HUMAN-AI INTERACTION

Eunsuk Kang

Required reading:

Building Intelligent Systems by Geoff Hulten (2018), Chapter 8.

Guidelines for Human-AI Interaction. Saleema Amershi, et al., in CHI 2019.

Optional reading:

Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-user Expectations of AI Systems. Kocielnik, et al., in CHI 2019

LEARNING GOALS

- Understand the risks of poor interaction design
- Understand the challenges behind designing human-AI interactions
- Understand the basic elements of user interaction design
- Consider design considerations for AI-based systems
    - Modes of interaction: Automate or augment?
    - Mental model: User understanding of what AI is doing
    - Dealing with errors: Guide user towards prevention & recovery
    - Feedback and control: Align user feedback with AI improvement

WHAT'S COMING NEXT

Fundamentals of Engineering AI-Enabled Systems

Holistic system view: AI and non-AI components, pipelines, stakeholders, environment interactions, feedback loops

Requirements:
System and model goals
User requirements
Environment assumptions
Quality beyond accuracy
Measurement
Risk analysis
Planning for mistakes

Architecture + design:
Modeling tradeoffs
Deployment architecture
Data science pipelines
Telemetry, monitoring
Anticipating evolution
Big data processing
Human-AI design

Quality assurance:
Model testing
Data quality
QA automation
Testing in production
Infrastructure quality
Debugging

Operations:
Continuous deployment
Contin. experimentation
Configuration mgmt.
Monitoring
Versioning
Big data
DevOps, MLOps

Teams and process: Data science vs software eng. workflows, interdisciplinary teams, collaboration points, technical debt

Responsible AI Engineering

Provenance, versioning, reproducibility

Safety

Security and privacy

Fairness

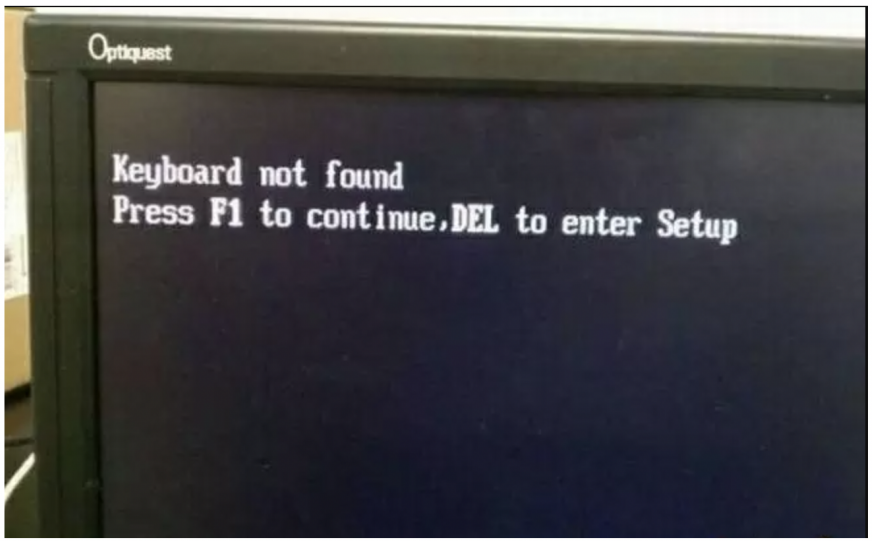
Interpretability and explainability

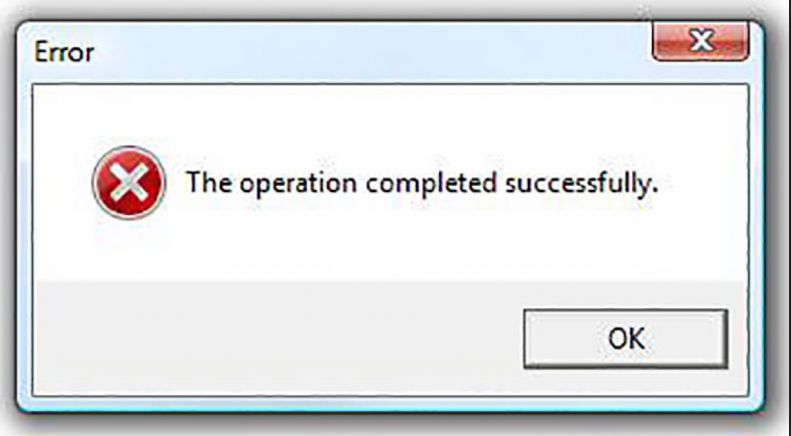
Transparency and trust

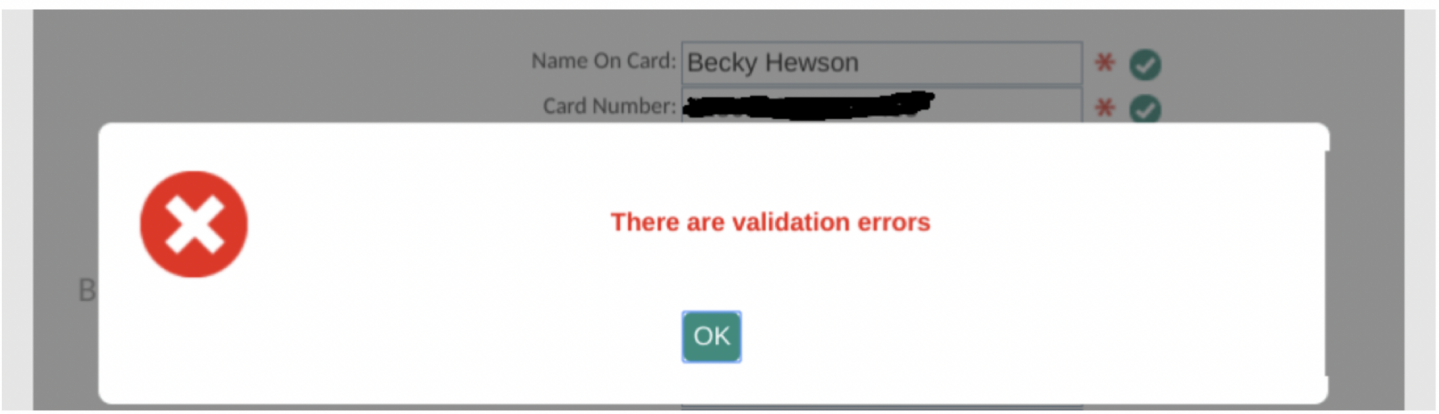
Ethics, governance, regulation, compliance, organizational culture

