So far …

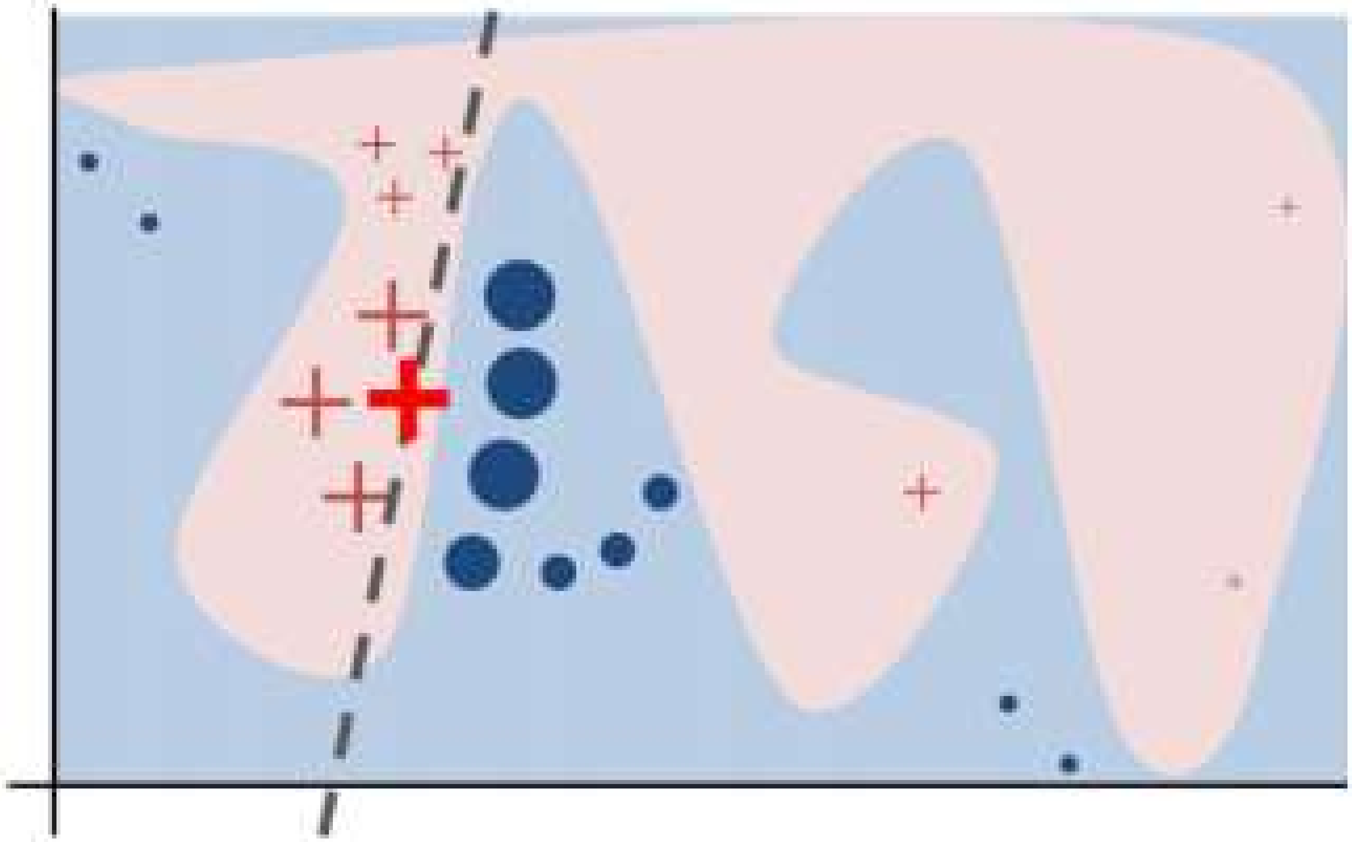
- HCI researchers have used ML to increase interaction possibilities
  - Invisible
    - Realizing variety of input methods
      (recognition technologies for voice, gesture, …)
    - Adaptive interfaces that reduce users' efforts and automate tasks
    - Infer people's states, like their interruptability, their routines, and other important contextual information
    - Other buried? (invisible) low level services
      - Search, travel time estimator, activity tracker
  - Visible (with some kind of presence or virtual form)
    - Almost? Independent intelligent agents
      - Like Siri and Amazon Echo

Visible or not:
When interacting with these AI things:
Interpretability / Understandability

- Need to be able to interpret (understand) the results (made by AI)
  - System should offer explanations
  - this way we can understand mental model of AI
  - that forms the mental model of my own
- what exactly is explanation?
  - What ever it is, needs to be somewhat convincing
    - People will not usually blindly (for some reason) trust the system (or AI)
    - So understandability goes with trust
      - Seemingly obvious bias will lower the trust too (but this is another problem with data themselves)

