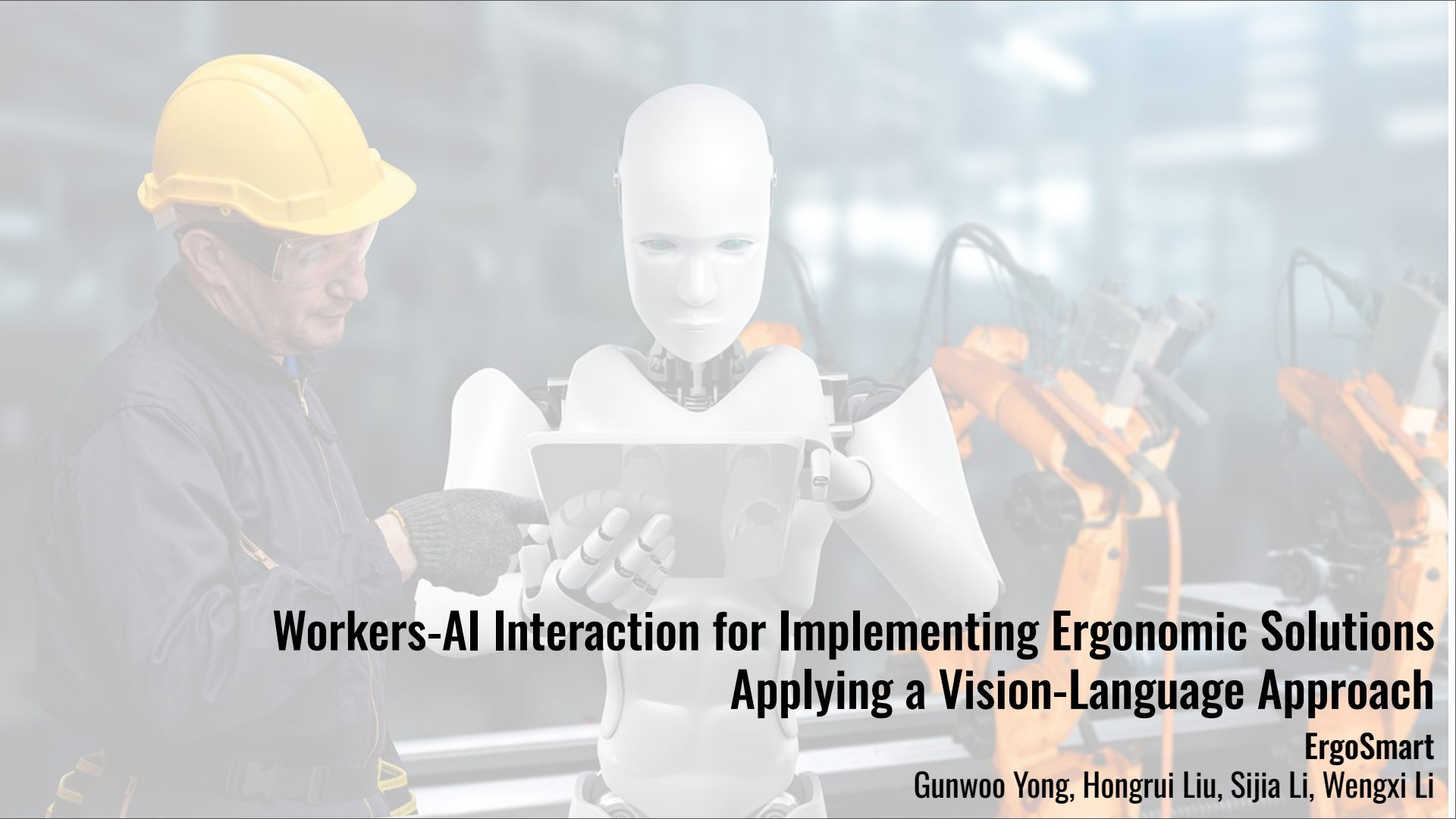
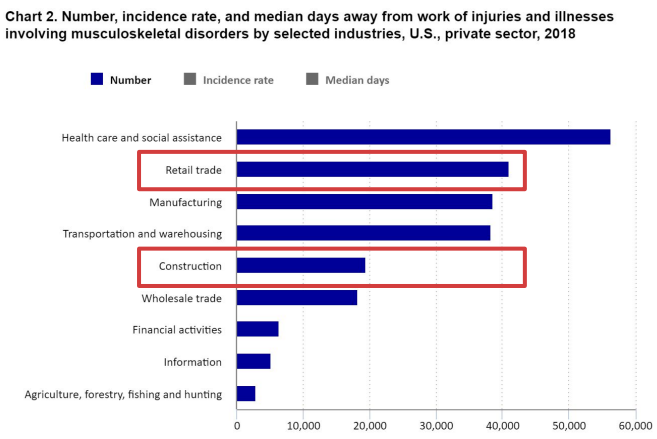