RISKS OF POOR INTERACTION DESIGN

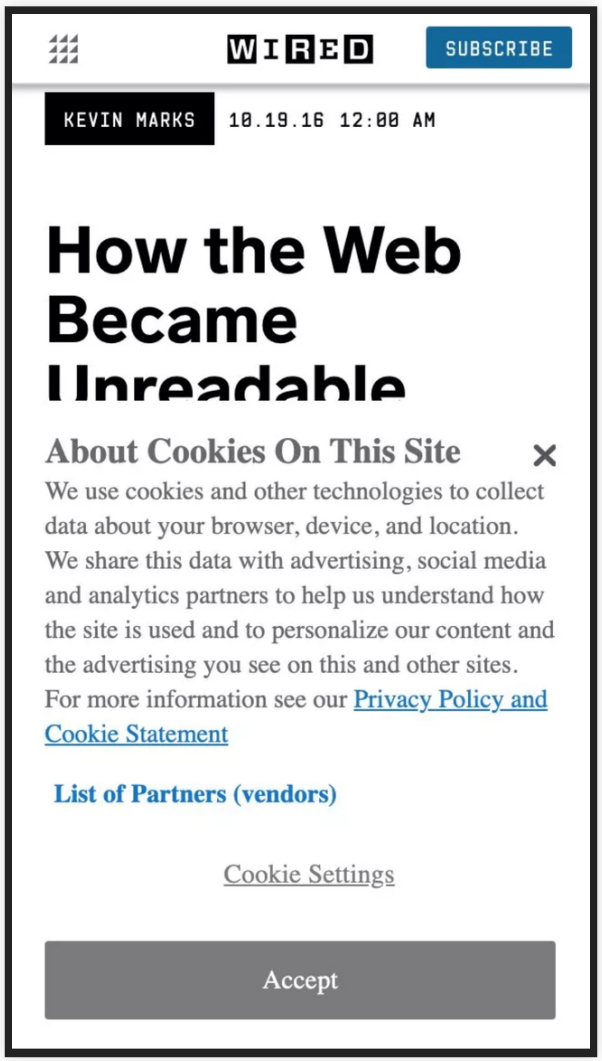
POOR INTERACTION DESIGN CONFUSES USERS







POOR INTERACTION DESIGN ANNOYS USERS



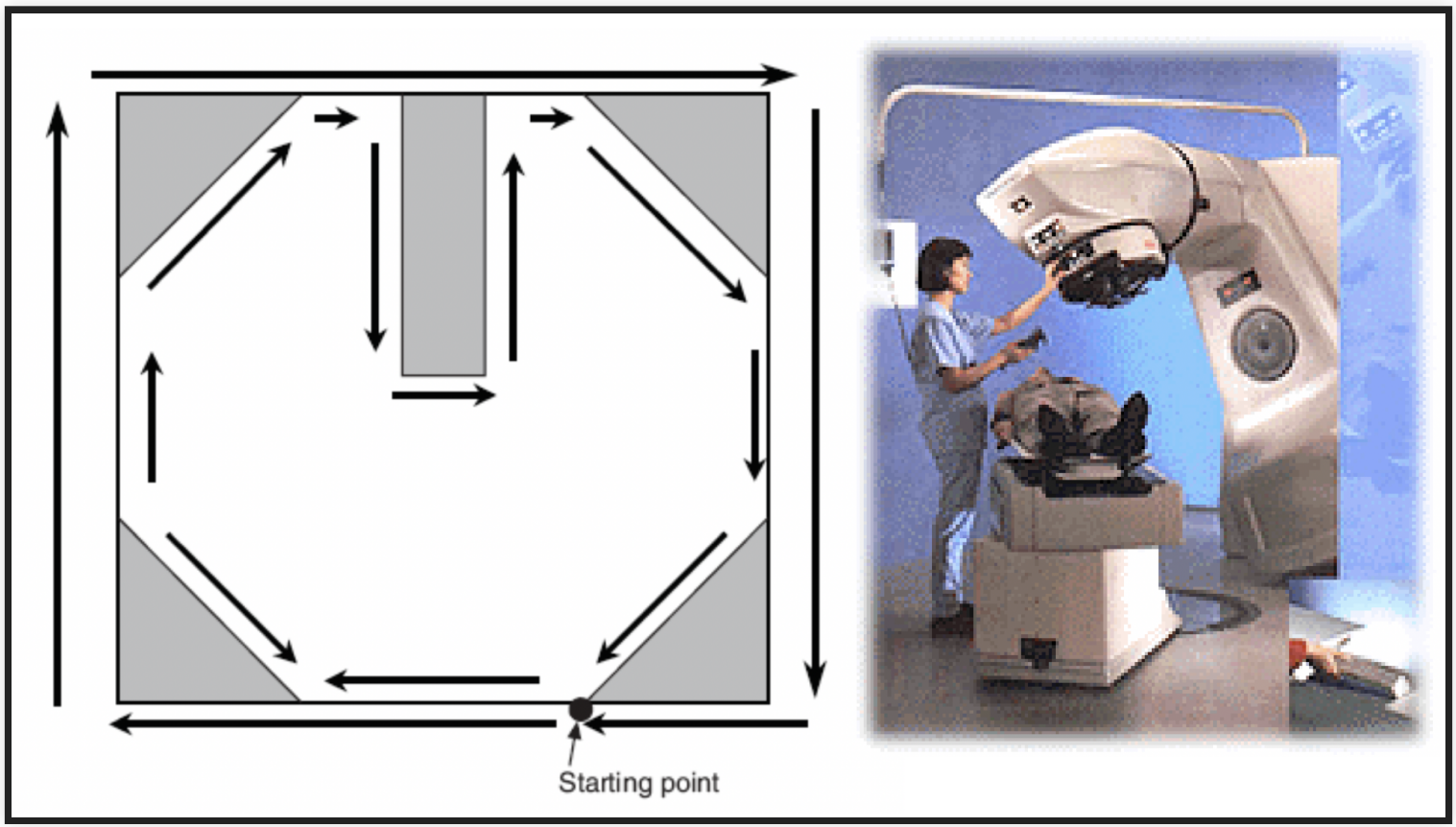
POOR INTERACTION DESIGN CAUSES HARM

Alexa recorded a woman's private conversation and sent it to a random contact

Kyle Wiggers @Kyle_L_Wiggers May 24, 2018 7:38 AM

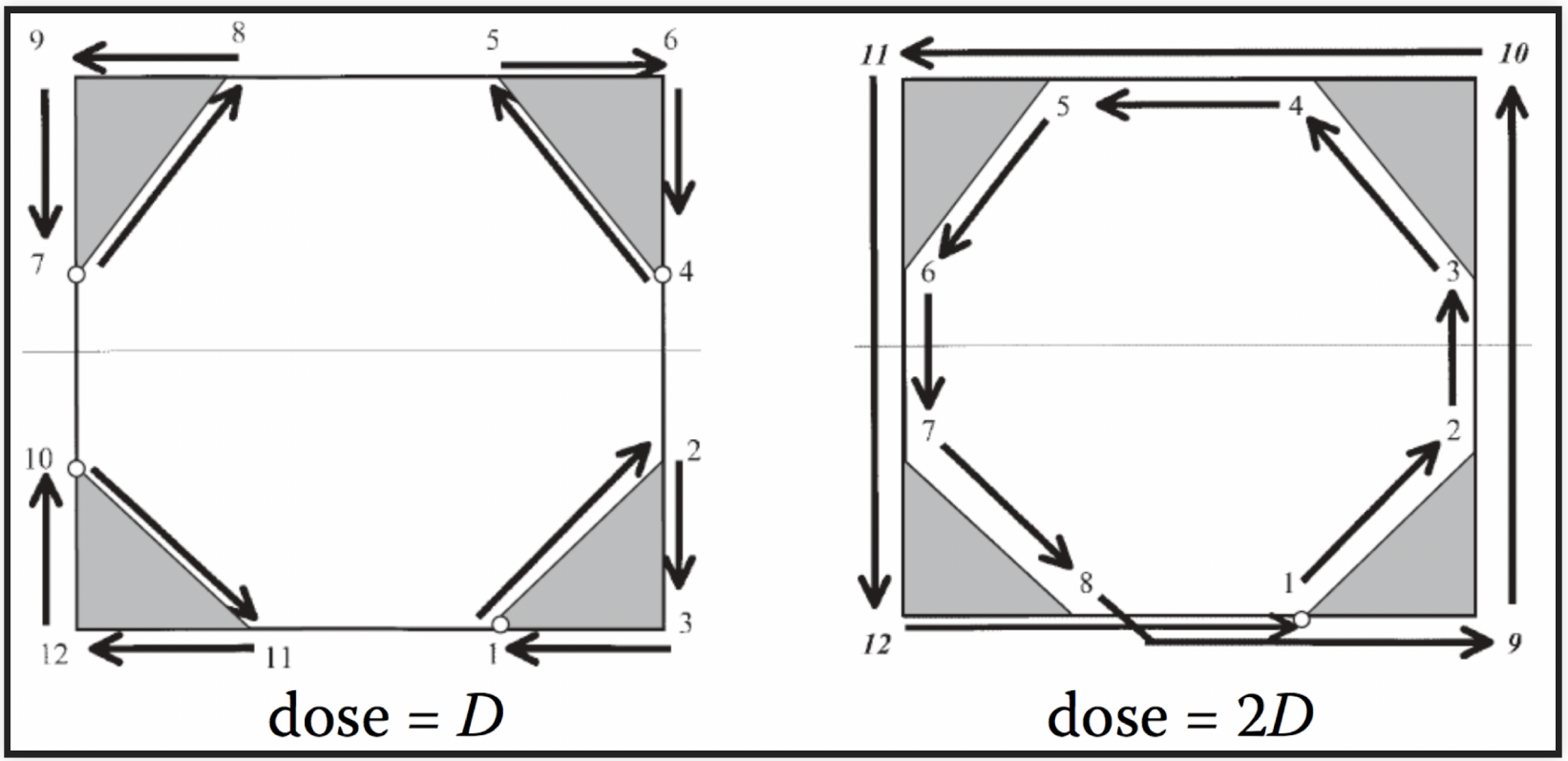
POOR INTERACTION DESIGN CAUSES HARM

POOR INTERACTION DESIGN CAUSES HARM



- Radiation therapy system at Panama City public hospital (2001)
  - Therapist draws block shapes to determine treatment area
  - Software computes final radiation settings

POOR INTERACTION DESIGN CAUSES HARM



- Same shape drawn in different order, double the radiation dose
- 28 patients overdosed; 8 dead
    - Therapists charged with 2nd degree murder (but are they really to blame?)

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!
- Poor interaction design can:

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!
- Poor interaction design can:
    - Cause confusion or misunderstanding

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!
- Poor interaction design can:
    - Cause confusion or misunderstanding
    - Prevent the user from effectively performing their task

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!
- Poor interaction design can:
    - Cause confusion or misunderstanding
    - Prevent the user from effectively performing their task
    - Increase mental and physical burden

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!
- Poor interaction design can:
    - Cause confusion or misunderstanding
    - Prevent the user from effectively performing their task
    - Increase mental and physical burden
    - Drive users away from the product

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!
- Poor interaction design can:
    - Cause confusion or misunderstanding
    - Prevent the user from effectively performing their task
    - Increase mental and physical burden
    - Drive users away from the product
    - Contribute to security or privacy issues

RISKS OF POOR INTERACTION DESIGN

- Interaction design is not just about visual presentation!
- Poor interaction design can:
    - Cause confusion or misunderstanding
    - Prevent the user from effectively performing their task
    - Increase mental and physical burden
    - Drive users away from the product
    - Contribute to security or privacy issues
    - Cause physical (injuries, deaths) and societal harms (bias, misrepresentation)

USABILITY CONCEPTS

(This will be a brief tour to a complex subject. If you are interested, consider taking 05-318/618: Human-AI Interaction)

DIMENSIONS OF USABILITY

https://www.nngroup.com/articles/usability-101-introduction-to-usability/

DIMENSIONS OF USABILITY

- Learnability: How easy is it for users to accomplish tasks the first time?

https://www.nngroup.com/articles/usability-101-introduction-to-usability/

DIMENSIONS OF USABILITY

- Learnability: How easy is it for users to accomplish tasks the first time?
- Efficiency: After learning, how efficiently can users perform the tasks?

https://www.nngroup.com/articles/usability-101-introduction-to-usability/

DIMENSIONS OF USABILITY

- Learnability: How easy is it for users to accomplish tasks the first time?
- Efficiency: After learning, how efficiently can users perform the tasks?
- Memorability: Can users remember to perform the tasks after a period of not using the system?

https://www.nngroup.com/articles/usability-101-introduction-to-usability/

DIMENSIONS OF USABILITY

- Learnability: How easy is it for users to accomplish tasks the first time?
- Efficiency: After learning, how efficiently can users perform the tasks?
- Memorability: Can users remember to perform the tasks after a period of not using the system?
- Errors: How often do users make errors, how severe are these errors, and how easily can they recover from the errors?

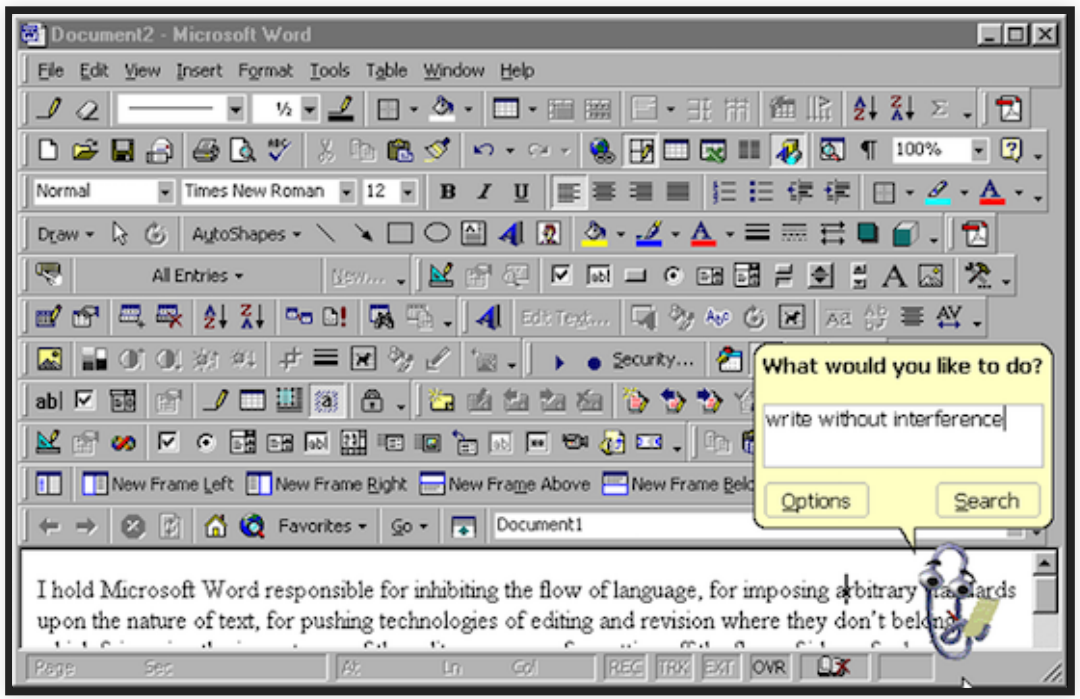
https://www.nngroup.com/articles/usability-101-introduction-to-usability/

DIMENSIONS OF USABILITY

- Learnability: How easy is it for users to accomplish tasks the first time?
- Efficiency: After learning, how efficiently can users perform the tasks?
- Memorability: Can users remember to perform the tasks after a period of not using the system?
- Errors: How often do users make errors, how severe are these errors, and how easily can they recover from the errors?
- Satisfaction: How pleasant is it to use the design?

https://www.nngroup.com/articles/usability-101-introduction-to-usability/

INTERACTION COST



- Mental and physical effort needed to perform a desired task
  - Task memorization & recall, context switch, track system state
  - Reading, scrolling, clicking, typing, waiting for UI changes
- **Goal of usable design**: Minimize interaction cost while allowing users to perform their tasks

USABILITY & AI

- AI has potential to greatly reduce interaction costs
    - Automate tasks through personalization & predictions
- But also introduces new usability challenges
    - **Q. What's new or hard about AI-based systems?**

USABILITY & AI

- AI has potential to greatly reduce interaction costs
    - Automate tasks through personalization & predictions
- But also introduces new usability challenges
    - **Unpredictability**: AI makes mistakes, sometimes unexpectedly
    - **Opaqueness**: User has difficulty understanding how system works
    - **Evolution**: AI behavior changes over time, surprising users

DESIGN CONSIDERATIONS FOR AI

- **Modes of interaction**: Automate or augment?
- **Mental model**: User understanding of what AI is doing
- **Dealing with errors**: Guide user towards prevention & recovery
- **Feedback and control**: Align user feedback with AI improvement

MODES OF INTERACTION

MODES OF INTERACTION

MODES OF INTERACTION

- Automate: Take action on user's behalf

MODES OF INTERACTION

- Automate: Take action on user's behalf
- Augment: Provide options or additional information
    - Prompt: Ask the user if an action should be taken
    - Organize: Display a set of items in an order
    - Annotate: Add information to a display

MODES OF INTERACTION

- Automate: Take action on user's behalf
- Augment: Provide options or additional information
    - Prompt: Ask the user if an action should be taken
    - Organize: Display a set of items in an order
    - Annotate: Add information to a display
- Hybrid of above

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:
    - User lacks knowledge/ability to perform the task (e.g., prediction)

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:
    - User lacks knowledge/ability to perform the task (e.g., prediction)
    - Boring, repetitive, dangerous tasks

