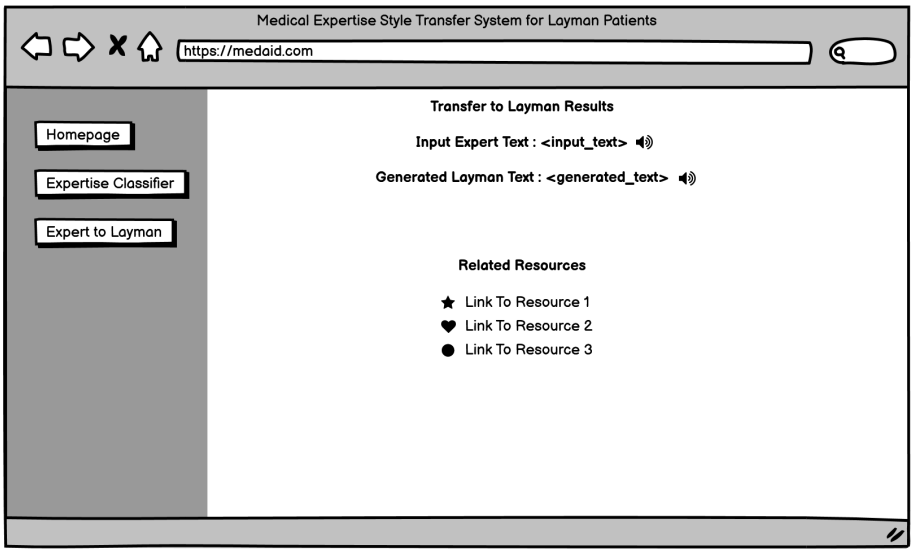

Figure 8: Updated LoFi Prototype – Transfer to Layman

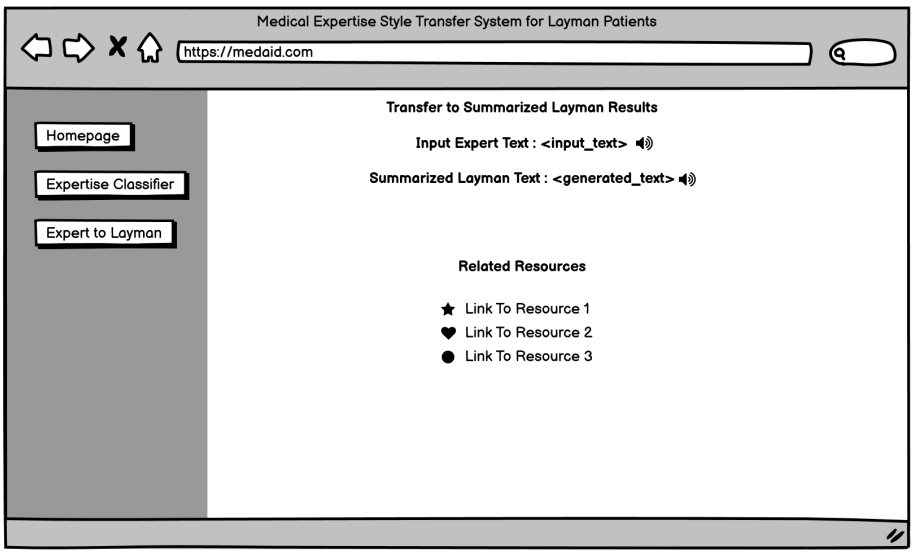

Figure 9: Updated LoFi Prototype – Transfer to Summarized Layman