Workers-AI Interaction for Implementing Ergonomic Solutions Applying a Vision-Language Approach

ErgoSmart

Gunwoo Yong, Hongrui Liu, Sijia Li, Wengxi Li

Motivation



Problems

- More than 60,000 workers in the retail trade and construction industry suffer from work-related musculoskeletal disorders (WMSDs) in a single year (U.S. BLS. 2020) due to ergonomic risk factors (e.g., awkward postures).
- Despite the severity of WMSDs, there are not enough ergonomic experts who can solve ergonomic problems in worksites (Sneller et al. 2018).

Therefore, an **automated explanation** of ergonomic problems and solutions to workers is required.

Findings from Interviews and Literature Review

- How can they recognize ergonomic problems and solutions?
  - Workers tend to recognize problems after one of their body parts are painful.
  - Existing risk assessment tools could only measure the risk level (Wang et al. 2021).
    - Users still cannot identify solutions.
  - Workers are not familiar with ergonomics (Sneller et al. 2018).
    - They have come up with solutions based on their common sense.
- What functions are necessary for end-users?

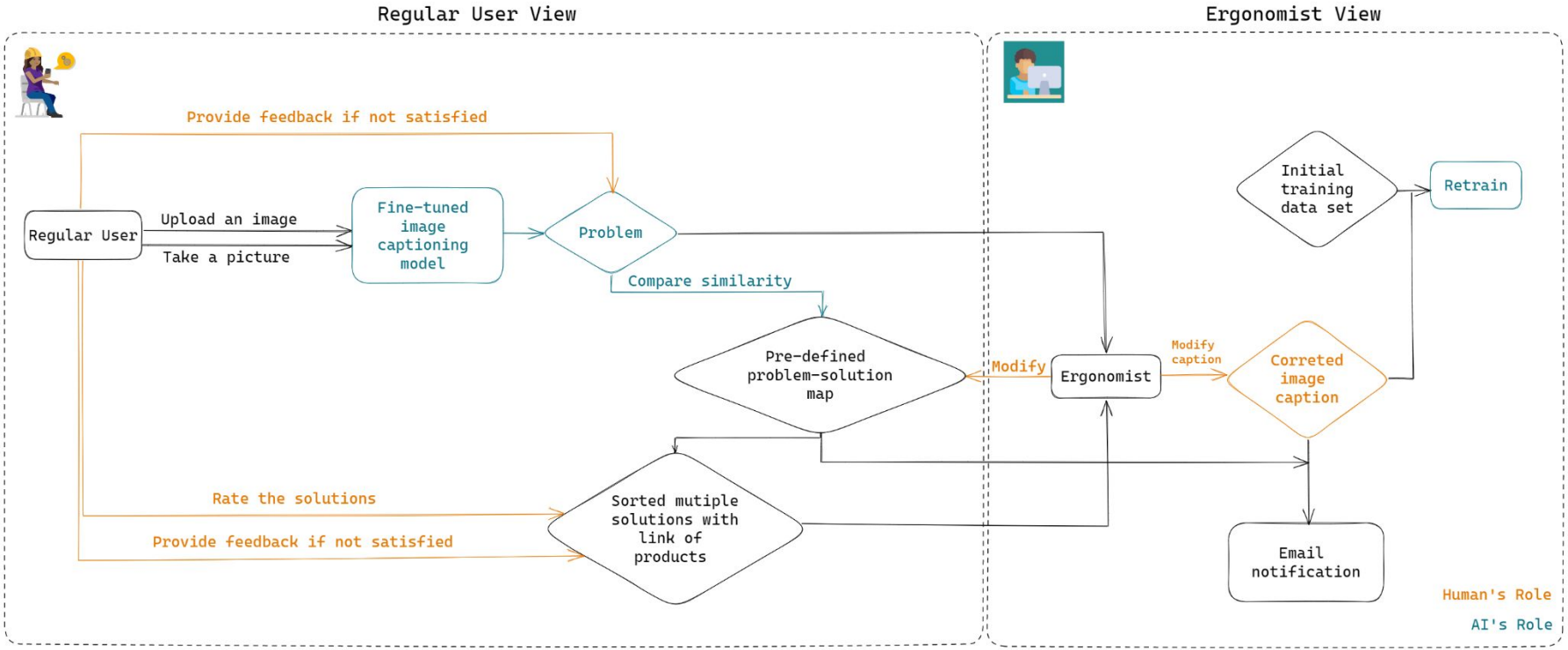
✅ Self-camera input system.