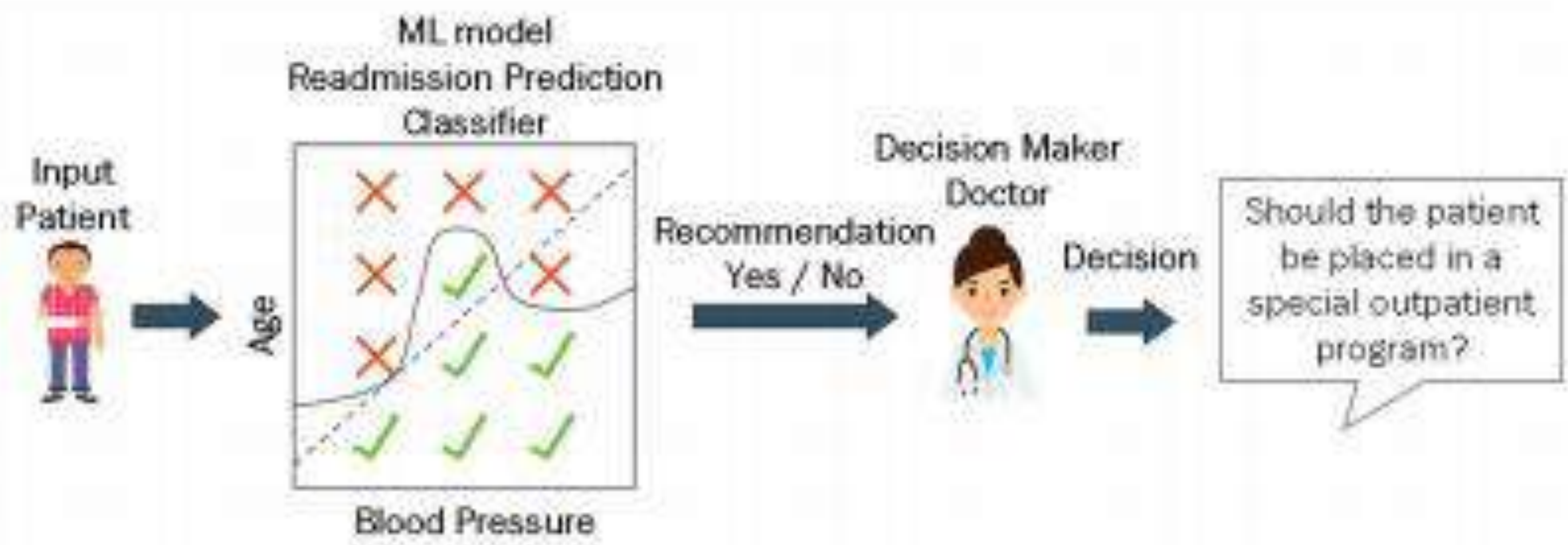
Visible or not:
When interacting with these AI things:
Interpretability / Understandability

- causality: problem is causality is difficult to provide (especially for black box DL systems)
  - Perhaps we somehow offer ad hoc 2nd best guessed explanations? (can be misleading sometimes) → See figure in the next page
  - Although this is not the scope of this course …
- Not explanation of the algorithm but the logical reason of the results
  - Although if above cannot be given, even algorithmic explanation may be ok (transparency)
  - Note that many ML (e.g. DT, SVM?) can offer some explanation while the black box DL seems to have problems
- if causal explanation is not really available, use graphics/Viz, natural language rather than just list of features and input



Visible or not:
When interacting with these AI things:
Interpretability / Understandability

- Remember mental model and Gulf of Evaluation and Execution?
  - Interaction model of system ~= Mental model of user
- System includes AI …
  - but often times, AI's operation model is not known …
    (black box, e.g. image goes in, and it detects "cats" somehow)
    - Users get baffled (low understandability)
  - AI's exact performance is not known (even to developers sometimes who just use pre-existing models)
- If AI model details are not known, make the system easily guessable
  (learn the AI model along the way and cope with the system, while using it)
  - Unexpected results or error like things should be consistent, not erratic and stochastic
  - System should be not so complex (like when too many inputs are needed)
    - Modularize the system (or AI) as much
  - Users will come to establish some kind of a mental judgement/cost model
    - To follow AI or now based on the impact (Remember the cost of false positives > benefit of true positives)


ML model
Readmission Prediction
Classifier
Input
Patient
Age
Blood Pressure
Recommendation
Yes / No
Decision Maker
Doctor
Decision
Should the patient be placed in a special outpatient program?

Mixing AI and Manual/Human Intervention

- The challenge of integrating “intelligent” technologies into people’s lives can be distilled into the automation debate between ‘do it for me’ and ‘do it myself’, or the challenge of <u>when automation is desirable and when people want to feel in control.</u>
- Problem is now we have probablistic output …
- Based on output we have choices of:
  - Throw away AI
  - Engage in dialog to get more info and certify AI result
  - Just trust AI and go ahead
    - e.g. email system: AI thinks user wants to reply
      - get more info
      - just pop up the reply template?