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:
    - User lacks knowledge/ability to perform the task (e.g., prediction)
    - Boring, repetitive, dangerous tasks
    - The effect of action can be reversed

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:
    - User lacks knowledge/ability to perform the task (e.g., prediction)
    - Boring, repetitive, dangerous tasks
    - The effect of action can be reversed
- Augment when:

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:
    - User lacks knowledge/ability to perform the task (e.g., prediction)
    - Boring, repetitive, dangerous tasks
    - The effect of action can be reversed
- Augment when:
    - High stakes & accountability is needed

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:
    - User lacks knowledge/ability to perform the task (e.g., prediction)
    - Boring, repetitive, dangerous tasks
    - The effect of action can be reversed
- Augment when:
    - High stakes & accountability is needed
    - Difficult to communicate the user's need to AI

SELECTING APPROPRIATE MODE OF INTERACTION

- Identify the tasks that the user wants to achieve
- For each task, decide between **automate vs. augment**
- Automate when:
    - User lacks knowledge/ability to perform the task (e.g., prediction)
    - Boring, repetitive, dangerous tasks
    - The effect of action can be reversed
- Augment when:
    - High stakes & accountability is needed
    - Difficult to communicate the user's need to AI
    - User enjoys performing the task (e.g., driving)

OTHER FACTORS TO CONSIDER

OTHER FACTORS TO CONSIDER

- Forcefulness: How strongly to encourage taking an action?
  - Active: Automate action or interrupt user and ask for confirmation
  - Passive: Suggest action, but do not require immediate answer

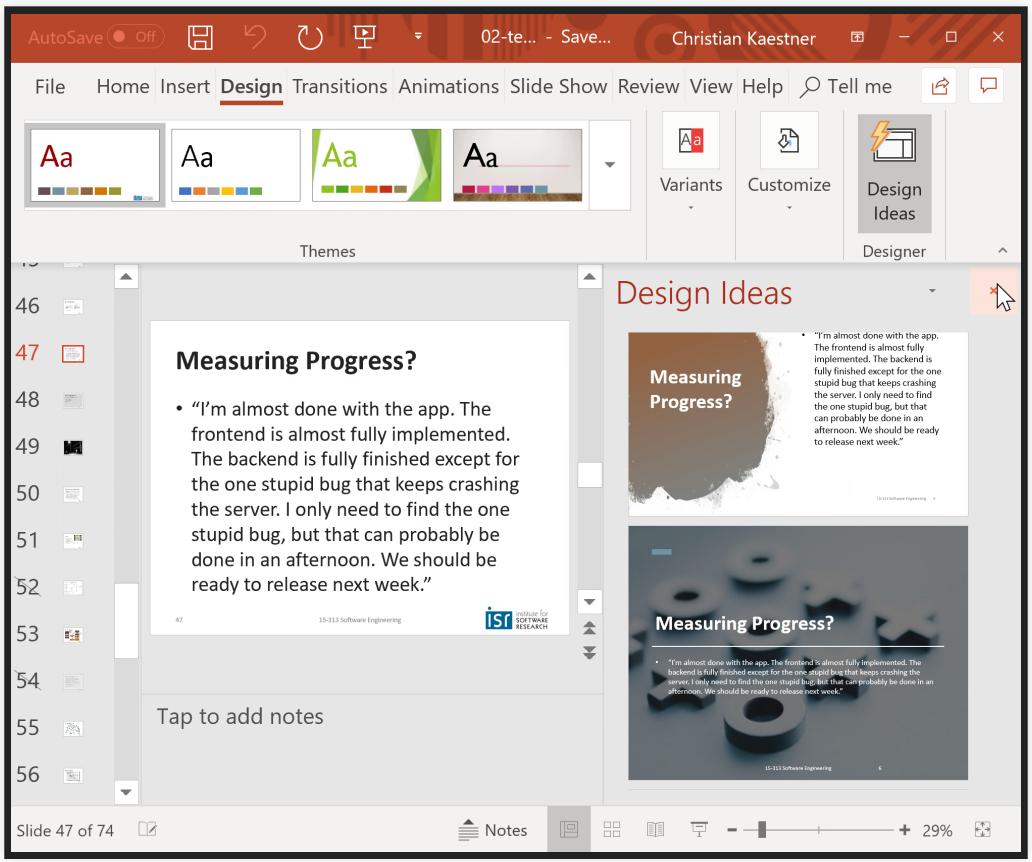
OTHER FACTORS TO CONSIDER

- Forcefulness: How strongly to encourage taking an action?
  - Active: Automate action or interrupt user and ask for confirmation
  - Passive: Suggest action, but do not require immediate answer
- Frequency: How often does interaction occur?
  - When a new prediction is available or model changes
  - Periodically (e.g., suggest action every hour)
  - Only when explicitly initiated by user

OTHER FACTORS TO CONSIDER

- Forcefulness: How strongly to encourage taking an action?
  - Active: Automate action or interrupt user and ask for confirmation
  - Passive: Suggest action, but do not require immediate answer
- Frequency: How often does interaction occur?
  - When a new prediction is available or model changes
  - Periodically (e.g., suggest action every hour)
  - Only when explicitly initiated by user
- Cost: What is the effect of a wrong prediction?
  - If cost is too high, consider augmenting rather than automating
  - If possible, provide a way to undo the action of AI

EXAMPLE: DESIGN SUGGESTIONS IN POWERPOINT



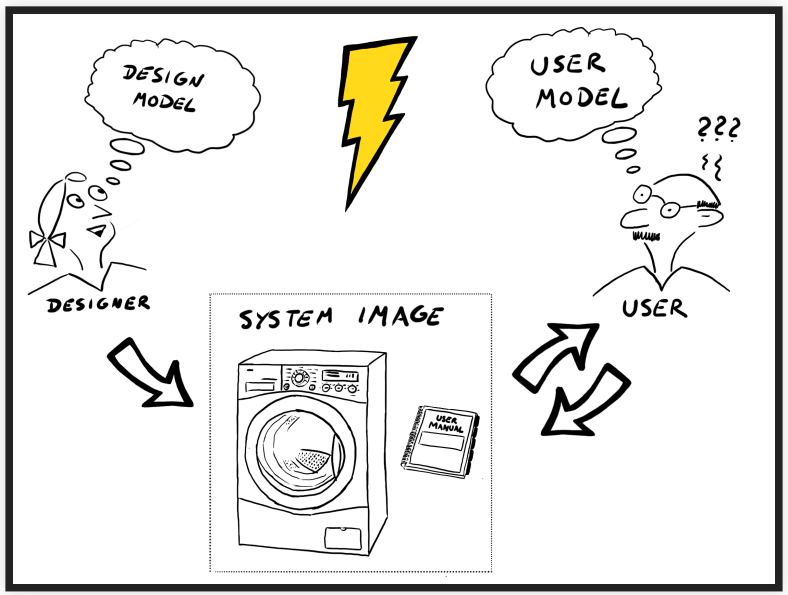
- Automate or Augment? Why?
- Forcefulness? (active vs. passive)
- Frequency?

EXAMPLE: FALL DETECTION

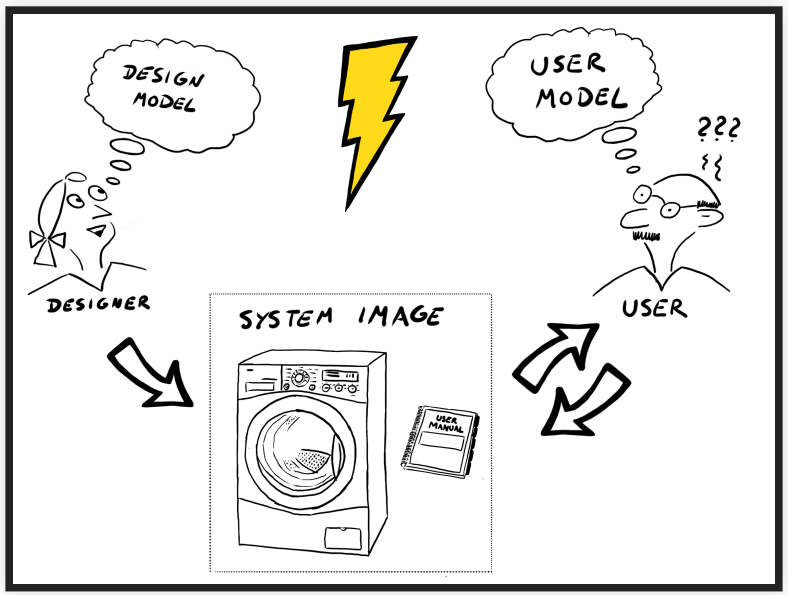
- Automate or Augment? Why?
- Forcefulness? (active vs. passive)
- Frequency?

MENTAL MODEL

MENTAL MODEL

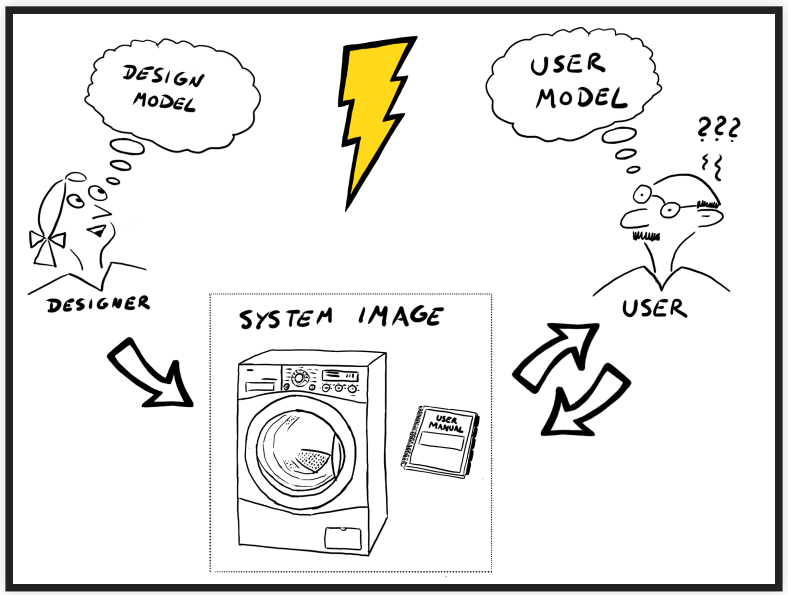


MENTAL MODEL



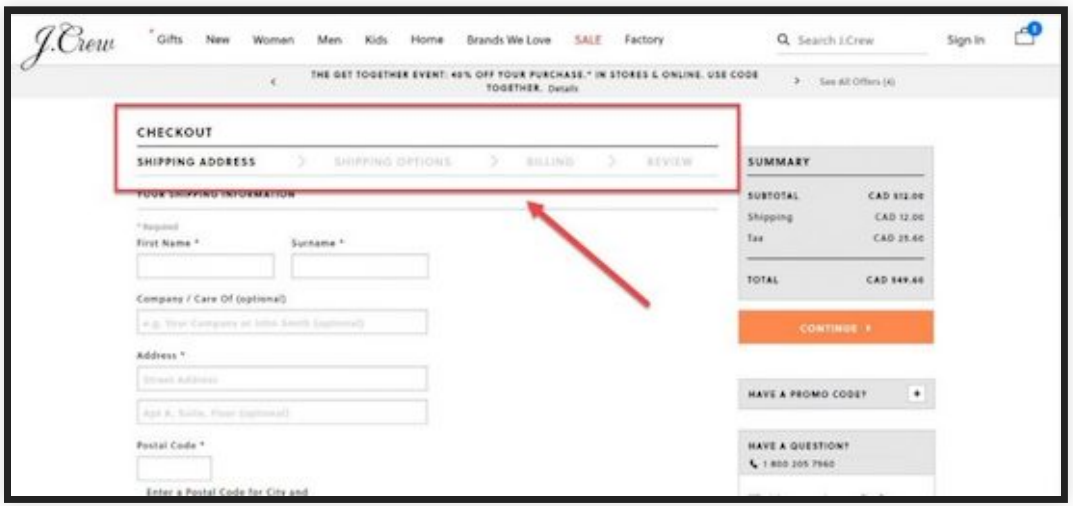
- What the user believes about the system
  - "How does the system work? How does it respond to my actions?"
  - User plans actions and reacts to system based on this mental model