✅ Quick response.

✅ Multiple solutions provision.

✅ Links associated with proposed solutions.

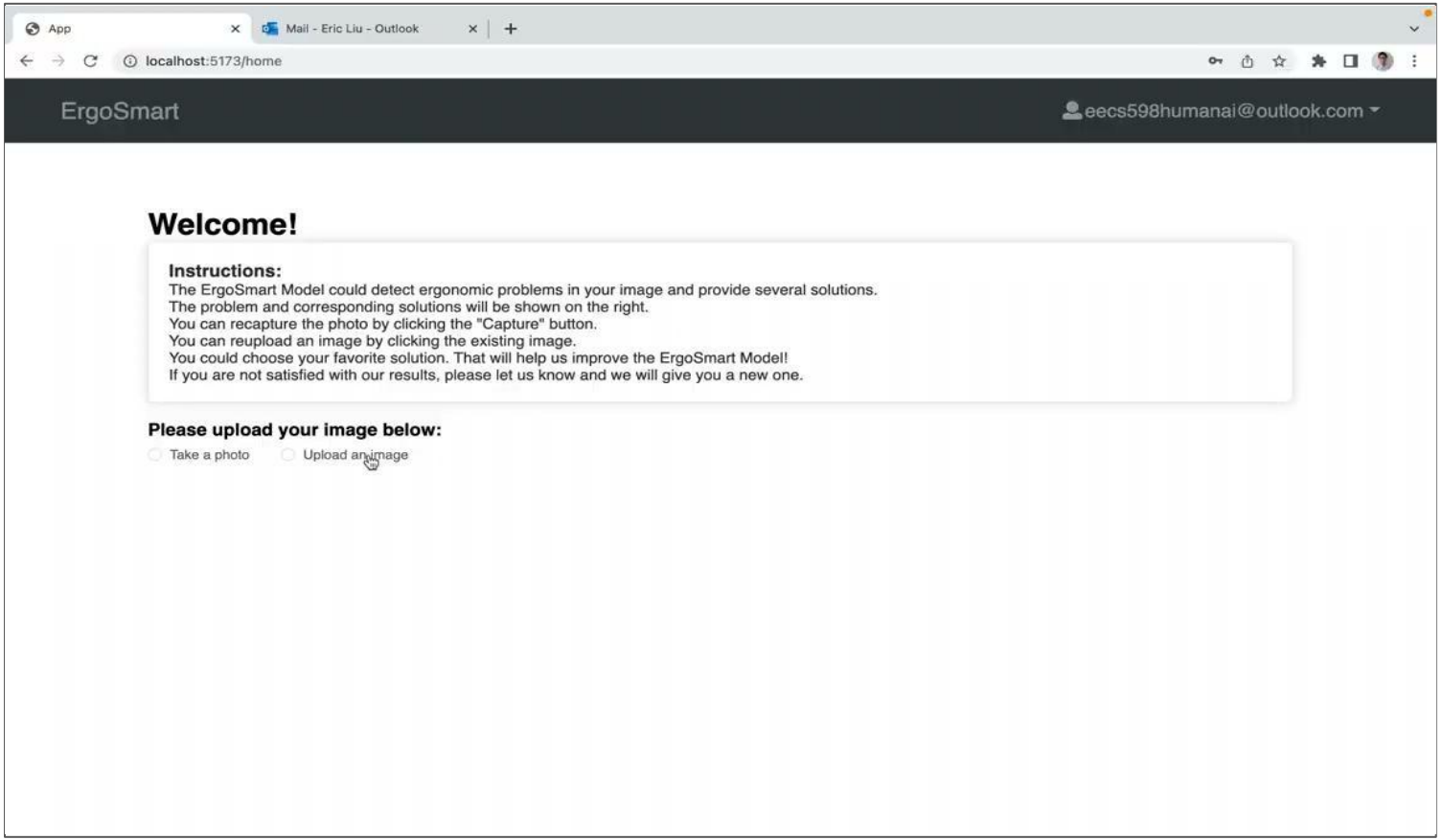
System Diagram



Vue.js Flask web development, one drop at a time SQLite PyTorch HTML JS CSS

https://adoption.microsoft.com/en-us/guides/construction/
https://www.library.manchester.ac.uk/services/research/

