COMP 3647
Human-AI Interaction Design

Topic 14
Guidelines for Human-AI Interaction

Prof. Effie L-C. Law

Durham
University

Lectures, Week 10

4 Dec 2023: Ethics in AI

8 Dec 2023: Future Directions of AI

Dr. Swaroop Panda

Durham
University

Dagstuhl Seminar:

Model Learning for Improved Trustworthiness in Autonomous Systems

Schloss Dagstuhl
Where Computer Scientists Meet

QUICK LINKS

Welcome to Schloss Dagstuhl   Leibniz Center for Informatics

What are Guidelines?

A **guideline** is a statement by which to determine a course of action. A guideline aims to streamline particular processes according to a set routine or sound practice. (*Cambridge English Dictionary*)

A **guideline** is similar to a rule, but are legally less binding as justified deviations are possible. *(European Commission)*

Design guidelines are sets of recommendations on how to apply design principles to provide a positive user experience. *(Interaction Design Foundation)*

Standards, Rules, Principles & Heuristics in HCI

Standards: ISO 9241-110:2020(en): Ergonomics of human-system interaction — Part 110: Interaction principles

Usability metrics: Effectiveness, Efficiency, Satisfaction

Rules: Ben Sheniderman's 8 Golden Rules for Interface Design (1987)

#1. Strive for Consistency

Principles: Don Norman's 7 Design Principles (1988)

#7. When all else fails, standardize.

Heuristics: Jakob Nielsen's 10 Usability Heuristics for UI Design (1994)

#4: Consistency and standards

Heuristic Evaluation Process

- **Establish an appropriate list of heuristics.**

- **Select your evaluators.** usability experts with domain expertise.

- **Brief your evaluators**. ask the evaluators to focus on a selection of tasks

- **First evaluation phase.** It takes ~ two hours, depending on the nature and complexity of the product. Use the product freely to gain a feel for the methods of interaction and the scope. Identify specific elements to evaluate.

- **Second evaluation phase.** Carry out another run-through when applying the chosen heuristics to the elements identified during the first phase. Focus on individual elements and look at how well they fit in the overall design.

- **Record problems.** Record problems as they carry out their various tasks to track any problems they encounter. as detailed and specific as possible when recording problems.

- **Make recommendations.** Suggest potential solutions for these problems based on the heuristics.

Durham
University

Examples of Nielsen's 10 Usability Heuristics

Heuristic Evaluation Articles & Videos
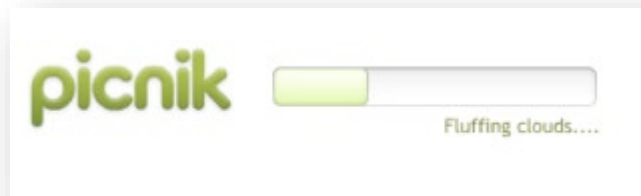https://www.nngroup.com/topic/heuristic-evaluation/