MENTAL MODEL



- What the user believes about the system
  - "How does the system work? How does it respond to my actions?"
  - User plans actions and reacts to system based on this mental model
- Challenge: Aligning system with the user's mental model
  - Inherent mismatch between user's & designer's models
  - User's model may be preconceived based on prior experience
  - User's model and/or system evolves over time

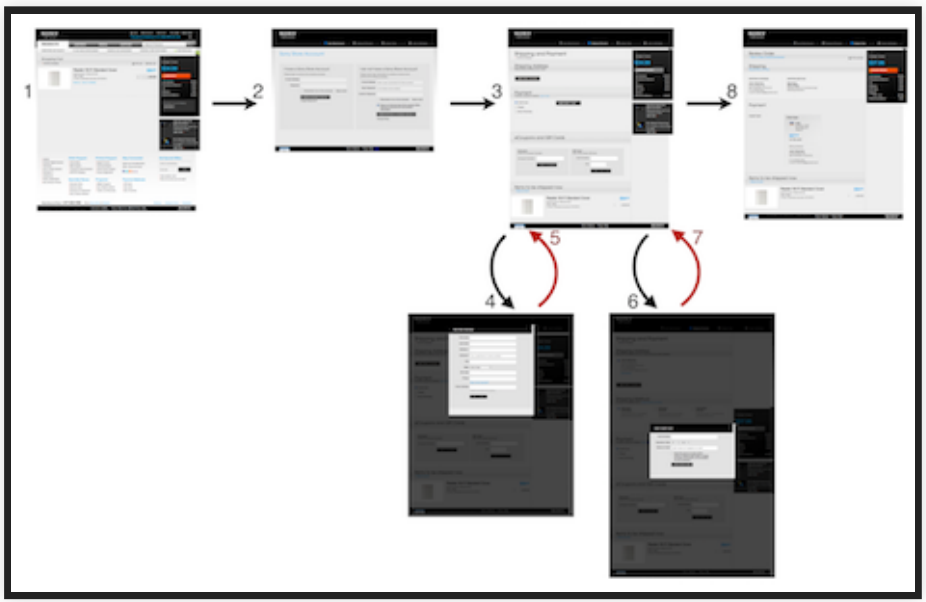
EXAMPLE: SHOPPING CART CHECKOUT



Mental model for shopping cart = A linear sequence of familiar steps

1. Browse for items
2. Add items to cart
3. Choose checkout
4. Enter shipping & billing data
5. Press Order
6. Get confirmation

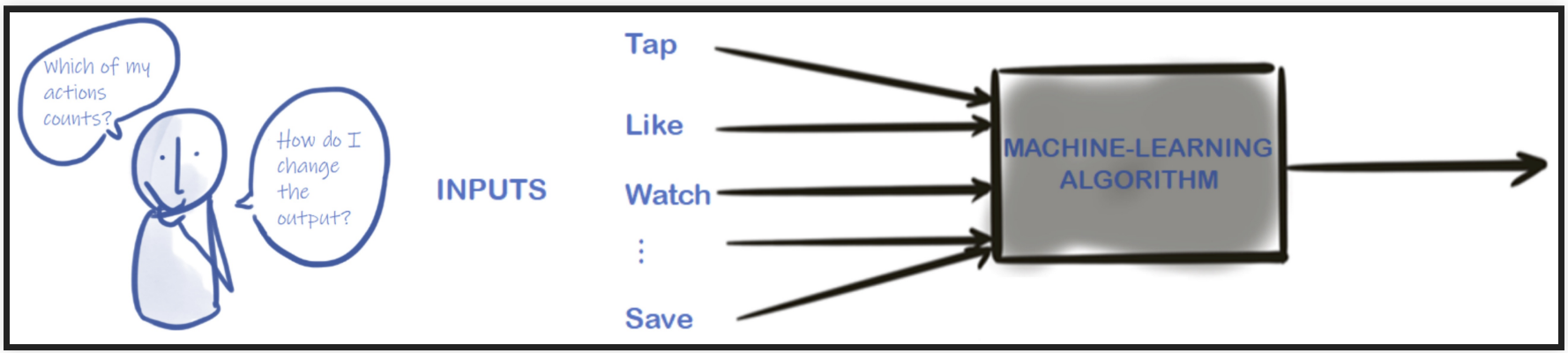
BREAKING MENTAL MODEL



- Anti-pattern: Interrupt linear flow & bring user back to a previous step
  - Create an account, open a new dialog to enter preferred address...
  - Breaks user's mental model => failure to convert into sales
- ~60% of customers abandon their shopping cart

https://baymard.com/blog/checkout-process-should-be-linear

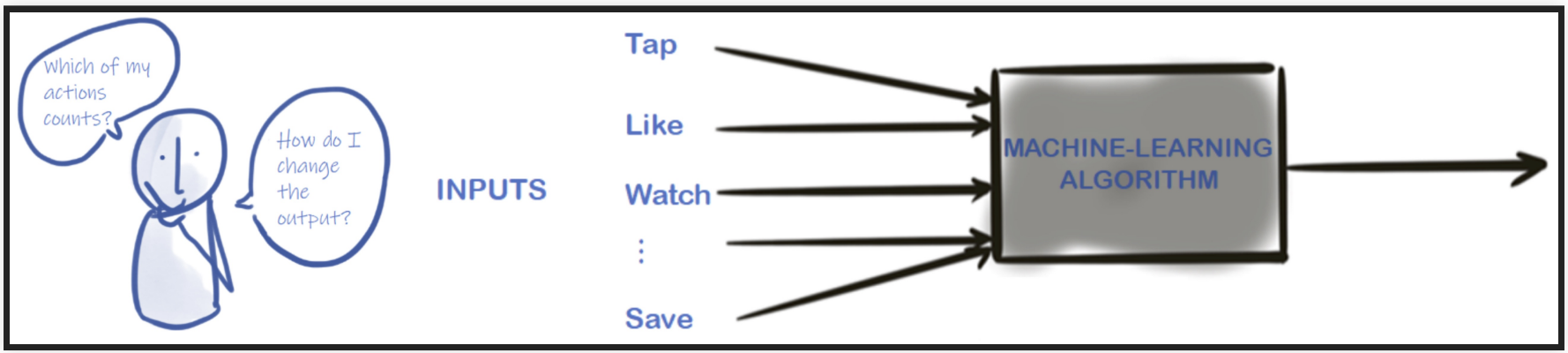
MENTAL MODEL FOR AI-BASED SYSTEMS



- User: "What is AI doing, and how do I use it?"
  - Opaqueness: Typically less transparent than traditional apps
  - AI will make mistakes, often unpredictably

https://www.nngroup.com/articles/machine-learning-ux/

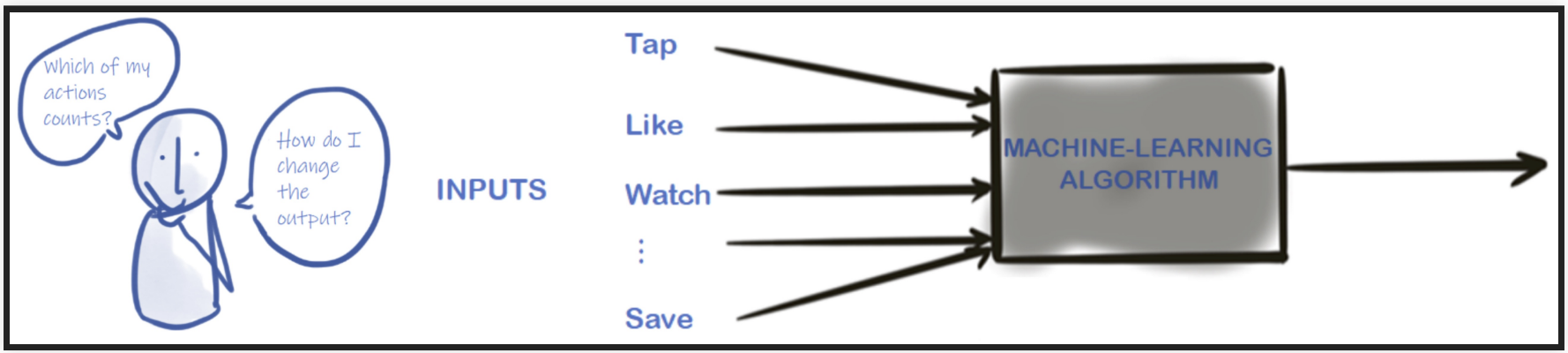
MENTAL MODEL FOR AI-BASED SYSTEMS



- User: "What is AI doing, and how do I use it?"
  - Opaqueness: Typically less transparent than traditional apps
  - AI will make mistakes, often unpredictably
- Unclear inputs: What are possible actions? Which of these actions matter? When does my action take effect?

https://www.nngroup.com/articles/machine-learning-ux/

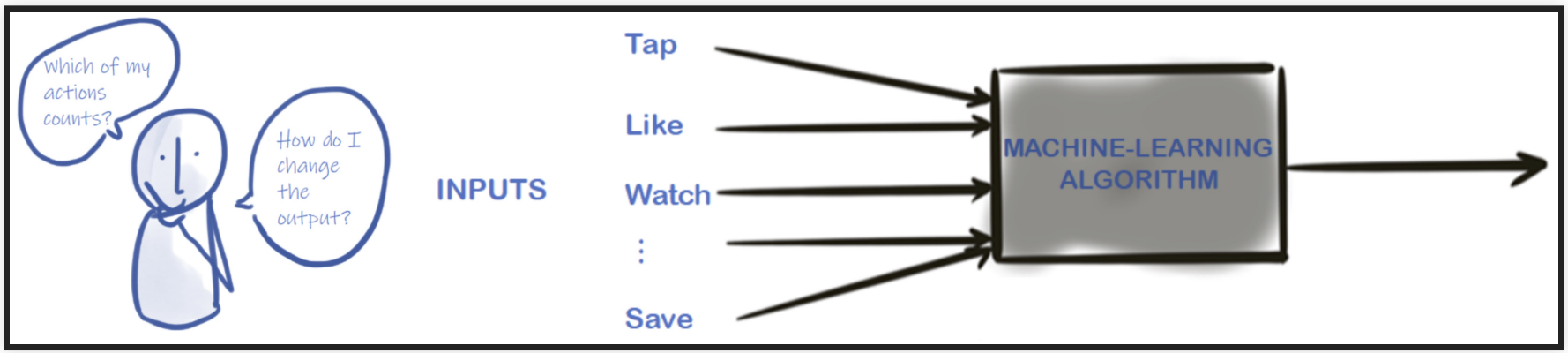
MENTAL MODEL FOR AI-BASED SYSTEMS



- User: "What is AI doing, and how do I use it?"
  - Opaqueness: Typically less transparent than traditional apps
  - AI will make mistakes, often unpredictably
- Unclear inputs: What are possible actions? Which of these actions matter? When does my action take effect?
- Lack of control over output: Why am I being given these recommendations? Why is the output displayed in this order?

https://www.nngroup.com/articles/machine-learning-ux/

MENTAL MODEL FOR AI-BASED SYSTEMS



- User: "What is AI doing, and how do I use it?"
  - Opaqueness: Typically less transparent than traditional apps
  - AI will make mistakes, often unpredictably
- Unclear inputs: What are possible actions? Which of these actions matter? When does my action take effect?
- Lack of control over output: Why am I being given these recommendations? Why is the output displayed in this order?
- Lack of trust over output: How do I know the output is correct?

https://www.nngroup.com/articles/machine-learning-ux/

MENTAL MODEL FOR VOICE ASSISTANTS?