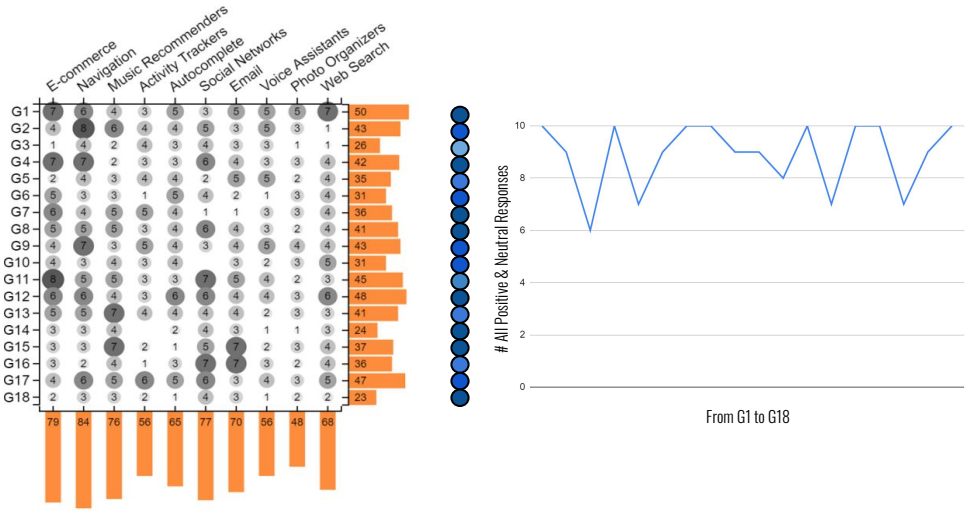
User Journey



Evaluation & Result

Metrics: 3-scale survey based on "18 Guidelines for Human-AI Interaction (Amershi et al. 2019)" and Human-evaluated Accuracy

- To what extent does our model follow the Microsoft's Human-AI Interaction guideline
- To identify the accuracy of ergonomic problems and solutions detection



Comparison of # All Positive & Neutral Responses Between Amershi et al. (2019)' results and Ours

Result

- 10 participants who knows/learns the problem.
- Our model achieved high human-AI interaction performance.
  - More than 80% of participants said our model followed 10 guidelines ("Yes") out of 18.
- Our model is vulnerable to get users' current situation (e.g., time and location).
  - 80% of users replied our model did not follow G3 guideline.
- The overall accuracy of our model is 0.7339.

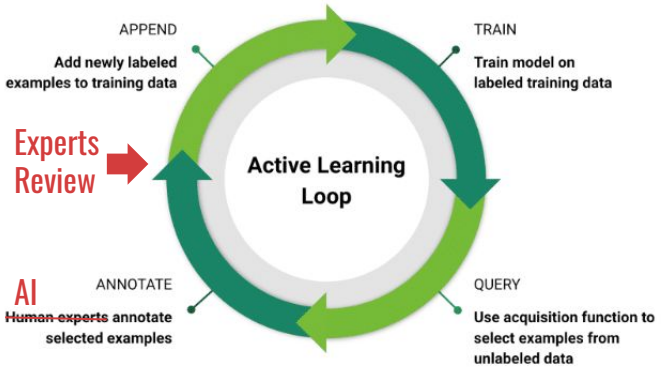
Discussion

- Human feedback is necessary to improve the performance of AI models
  - If users do not provide feedback, ergonomics will never find the wrong captions and retrain the model
- Designing a feedback loop can be difficult when users don't have enough background knowledge
  - It is hard for users to find out whether the problems or solutions are correct
- Our system cannot cover all 18 HCI guidelines and is not necessary to
- More features are added after bake-off's iteration, which improves user experience
  - Add camera capturing and allows multiple solutions
- A well-designed interface can allocate AI's and human's responsibility reasonably

Future Work

- Processing multiple problems simultaneously.
  - By combining our model with object detection and pose estimation, we can extract all workers taking awkward postures.
- Improving the ambiguous five-rating system.
  - By describing a specific instruction of the rating system, we can notify the function of the five-scale rating system.
- Including the active retraining system.
  - By retraining with positive examples continuously, our model cannot only obtain more data, but also enhance our model's ability.