Durham University

1. Visibility of system status (Feedback)

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
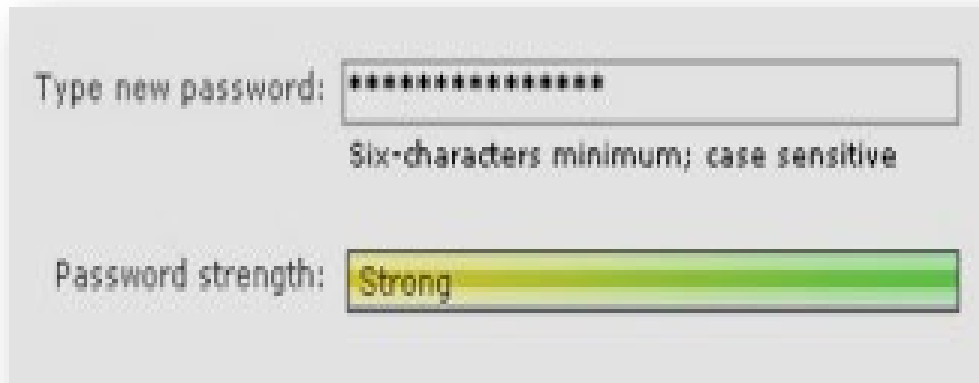
Windows Live Account
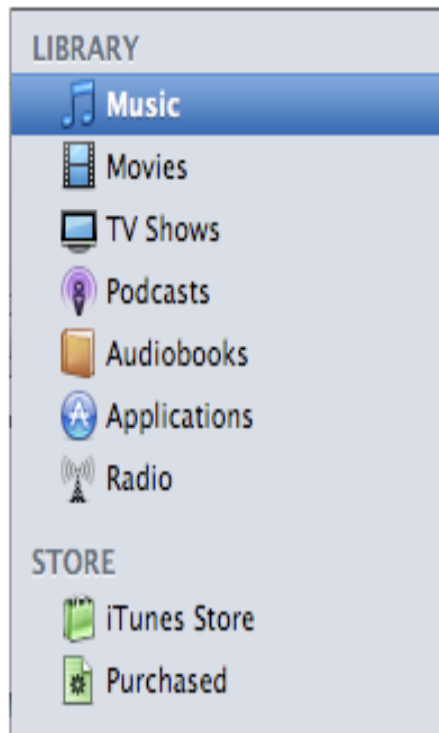Password strength is shown as the password is entered

Picnik
Progress message and indicator shows while the application loads

2. Match between system and the real world (METAPHOR)

The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order

LIBRARY
- Music
- Movies
- TV Shows
- Podcasts
- Audiobooks
- Applications
- Radio

STORE
- iTunes Store
- Purchased

iTunes
Organized as a library that contains your media library: music, movies, shows, audio-books. Beneath the Library is the Store where you can buy more media to put in your Library.

3. User control and freedom (NAVIGATION)

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue.
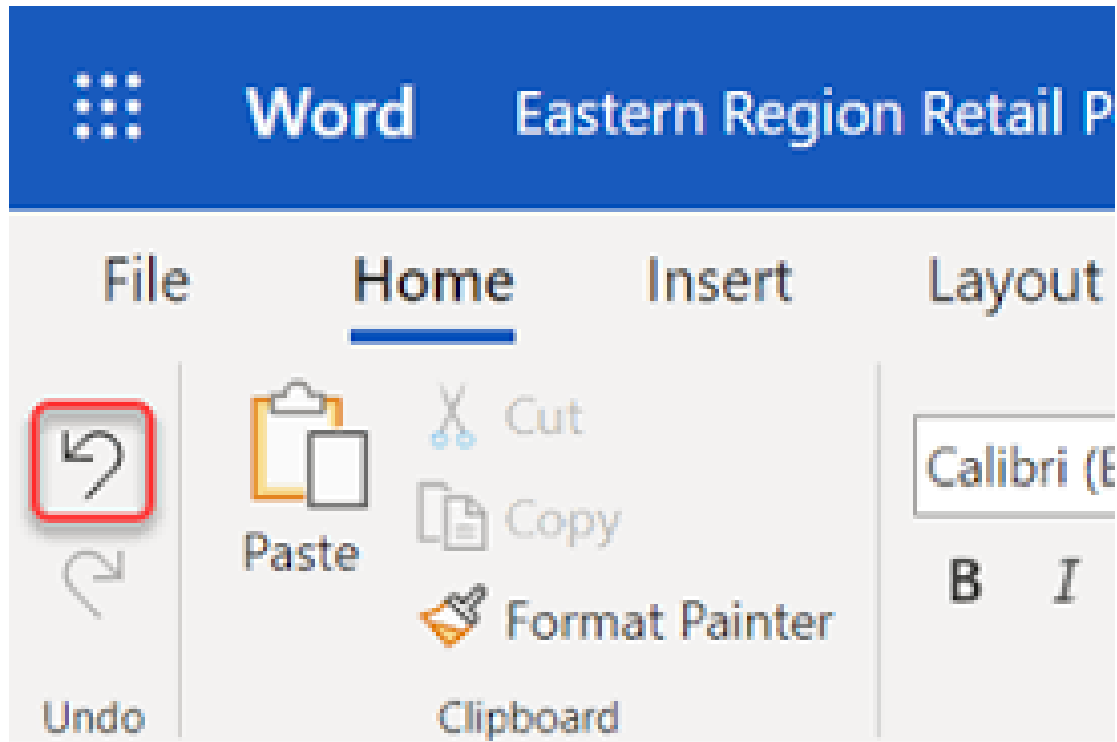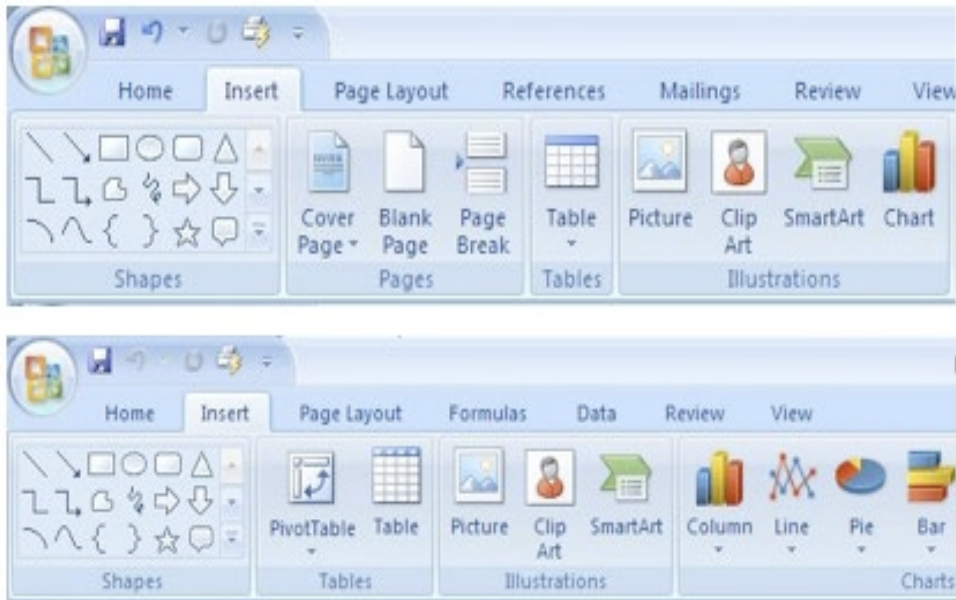