Mixing AI and Manual/Human Intervention

- As a developer, we might consider applying cost analysis (if such info is available … rather than have the user learn it along the way and cope)
- Utility
  - u (A, G) = utility of take action when AI is right 50
  - u (A, ~G) = utility of take action when AI is wrong -200
  - u (~A, G) = utility of no action when AI is right -10
  - u (~A, ~G) = utility of no action when AI is wrong 30
- Expected utility:
  - P(AI giving right answer) * u (A, G) + P(AI giving wrong answer) * u (A, ~G)
    - 0.75 * 50 + 0.25 * -200 = -12
  - P(AI giving right answer) * u (~A, G) + P(AI giving wrong answer) * u (~A, ~G)
    - 0.75 * -10 + 0.25 * 30 = 0
  - Better not to listen to AI …
  - Things can change based on utility value and accuracy of system

Hard performance data or Self certainty data

Mixing AI and Manual/Human Intervention

- Thoughts
  - AI perhaps should have been trained based on this cost model in the first place
    - With what knowledge of what user wants !
  - After all, it merely means that some error handling is needed! (Obviously)
  - The situation just got worse with AI because the output is probabilistic!
    - E.g. Certainty 100% → do something, 99~90% do something else, 80~85% do something else, …
    - Situation is similar to user having to do prompt engineering except that it is not included in ChatGPT itself (users have to figure it out)
      - This is difficult ! Even for developers who will just use it as black box
      - Even if ChatGPT was 100%, the impression might not be so …
        (what to do?)

Mixing AI and Manual/Human Intervention

Sheridan and Verplank (1978) ten levels from human control to computer automation/autonomy

Table 1. Summary of the widely cited, but mind-limiting 1-dimensional Sheridan-Verplank levels of automation/autonomy (Parasuraman et al., 2000).

Level	Description
High	10. The computer decides everything and acts autonomously, ignoring the human.
	9. The computer informs the human only if it, the computer, decides to.
	8. The computer informs the human only if asked, or
	7. The computer executes automatically, then necessarily informs the human, and
	6. The computer allows the human a restricted time to veto before automatic execution, or
	5. The computer executes that suggestion if the human approves, or
	4. The computer suggests one alternative, or
	3. The computer narrows the selection down to a few, or
	2. The computer offers a complete set of decision/action alternatives, or
Low	1. The computer offers no assistance; the human must take all decisions and actions.

Table 2. Persistent, but still misleading, 1-dimensional thinking about levels of autonomy for self-driving cars (Brooks, 2017; Society of Automotive Engineers, 2014).

Level	Description
5.	Full autonomy: equal to that of a human driver, in every driving scenario.
4.	High automation: Fully autonomous vehicles perform all safety-critical driving functions in certain areas and under defined weather conditions.
3.	Conditional automation: Driver shifts "safety critical functions" to the vehicle under certain traffic or environmental conditions.
2.	Partial automation: At least one driver assistance system is automated. Driver is disengaged from physically operating the vehicle (hands off the steering wheel AND foot off the pedal at the same time).
1.	Driver assistance: Most functions are still controlled by the driver, but a specific function (like steering or accelerating) can be done automatically by the car.
0.	No Automation: Human driver controls all: steering, brakes, throttle, power.

Mixing AI and Manual/Human Intervention

One-dimensional levels of automation

➔ Two-dimensional Human-Centered Artificial Intelligence (HCAI) framework

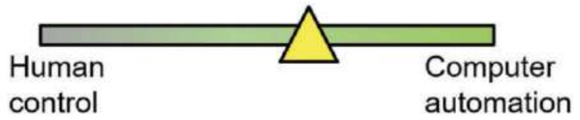


Figure 1. One-dimensional thinking suggest that designers must choose between human control and computer automation.

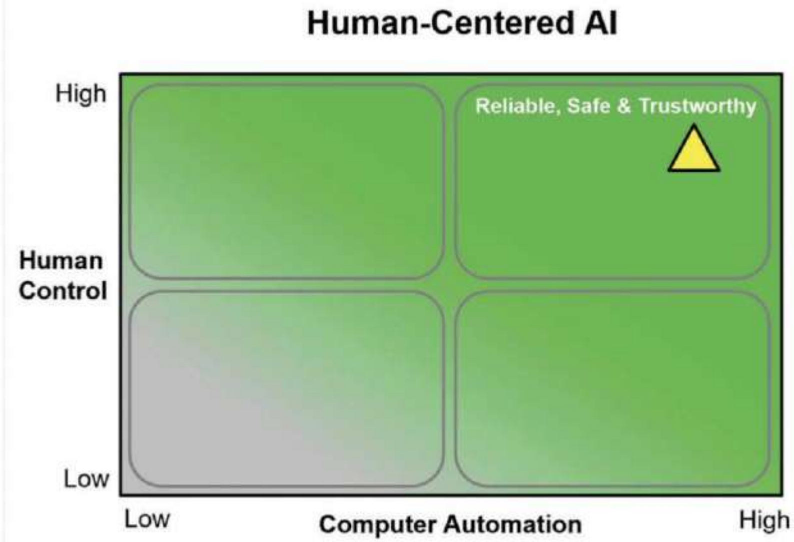


Figure 2. Two-dimensional framework with the goal of Reliable, Safe & Trustworthy, which is achieved by a high level of human control and high level of computer automation (yellow triangle).