Example of multiple ergonomic problems in a single scene.



Modified Active Learning Loop

In-Class Comments

Interaction Perspective

- (Solved) In the user's view, it will be better to show the functions (e.g., rating) after generating a caption.
- (Solved) Emails do not have the content.
- (Solved) There will be better if you include instructions of your functions.
- (Solved) Split the ergonomic problem and the solution.
- (Solved) Think about how to employ users' feedback in your Human-AI joint system.

Model/Dataset Perspective

- (Unresolved) It will be better to add a function using text as an input.
  - Workers cannot describe their problems due to the lack of knowledge in ergonomics, we used images as an input, which does not require such knowledge.
- (Unresolved) How can your model handle multiple problems or zero problem case?
  - As our model is based on the prior work which detects a single ergonomic problem for the risk assessment, we assume that there is only one problem.
- (Solved) Add more solutions for each problem and give a link for the solution.
- (Unresolved) Check the change of wrong answers after retraining.
  - As there are not enough corrections for wrong answers, we could not show a dramatic performance change.

Problems

Many construction workers suffer from musculoskeletal disorders (e.g., the back pain) due to awkward postures, repetitive tasks, and etc.
Ergonomics aims to reduce physical injuries and stress, while improving the productivity in the workplace.
Identifying ergonomic problems and solutions can reduce ergonomic risks.
However, workers are unaware of ergonomics due to the lack of ergonomists and ergonomic programs.
Thus, the **automated explanation** of ergonomic problems and solutions to workers is required.

Target Users

1. **Construction workers** who are not familiar with ergonomics.
2. **Managers** who supervise construction workers without ergonomic knowledge.
3. **Ergonomists.**

By talking to them, we aim to learn the practical impact of AI-based ergonomic risk analysis.

System

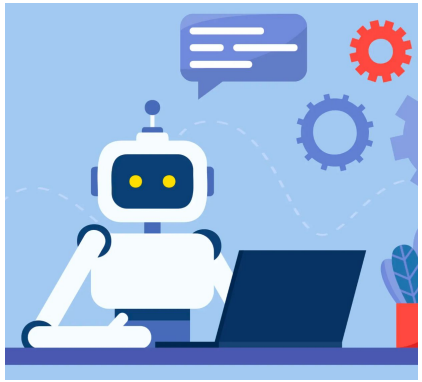
This project is a web application which conducts image captioning to add an explanation of ergonomic problems as well as corresponding solutions.
Also, our application will receive users' feedback to improve the performance.

User Requirements

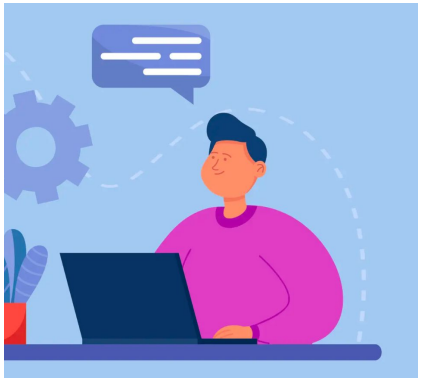
- Process
  - Interviewed three potential users: two are office workers and the other one is a bus driver.
  - Two of them did not know AI well
  - They were showed a demo video of a computer vision-based ergonomic assessment tool and were informed our work at Stage 1.

Storyboard

After a day's ceiling work, Peter Parker not just feels tired but has a dull pain in his shoulder. He wants to know what causes this ache and how to relieve it. He heard ErgoSmart from MJ.



ErgoSmart will take a photo as input and determine which action caused the pain and provide some feasible suggestions instantly. It will also listen to feedbacks. It could provide a new solution if the former is useless.