Undo/Redo

Supports undo and redo and a clear way to navigate.

.

4. Consistency and standards (CONSISTENCY)

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
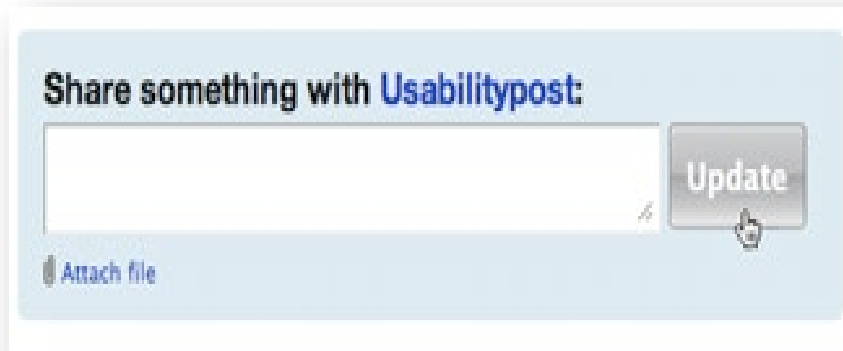

Microsoft Office
Word, Excel, and PowerPoint all use the same style toolbar with the same primary menu options: Home, Insert, Page Layout… Consistency results in efficiency and perceived intuitiveness.

5. Error prevention (PREVENTION)

Even better than good error messages is a careful design, which prevents a problem from occurring in the first place.