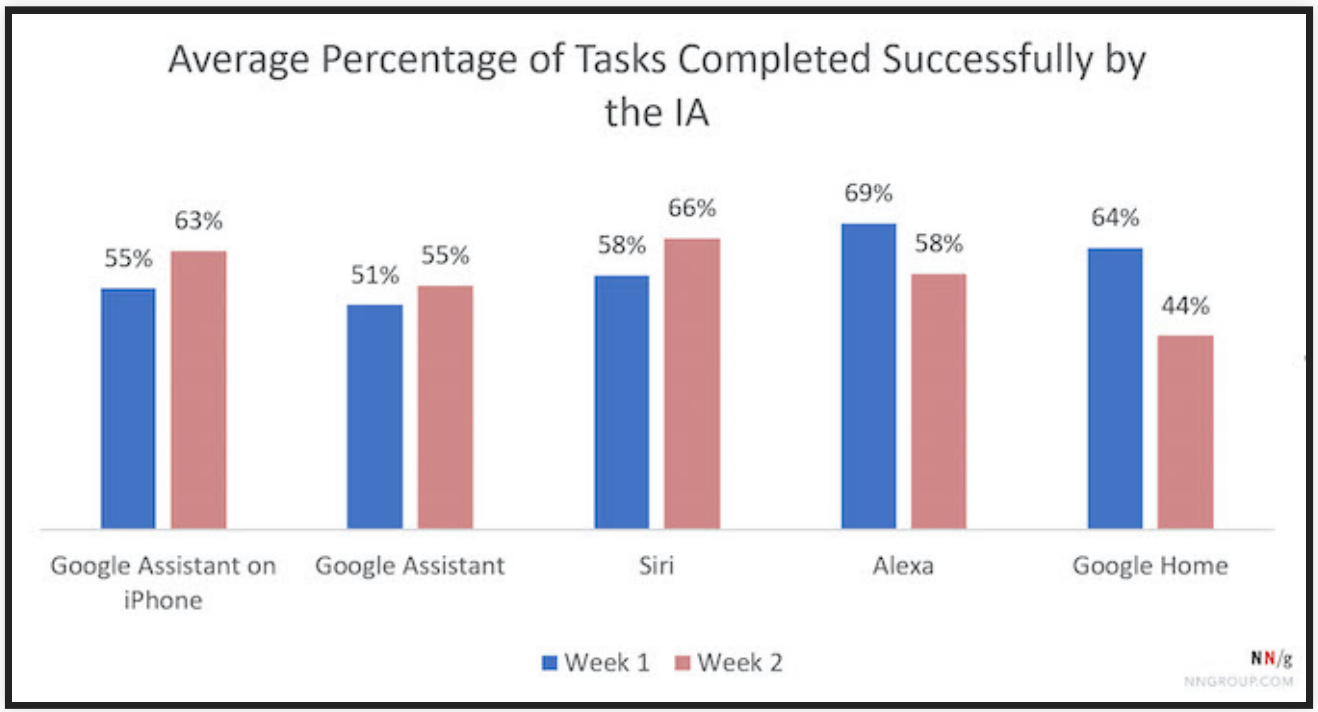
Q. Can you describe what it does? What it cannot do?

MENTAL MODEL FOR VOICE ASSISTANTS?

- Unclear, inconsistent mental model
  - An interface for other services?
  - "Handy helper"?
  - Knowledge repository? Fact-finding tool?

https://www.nngroup.com/articles/mental-model-ai-assistants/

MISALIGNMENT IN VOICE ASSISTANTS



- AI often fails to meet user expectations
  - (1) User doesn't know how to get AI to do X
  - (2) User says X, but AI can't do X well
- Users settle on simple tasks over time; small but limited improvements

https://www.nngroup.com/articles/mental-model-ai-assistants/

MISALIGNMENT IN MENTAL MODELS

> *“So, this week, I realized that I don't use my IA nearly as much as I thought I did. I do use it often. However it's very much normally the same like five things over and over again."*

- User settles on a suboptimal mental model & fails to benefit from the full capabilities of AI

https://www.nngroup.com/articles/mental-model-ai-assistants/

PRINCIPLES FOR ALIGNING MENTAL MODEL

PRINCIPLES FOR ALIGNING MENTAL MODEL

- Identify user's existing mental models
    - Find similar apps & identify common patterns
    - User interviews, walkthroughs, prototype testing

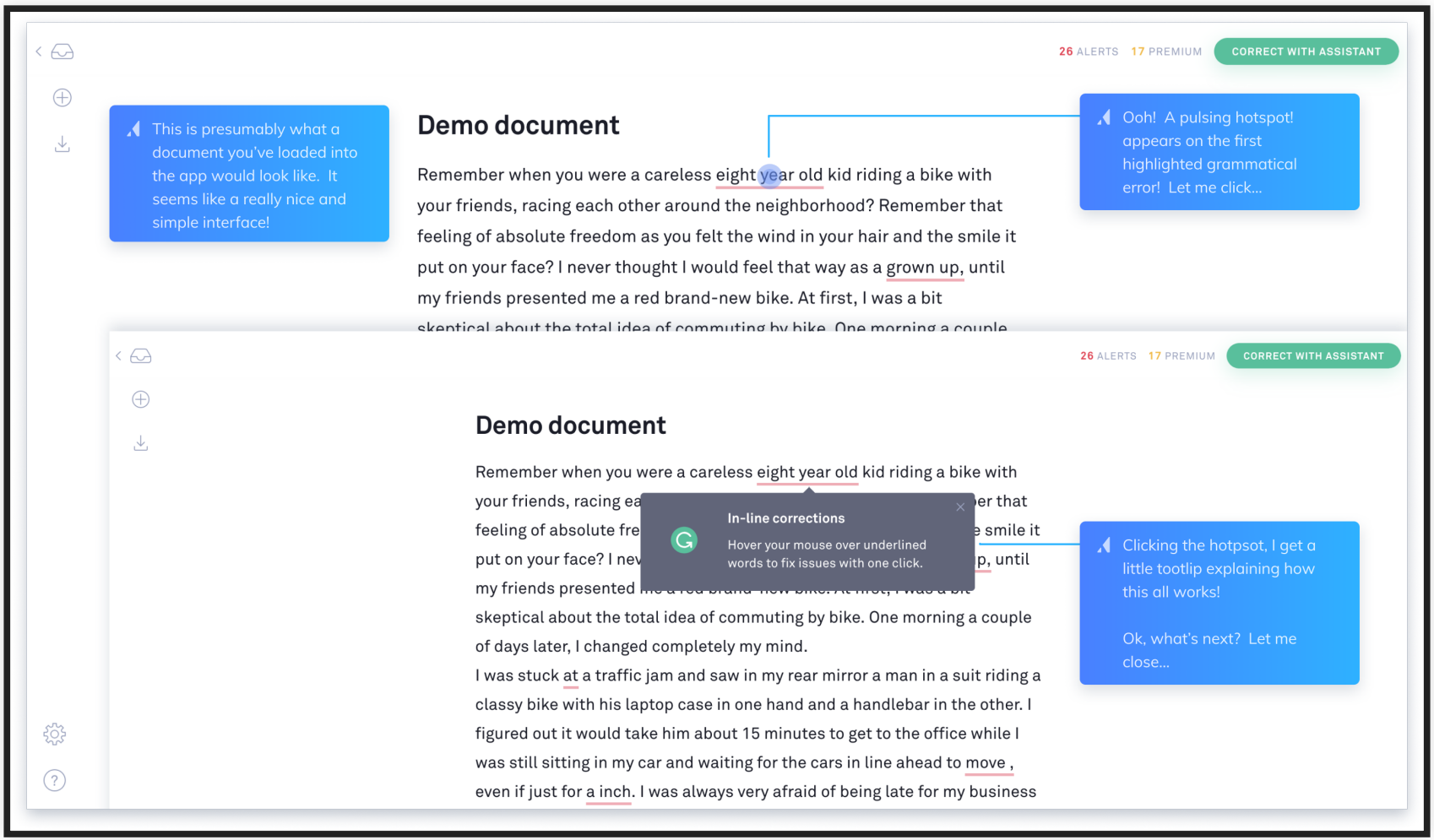
PRINCIPLES FOR ALIGNING MENTAL MODEL

- Identify user's existing mental models
  - Find similar apps & identify common patterns
  - User interviews, walkthroughs, prototype testing
- Design & evolve the system to conform to the user's model
  - Collect & analyze errors made by user
  - Identify potential mismatch vs. user's mental model

PRINCIPLES FOR ALIGNING MENTAL MODEL

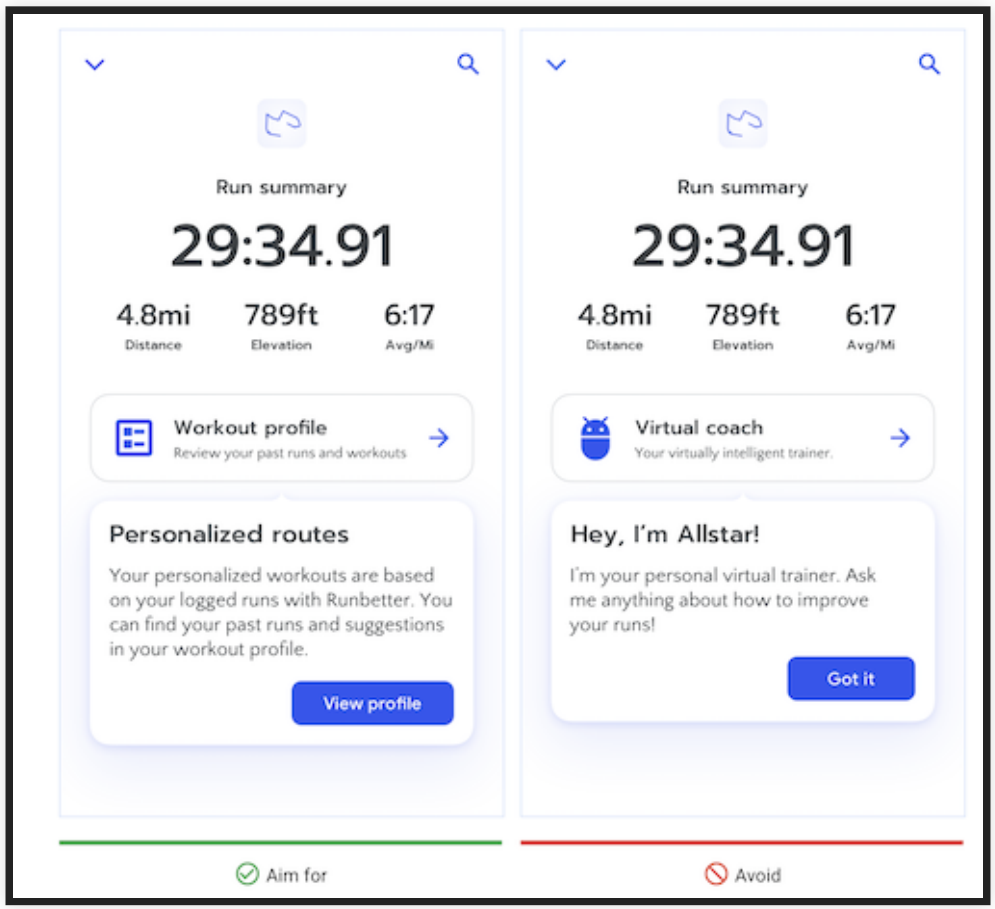
- Identify user's existing mental models
  - Find similar apps & identify common patterns
  - User interviews, walkthroughs, prototype testing
- Design & evolve the system to conform to the user's model
  - Collect & analyze errors made by user
  - Identify potential mismatch vs. user's mental model
- Improve/adjust the user's mental model
  - Set the user's expectations through onboarding
  - Increase transparency and explain decisions made by AI
  - Allow user to adjust system behavior to match their expectations