Human-Centered Artificial Intelligence (HCAI) framework

(1) Human performance을 높이기 위해 높은 수준의 인간 제어와 높은 수준의 컴퓨터 자동화를 설계하는 방법 제시

(2) 완전한 인간 제어 또는 완전한 컴퓨터 제어가 필요한 상황을 이해

(3) 과도한 인간 제어나 과도한 컴퓨터 제어의 위험을 피한다.

- Boeing 737 MAX's MCAS system
- Tesla Autopilot system
- Deadly human error

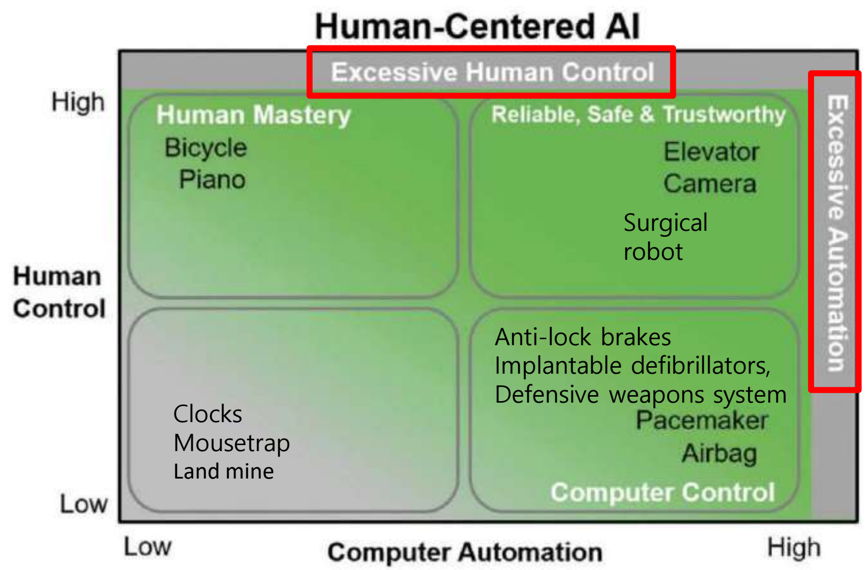


Figure 4. Designers need to prevent the failures from excessive automation and excessive human control (gray areas).

HCAI Examples

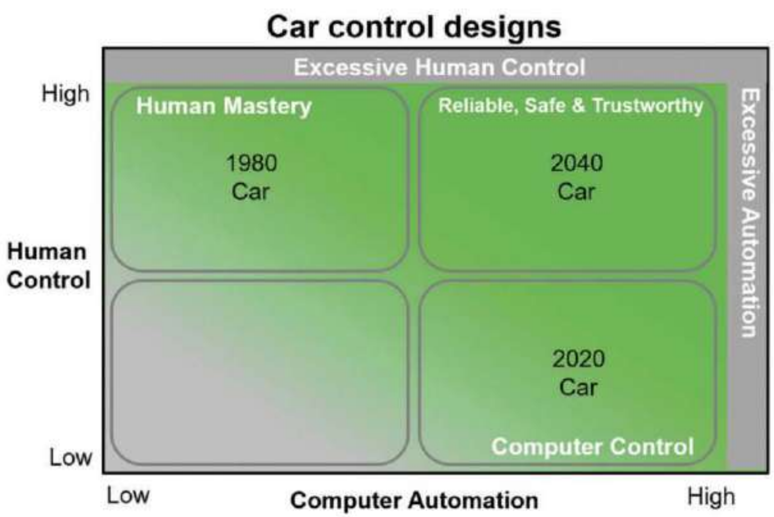


Figure 5. 1980 cars had modest computer automation with high levels of human control, while 2020 self-driving cars have high computer automation, but inadequate human control. Achieving Reliable, Safe & Trustworthy (RST) self-driving cars by 2040 could be done with high computer automation and high human control, while avoiding the dangers of excessive automation or excessive human control.