Yammer
Disables the update button after it is clicked, so the person cannot update the post twice by accident
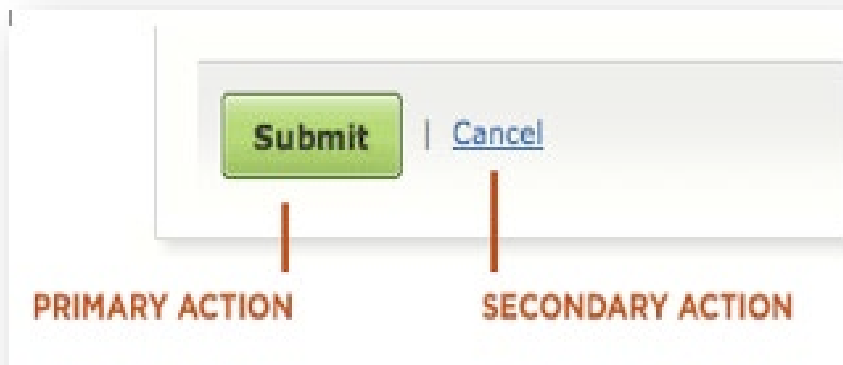
Example from "Web form Design:Filling in the Blanks" by Luke W.
Make the primary action prominent with a larger click area. Cancel and secondary actions are just shown as links

6. Recognition rather than recall (MEMORY)

Minimize the user's memory load. Make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
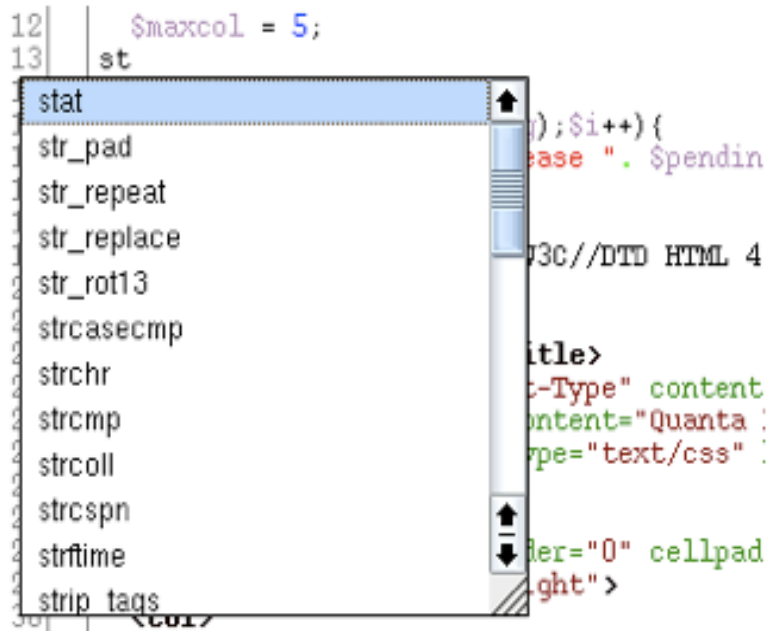
Quanta IDE
Type ahead for coding in a development environment
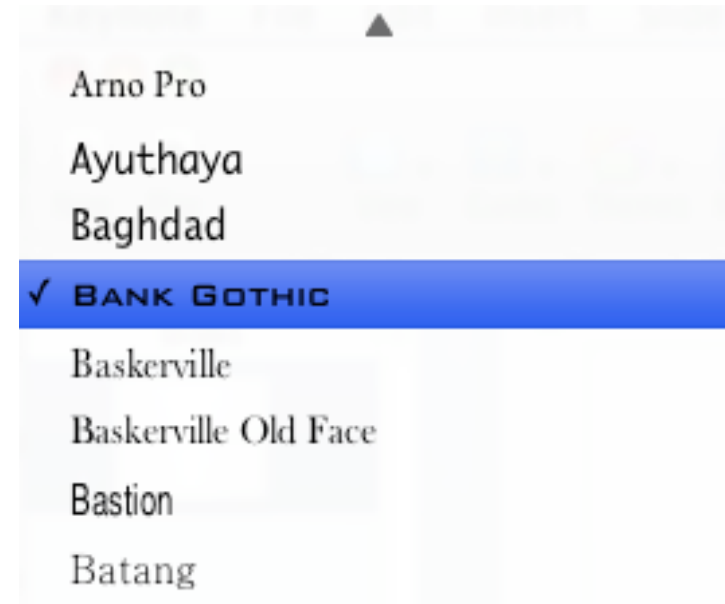
Keynote
Previews the fonts you can pick from, instead of just the font name