ONBOARDING: SET USER'S MENTAL MODEL



- Provide examples of how the system works

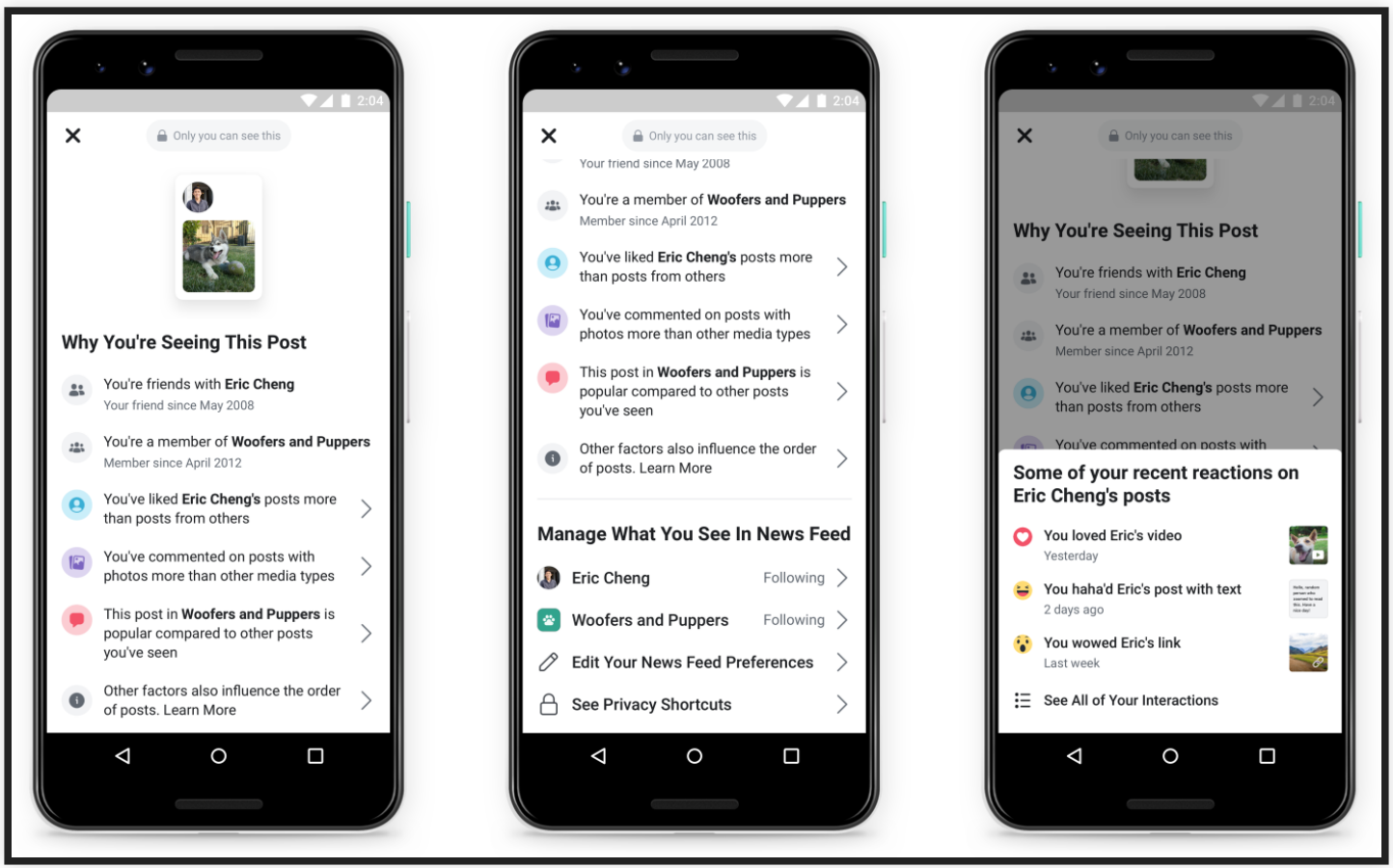
ONBOARDING: SET USER'S MENTAL MODEL



- Be explicit about what system can and cannot do

https://pair.withgoogle.com/chapter/mental-models/

TRANSPARENCY: EXPLAIN HOW DECISIONS ARE MADE



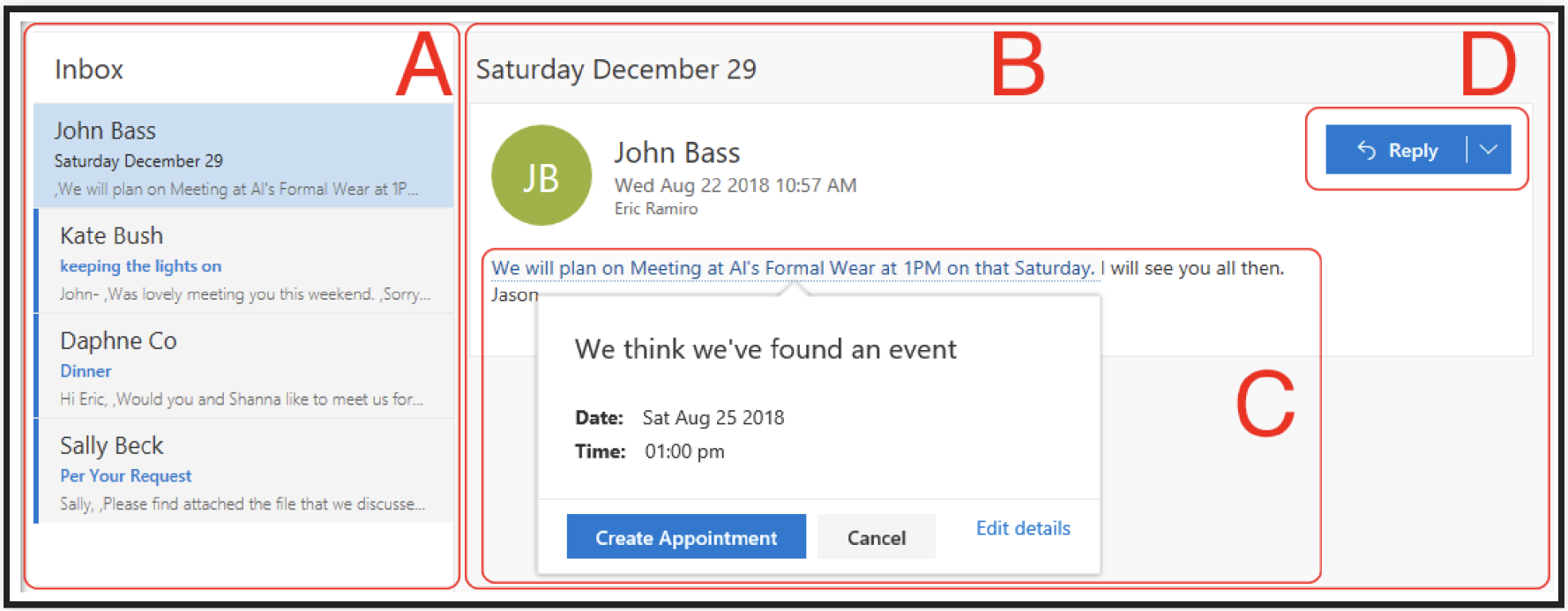
- Explain how the user's input actions influence output

DEALING WITH ERRORS

DEALING WITH ERRORS

- User errors: Mistakes made by users (e.g., click on a wrong button)
  - Lots of work in cognitive science & human factors
  - Error taxonomies, human performance modeling, task analysis, ergonomic analysis, etc.,
  - Often due to misalignment of mental models
- System errors: Failure to provide an outcome expected by the user
  - Due to mistakes made by an ML model
  - **Our focus in this lecture**

EXAMPLE: SCHEDULING ASSISTANT



- Analyze e-mail content for possible meeting scheduling
- Suggest creating a new meeting based on inferred information

Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-user Expectations of AI Systems. Kocielnik, et al. (CHI 2019)

DEALING WITH ERRORS IN ML

DEALING WITH ERRORS IN ML

- Define types of errors & their costs
    - False positives vs. false negatives
    - Optimize for one with lower costs
    - Q. For meeting scheduling, which are more acceptable?

DEALING WITH ERRORS IN ML

- Define types of errors & their costs
    - False positives vs. false negatives
    - Optimize for one with lower costs
    - Q. For meeting scheduling, which are more acceptable?
- Detect & record occurrences of errors
    - Collect telemetry or user feedback
    - Q. Telemetry to collect for meeting scheduler?

DEALING WITH ERRORS IN ML

- Define types of errors & their costs
    - False positives vs. false negatives
    - Optimize for one with lower costs
    - Q. For meeting scheduling, which are more acceptable?
- Detect & record occurrences of errors
    - Collect telemetry or user feedback
    - Q. Telemetry to collect for meeting scheduler?
- Identify sources of errors
    - Poor/bias training data, noise in data, data drifts

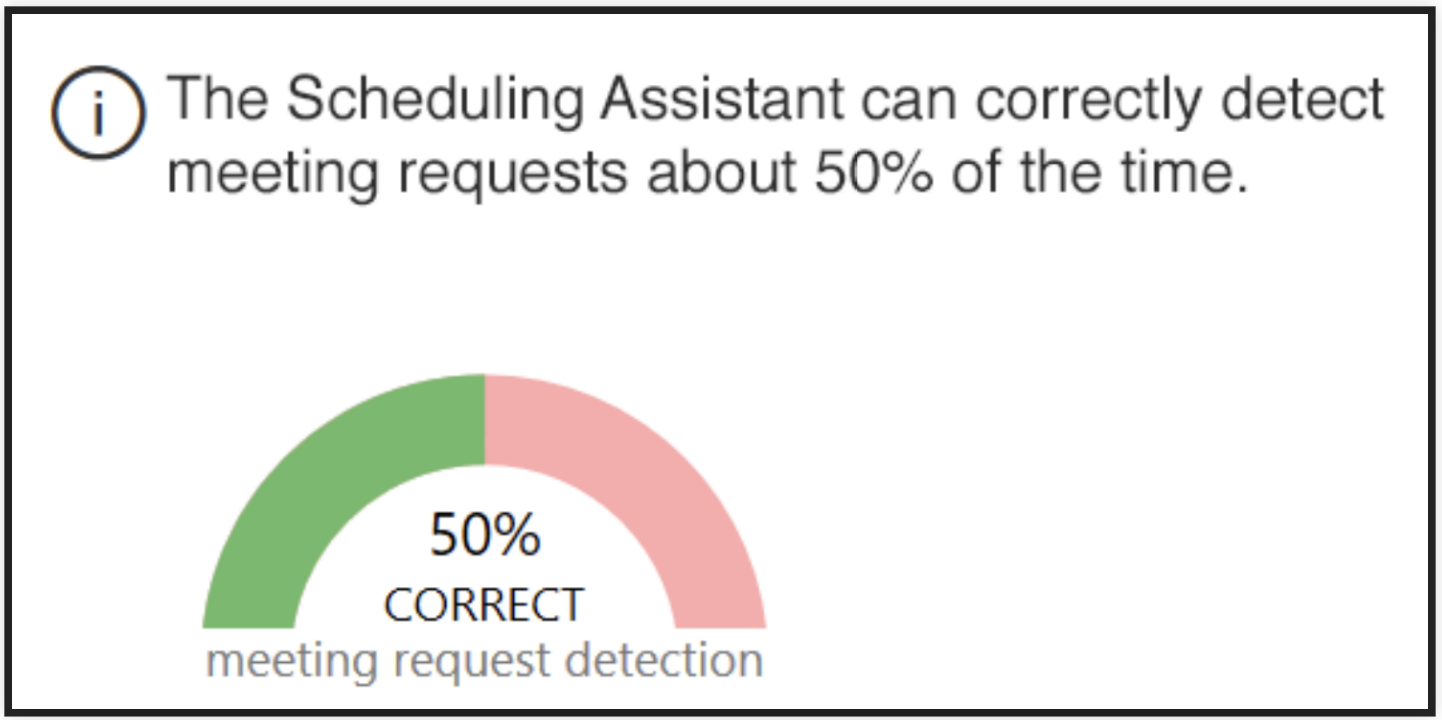
DEALING WITH ERRORS IN ML

- Define types of errors & their costs
    - False positives vs. false negatives
    - Optimize for one with lower costs
    - Q. For meeting scheduling, which are more acceptable?
- Detect & record occurrences of errors
    - Collect telemetry or user feedback
    - Q. Telemetry to collect for meeting scheduler?
- Identify sources of errors
    - Poor/bias training data, noise in data, data drifts
- Provide meaningful error messages to the user
    - Provide an explanation for the error
    - Suggest actions to fix the error (e.g., "Edit details" option)