Patient Controlled Analgesia (PCA) device

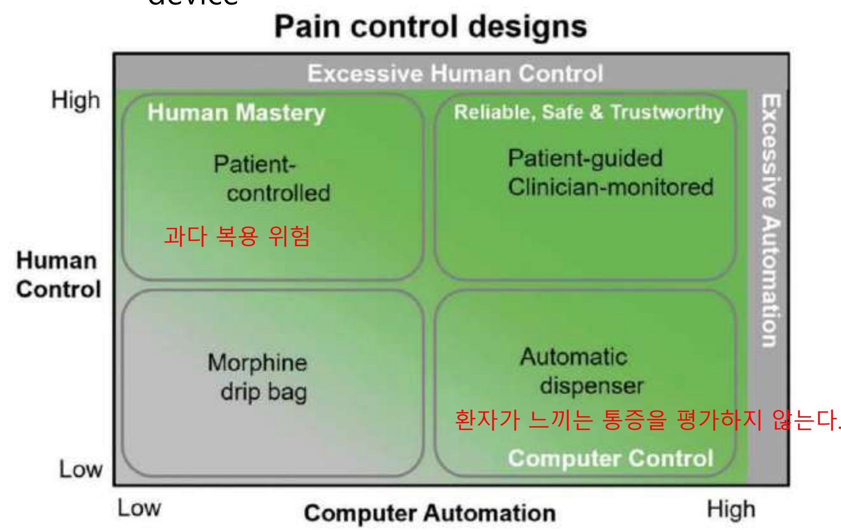
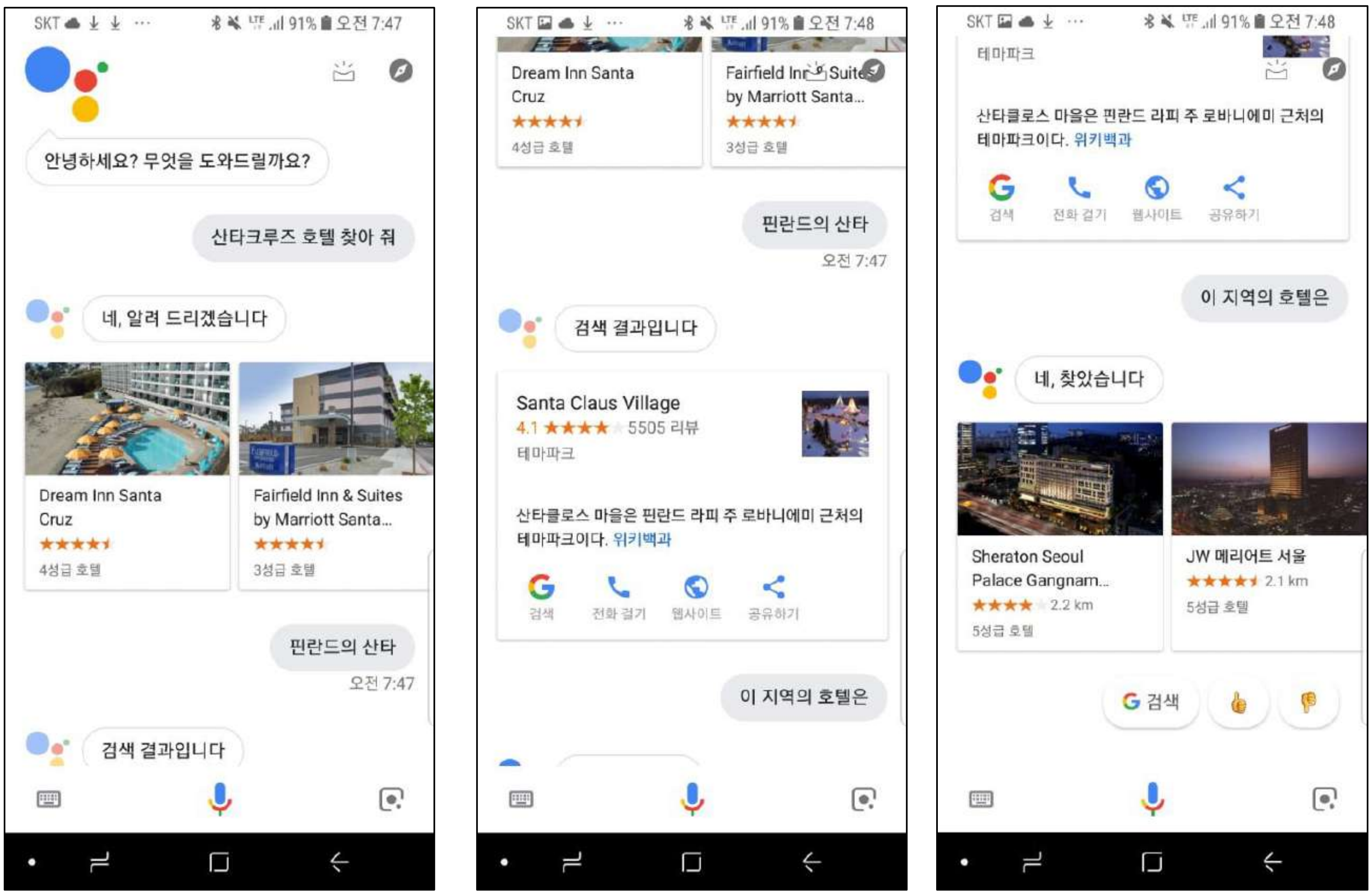


Figure 6. Four approaches to pain control designs.

- Note that these are not so new
- Goes well with traditional HCI principles
- Really just another reiteration



Interaction model for spelling error

Interaction model for contextual handling (note that this is not really an error)

…

AI and UX

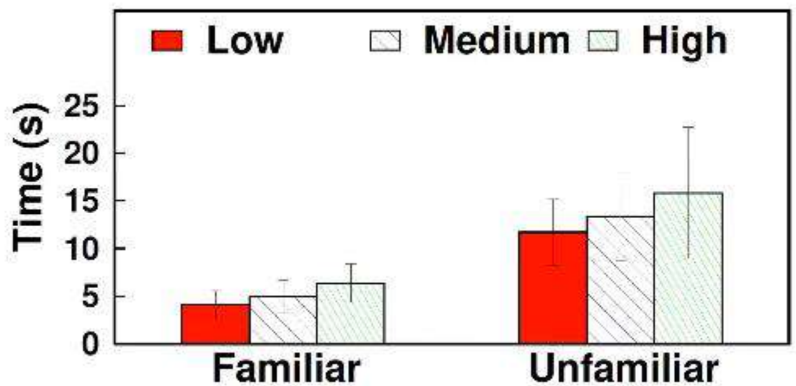
서울대학교 이영기 교수님
비디오 자료

Humans: How Fast and Accurate?