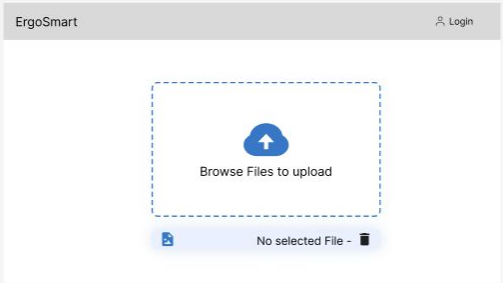


Peter uploads a photo of him at work. ErgoSmart tells him the pain is from holding hands up so try to use a proper tool, such as using a long plastering bar, or to use a higher ladder.

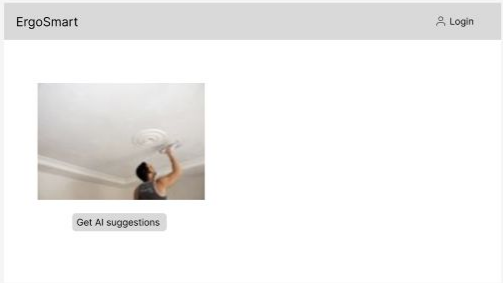
After some days' work, Peter feels better. He comments to ErgoSmart that the advice worked. ErgoSmart now has a positive feedback and will be more confident next time.

Prototype

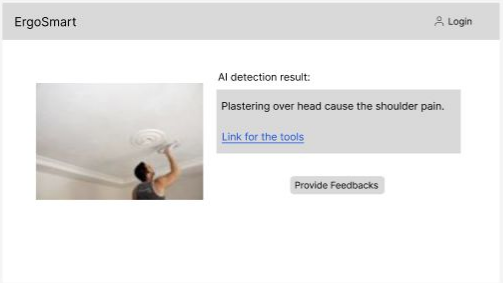
Wireframe - 1



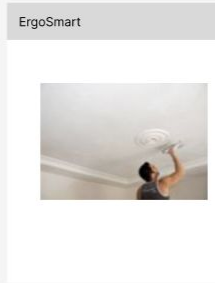
Wireframe - 2



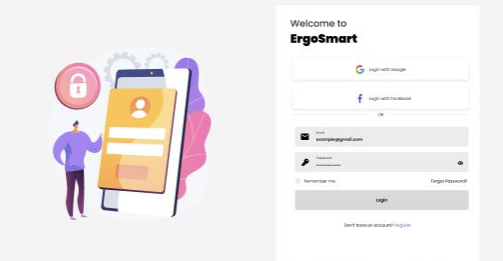
Wireframe - 3



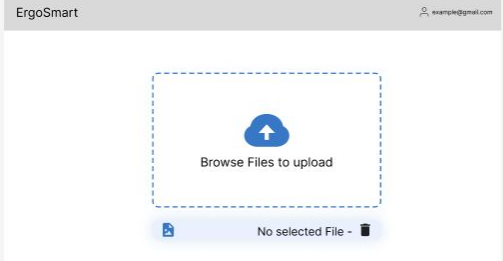
Wireframe - 7



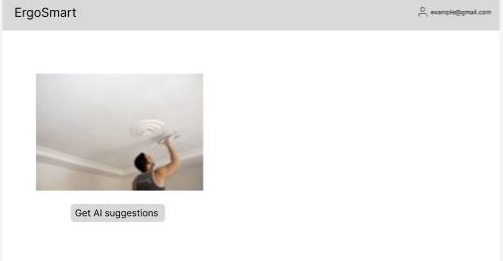
Login



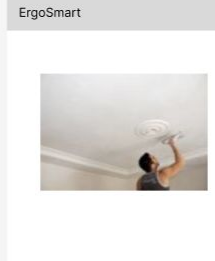
Wireframe - 4



Wireframe - 5



Wireframe - 10



https://www.figma.com/proto/302T9skNiKqW9d7c6tbOsr/ErgoSmart?node-id=1%3A5&scaling=contain&page-id=0%3A1&starting-point-node-id=1%3A5

Storyboard

Worker's working involves some awkward postures with ergonomic risks which might cause back pain due to lifting bricks from low

AI detects the problem and provide explanation as well as solution: a new tool for brick stacking

Worker is satisfied with the explanation but want something other than the new tool

AI provide new solution: a posture for lifting bricks with lower ergonomic risks

Human

AI

Reference

https://www.microsoft.com/en-us/research/project/guidelines-for-human-ai-interaction/
https://www.microsoft.com/en-us/research/uploads/prod/2019/01/Guidelines-for-Human-AI-Interaction-camera-ready.pdf