DEALING WITH ERRORS IN ML

- Define types of errors & their costs
    - False positives vs. false negatives
    - Optimize for one with lower costs
    - Q. For meeting scheduling, which are more acceptable?
- Detect & record occurrences of errors
    - Collect telemetry or user feedback
    - Q. Telemetry to collect for meeting scheduler?
- Identify sources of errors
    - Poor/bias training data, noise in data, data drifts
- Provide meaningful error messages to the user
    - Provide an explanation for the error
    - Suggest actions to fix the error (e.g., "Edit details" option)
- Give user controls to recover from and mitigate the effect of an error
    - e.g., delete or modify incorrect meeting schedule

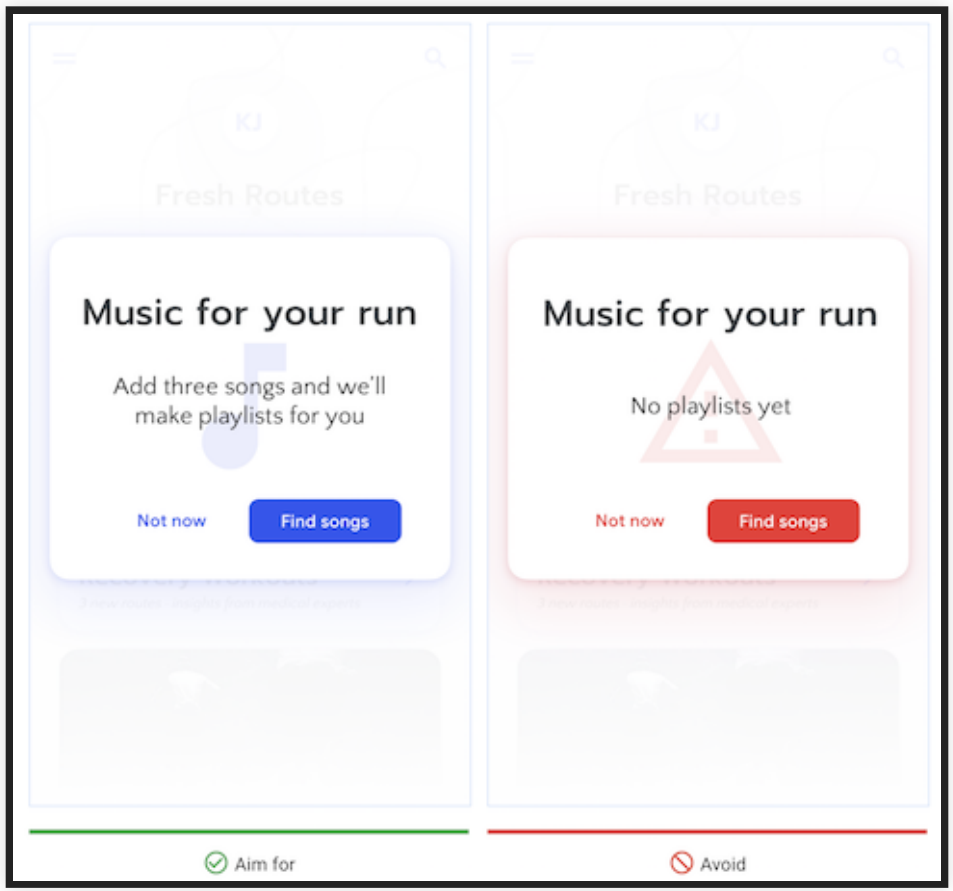
SETTING USER EXPECTATIONS FOR ML ERRORS



- Be upfront about how well the system performs (e.g., model accuracy)
- Temper the user's expectations and avoid surprises

Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-user Expectations of AI Systems. Kocielnik, et. al. (CHI 2019)

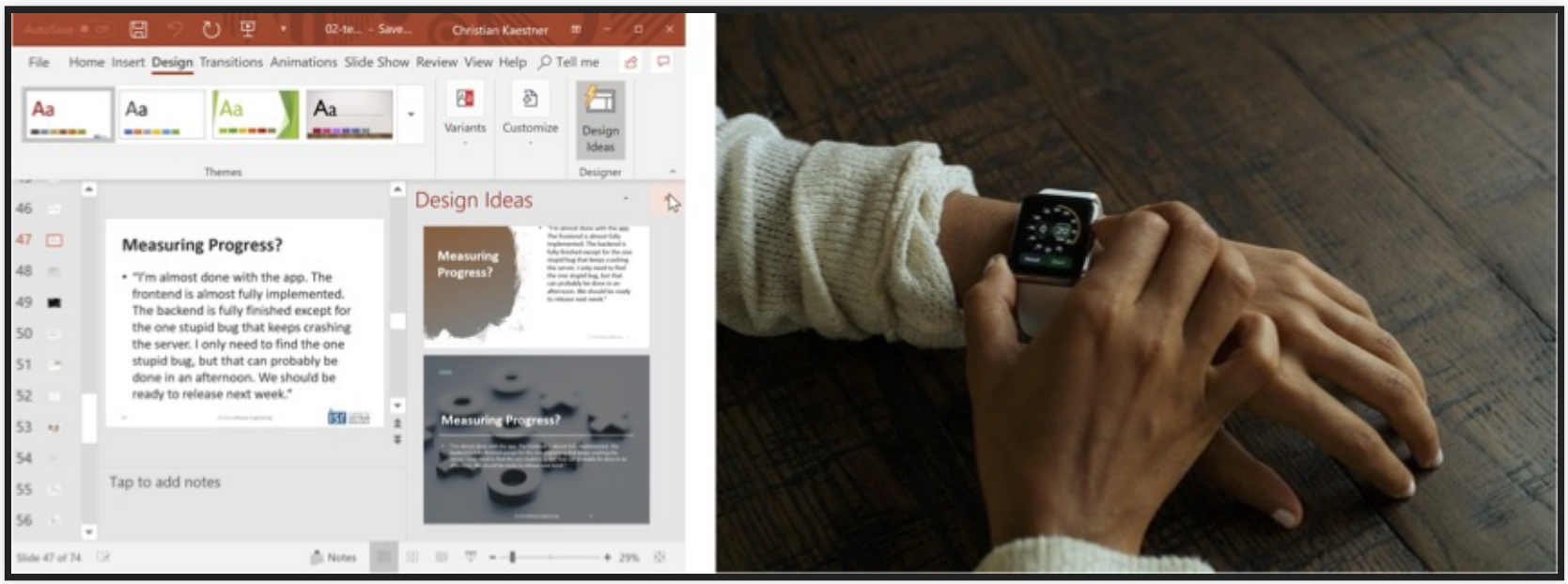
ERROR MESSAGES: SUGGEST USER ACTIONS



- Tell the user what the AI needs in order to behave as intended
- Guide the user towards ways to recover from/prevent further errors

https://pair.withgoogle.com/chapter/errors-failing/

BREAKOUT: DEALING WITH ERRORS



Design suggestions/fall detection

- In #lecture, type:
  - Possible error(s):
  - How to detect the error:
  - How to allow the user to recover from error:
  - What additional data to collect (from user) to reduce future errors:

FEEDBACK AND CONTROL

TYPES OF FEEDBACK

TYPES OF FEEDBACK

- Implicit feedback: Data about user behaviors collected by system
    - e.g., times of day, duration of usage, recommendations accepted/rejected, click patterns, etc.,

TYPES OF FEEDBACK

- Implicit feedback: Data about user behaviors collected by system
    - e.g., times of day, duration of usage, recommendations accepted/rejected, click patterns, etc.,
- Explicit feedback: Prompted or deliberately provided by user
    - Surveys, ratings, thumbs up, feedback forms, etc.,

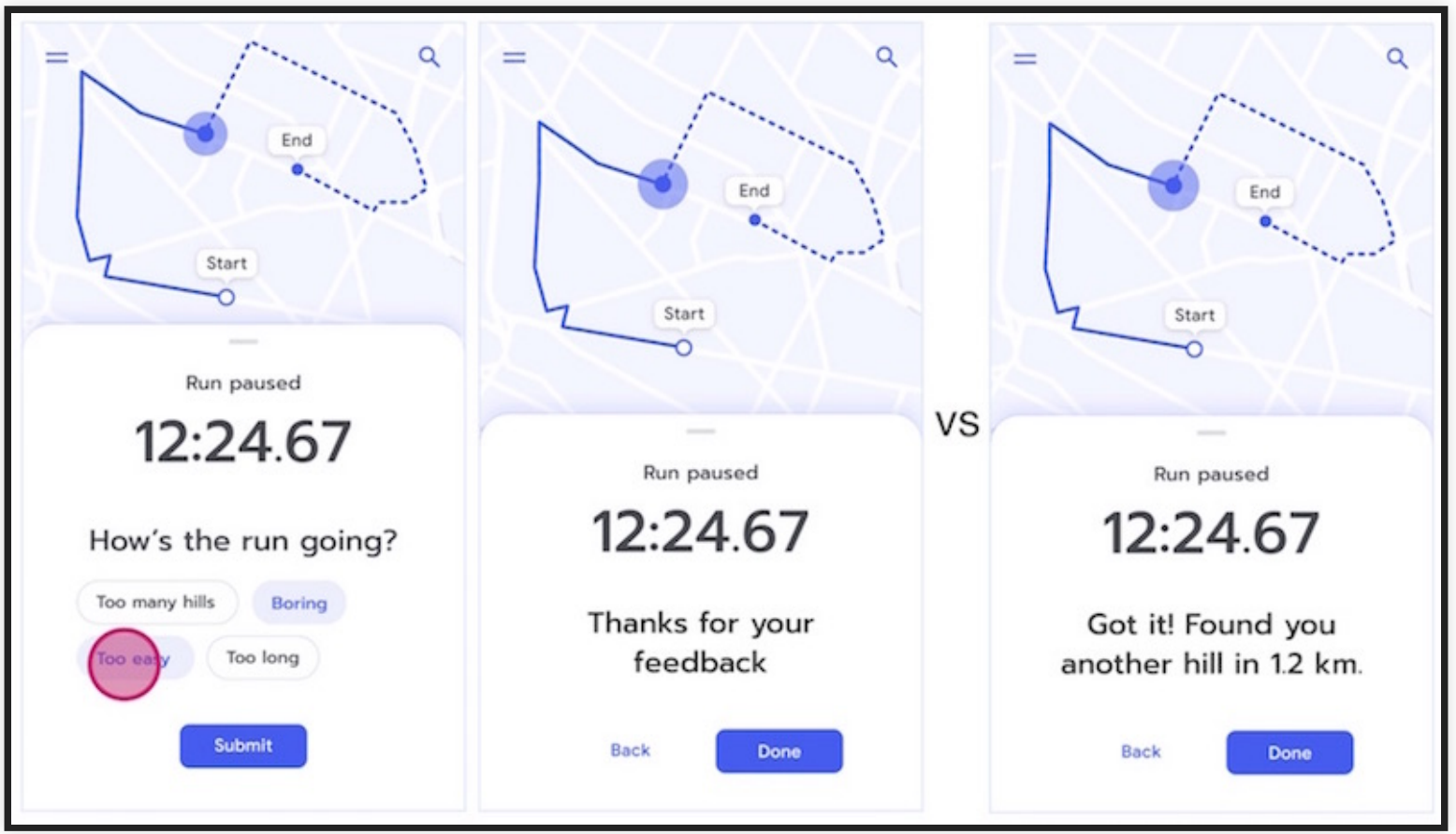
TYPES OF FEEDBACK

- Implicit feedback: Data about user behaviors collected by system
    - e.g., times of day, duration of usage, recommendations accepted/rejected, click patterns, etc.,
- Explicit feedback: Prompted or deliberately provided by user
    - Surveys, ratings, thumbs up, feedback forms, etc.,
- Design considerations for feedback
    - Align feedback with improving interactions (and AI)
    - Acknowledge user feedback & respond immediately