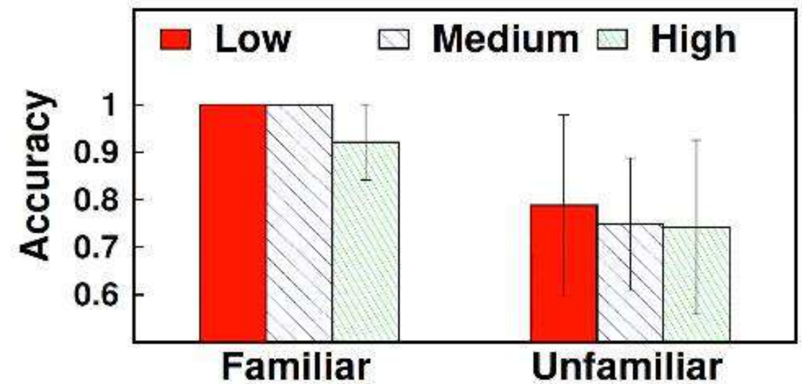
Difficult due to large number of faces and their small sizes

User Study Results

- 25 participants (age 24-32), 15 images per participant
- Crowdedness level: low/medium/high



Latency increases to **16 seconds**



Accuracy drops to **75%**

Humans are vulnerable to cognitive overload

EagleEye: Person Identification AR Service

Research Question 1:
How Can We Build Fast & Accurate Identification System with AI?

Research Question 2:
How Can We Design Intuitive User Interface for EagleEye?

EagleEye: Limitations

80% accuracy, 1 second delay, flickering

Can we design an
effective user interface to overcome
the limitations of EagleEye?

AI UI/UX Design Principles

- **Affordance**
  Clear connection between a user interface and its functional properties
- **Error Tolerance**
  Show how the system is prone to error
- **Cognitive Guidance**
  Enable the user to focus on the actual task
  and reduce cognitive overhead needed to interact with the application.

Saleema Amershi, et al. Guidelines for Human-AI Interaction. CHI 2019.
Liao, Q. V., et al. *Questioning the AI: Informing Design Practices for Explainable AI User Experiences*. CHI 2020.

How do we apply these principles for EagleEye?

Affordance

- Lack of clarity in what the bounding boxes are showing
  - Which one is the target?
  - What is the accuracy?

Baseline

Affordance

- Show **only the Top-1 result** indicating the target person
- Display the match score (0-100)
  - How is match score calculated? output feature distance in a normalized value

Error Tolerance

- How can we tolerate delayed and inaccurate bounding boxes?

$t = T$ $\quad$ $t = T + 1$ $\quad$ $t = T + 2$

Examples of inaccurate results

Error Tolerance

- Filter out jumping bounding boxes in consecutive frames
- Scope the results that are in doubt with **pulse boxes**

Cognitive Guidance

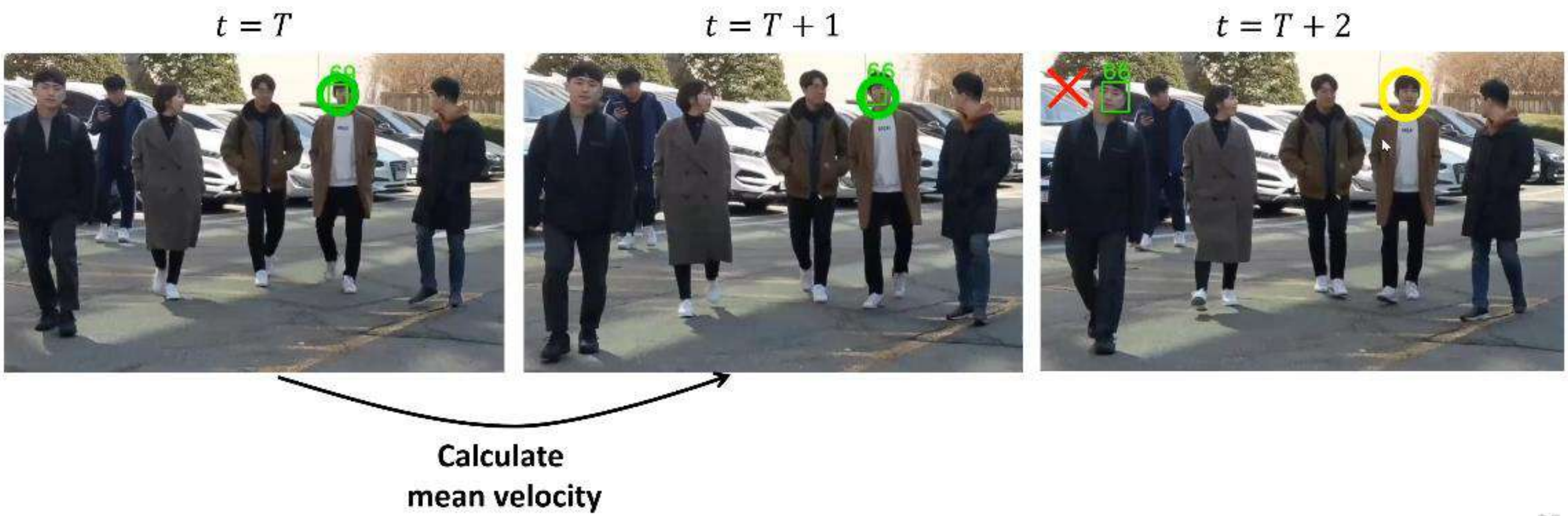
- How can we guide users to focus on tasks under ambiguous results suggested by AI system?