TYPES OF FEEDBACK

- Implicit feedback: Data about user behaviors collected by system
    - e.g., times of day, duration of usage, recommendations accepted/rejected, click patterns, etc.,
- Explicit feedback: Prompted or deliberately provided by user
    - Surveys, ratings, thumbs up, feedback forms, etc.,
- Design considerations for feedback
    - Align feedback with improving interactions (and AI)
    - Acknowledge user feedback & respond immediately
- In addition to feedback, provide a way for user to adjust AI behavior

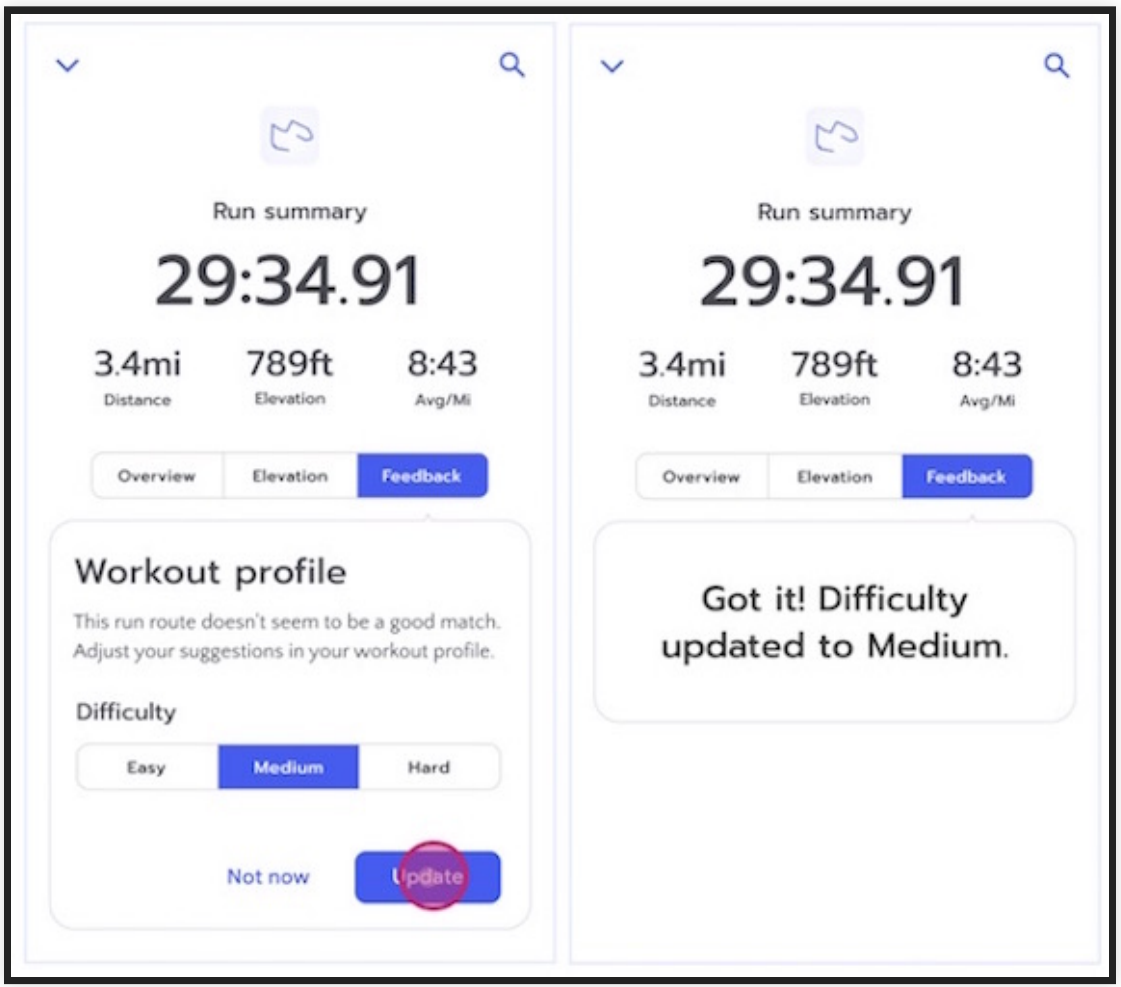
RESPONDING TO FEEDBACK



- When possible, respond to feedback with an adjustment to AI behavior

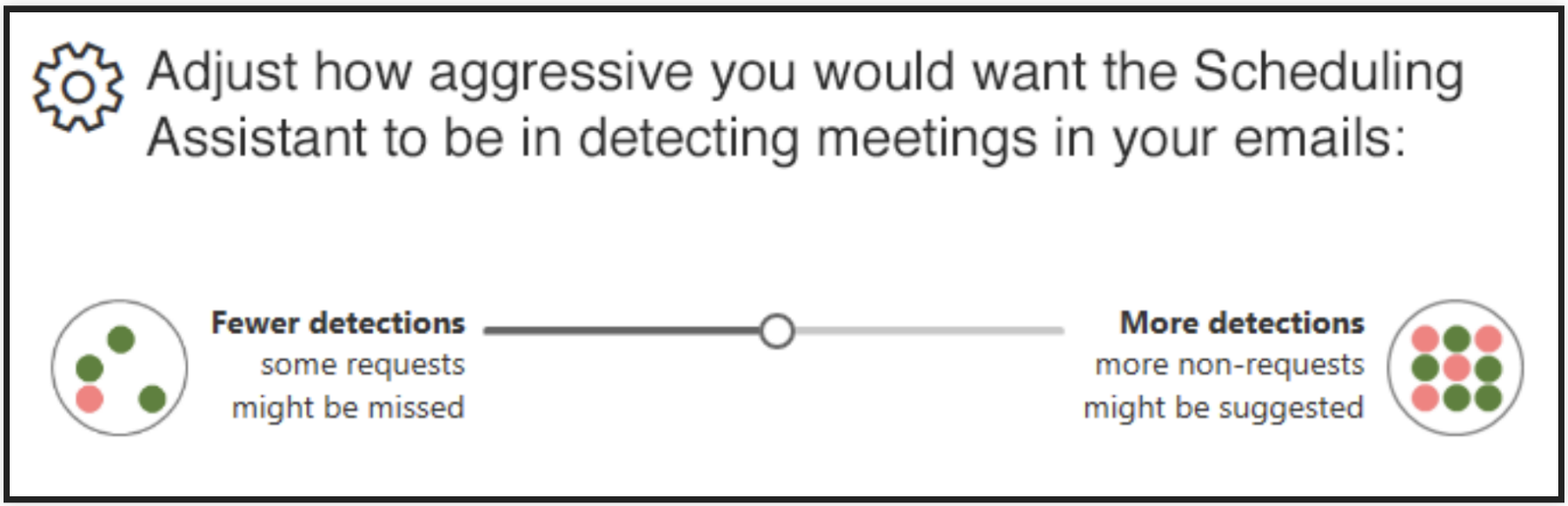
https://pair.withgoogle.com/chapter/feedback-controls/

GIVING USER CONTROL



- Provide a mechanism for user to adjust system behavior

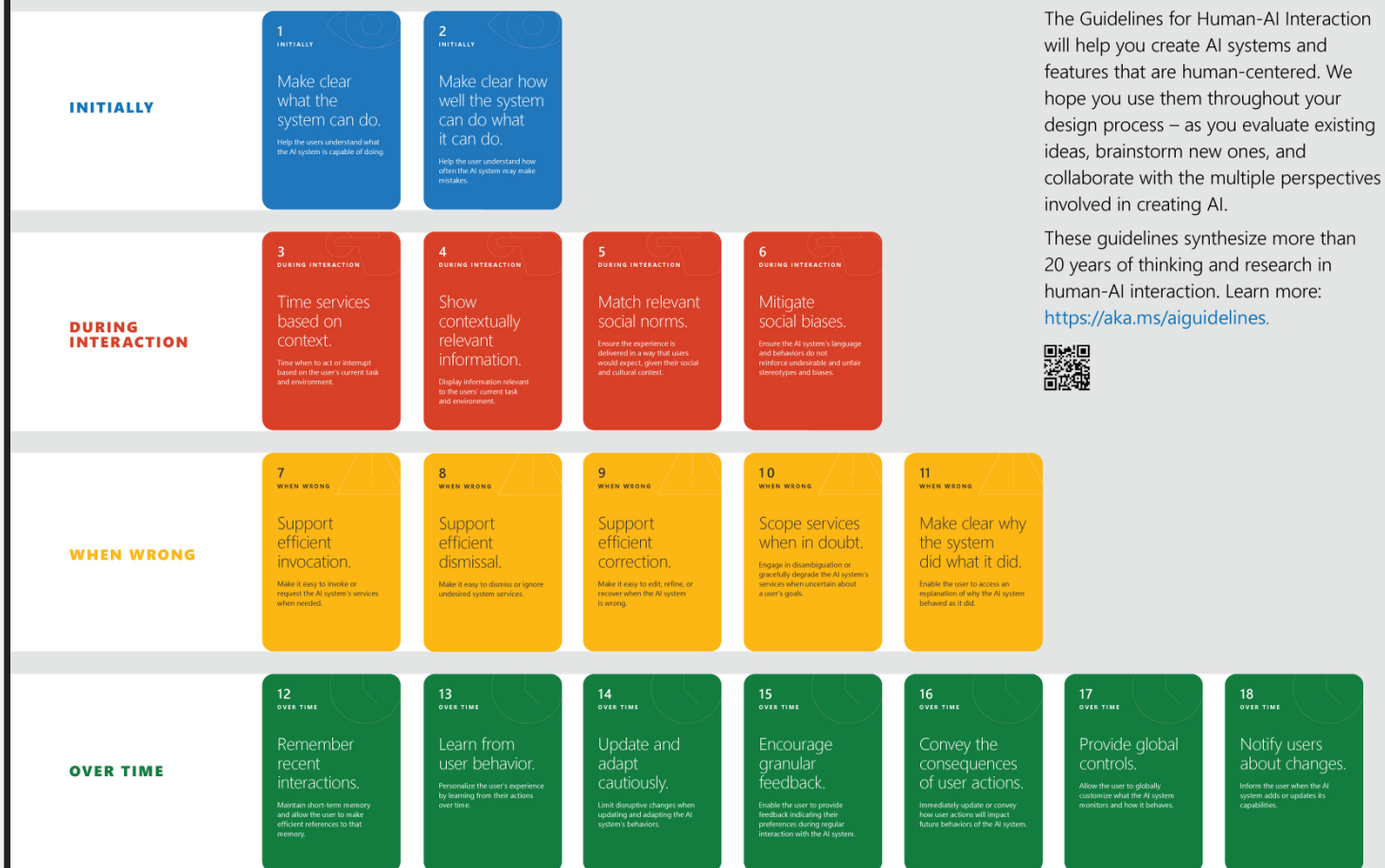
GIVING USER CONTROL OVER ML BEHAVIOR



- Provide a mechanism for the user to control the types of ML errors
- Scheduling assistant: Adjust thresholds to achieve trade-offs between precision vs recall

GUIDELINES FOR HUMAN-AI INTERACTIONS

Guidelines for Human-AI Interaction



The Guidelines for Human-AI Interaction will help you create AI systems and features that are human-centered. We hope you use them throughout your design process – as you evaluate existing ideas, brainstorm new ones, and collaborate with the multiple perspectives involved in creating AI.

These guidelines synthesize more than 20 years of thinking and research in human-AI interaction. Learn more: https://aka.ms/aiguidelines.

HUMAN-AI INTERACTIONS

Human-AI interactions must be considered throughout the entire ML lifecycle!

- Requirements & design
  - Understand user needs & their mental models
  - Explicitly design system to match the mental model
- During interaction
  - Consider factors for interaction (automate vs augment, forcefulness, frequency)
- When errors occur
  - Provide an explanation & actionable information
  - Provide ways for user to adjust AI behavior
- Maintenance and evolution
  - Collect user feedback and improve model
  - Adjust system design to reduce mental model mismatch

SUMMARY

- Goal of usable design: Minimize interaction cost
  - Automation does not necessarily imply reduced cost!
- Interaction design considerations for AI
  - Modes of interaction: Automate or augment?
  - Mental model: User understanding of what AI is doing
  - Dealing with errors: Guide user towards prevention & recovery
  - Feedback and control: Align user feedback with AI improvement