$t = T$ $\quad$ $t = T + 1$ $\quad$ $t = T + 2$

Requires User Attention

Cognitive Guidance

- Calculate and show the estimate location using the **temporal information**
- Color the pulse box in **a different color** for user attention



Preliminary User Study

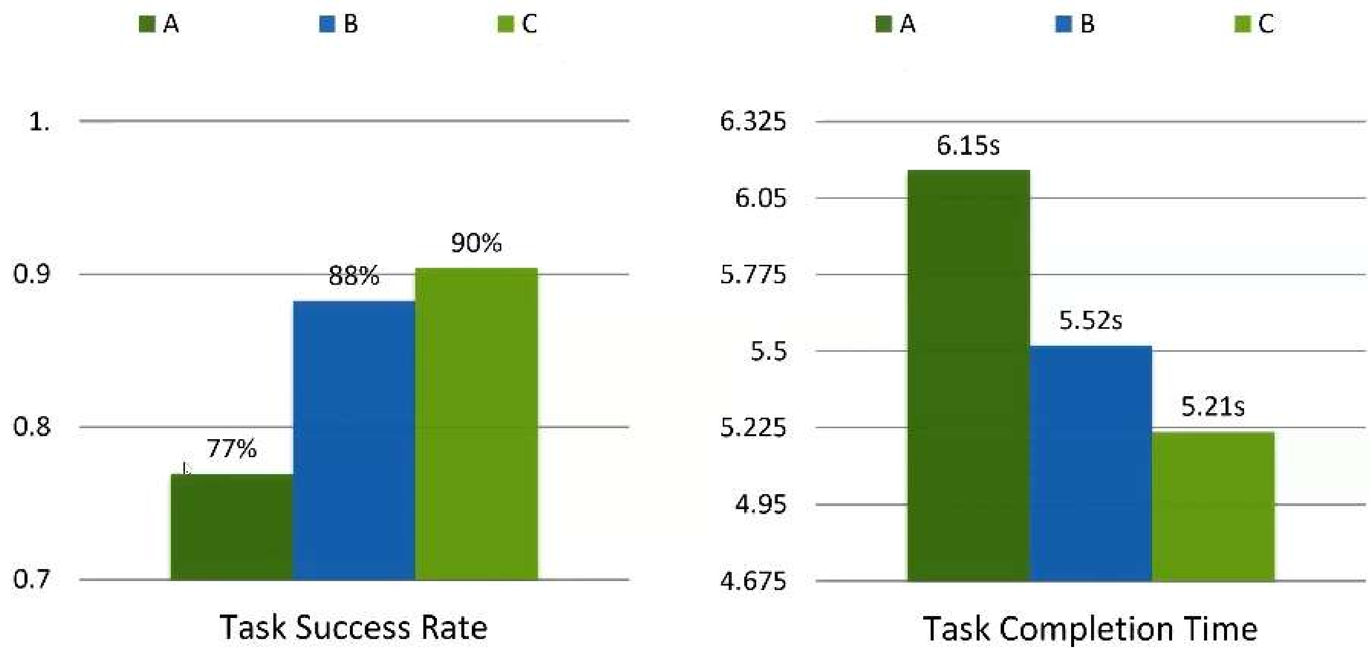
- 15 participants (age 20s), 12 videos per participant
- Comparison group (4 videos each)

A. Baseline EagleEye

B. Affordance

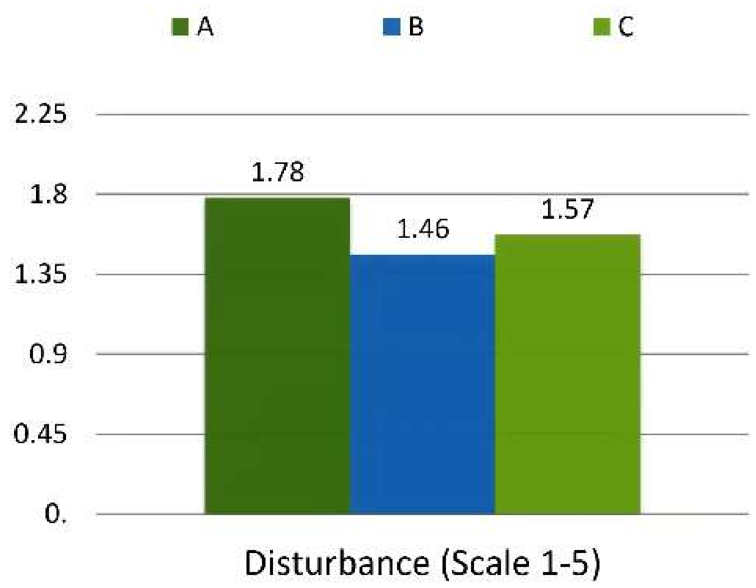
C. Affordance + Error Tolerance + Cognitive Guidance

Results: Task Performance

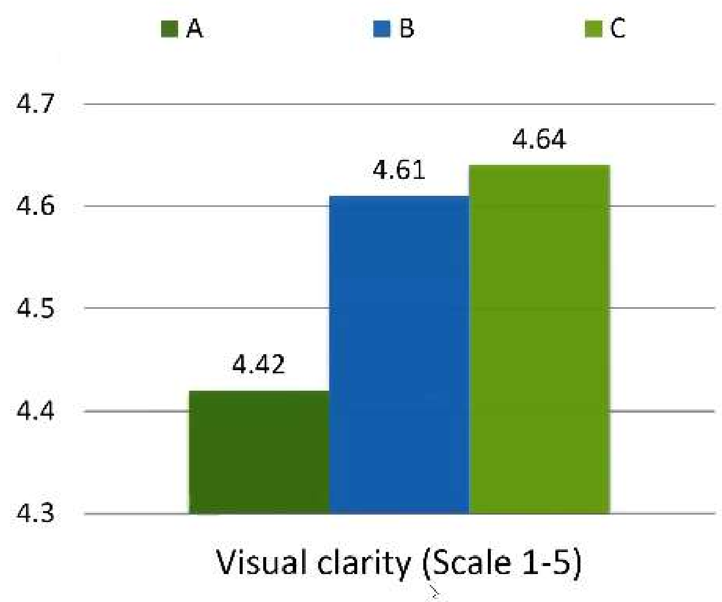


Effective UI complements system performance

Results: User Satisfaction



(1 – No disturbance at all, 5 – Very Disturbing)



(1 – Visualization is ambiguous, 5 – Visualization is very clear)

More study is needed for user satisfaction

Discussions

- Our initial findings
  - A well-designed UI can help the user perform the task better
  - A well-designed UI can improve user-perceived AI system performance
- Future work
  - Explore a more comprehensive set of UI design principles
  - Explore how much latency/accuracy problems can be overcome by better UI
  - Design XR-specific AI UI/UX principles

Concluding Remarks for HCI+AI

- Emerging AI services (especially XR apps) have tight user + system requirements for seamless user experience
- Designing an accurate and fast AI system is difficult
  - Low accuracy in in-the-wild settings
  - High computational complexity on resource-constrained consumer devices
- We can't just wait for AI algorithms and systems to be perfect ...
- Comprehensive HAI research should be done to improve usability of emerging AI-enabled systems

UX Design Innovation: Challenges for Working with Machine Learning as a Design Material

https://dl.acm.org/citation.cfm?id=3025739

- It is no longer enough for UX designers to only improve user experience by paying attention to usability, utility, and interaction aesthetics.
- Instead, the best user experiences may come from services that automatically personalize their offers to the user and context, and systems that leverage more detailed understandings of people and the world in order to provide new value.

Issues: Designing for “Intelligent” Interactive Systems

- Unlike heuristic driven systems, ML is very different from human intelligence. It applies statistical methods to produce output that can be difficult to explain, and that make seemingly strange errors
  - UX designers may therefore struggle to make designs that bridge the ML and human perspective .. But that is difficult without deep understanding of AI
  - They lack the formal training in AI and its connection to HCI
  - No clear understanding of what it can or cannot do … (notion of magic)

- They have tools for prototyping services and simulate the behavior of an app, but they have nothing that helps them quickly prototype and understand the UX impact of false negative and false positive responses from a ML service

- Intelligent systems sometimes encourage unrealistic expectations, leading to some reluctance to use them in complex or sensitive contexts
  - They may inadvertently display an inability to understand the intent behind users' behavior, which results in "intelligent" features being perceived as useless and unintuitive

- ML clearly demands a new type of prototyping, one that does not yet exist. We believe that there are a number of reasons for this.
  - First, "learning" implies that the system and data will change over time. Designers are not accustomed to designing a form for data that is dynamic at a large scale.
  - Second, the way in which ML and designers treat data is quite different from each other. Designers mostly visualize data and look for correlations and patterns that "make sense;" that fit with their understanding of how the world should and does work. ML in contrast finds machine-recognizable correlations and patterns in data.

Consider the interplay between ML statistical intelligence and human common sense intelligence

- The statistical intelligence displayed by ML may result in a very different interpretation of the same data than common sense human intelligence.
- So that, for example, we might better anticipate, mitigate or account for ML statistical errors, such as false positives or negatives, or design for perceptive qualities.
- Needs deeper understanding/training of AI technology

- Q: What is the most natural way to recover from AI errors?
  - AI is not so special in some sense, Humans make mistakes too!