7. Flexibility and efficiency of use (EFFICIENCY)

Accelerators — unseen by the novice user — may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
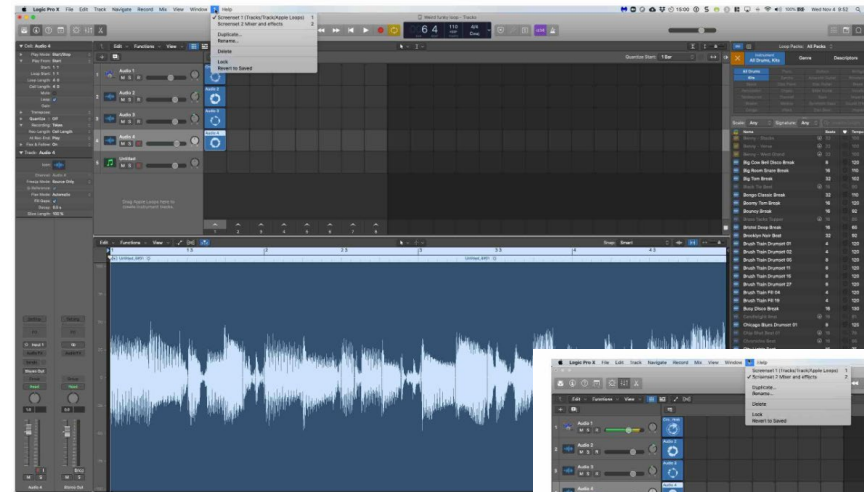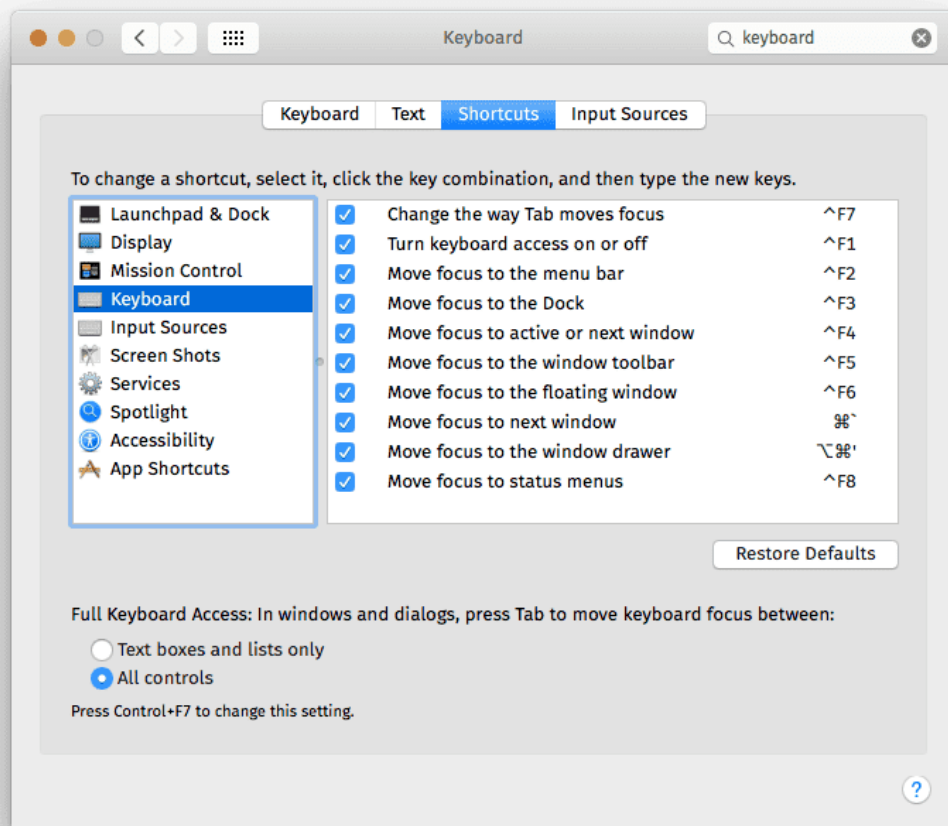
Screenset 1
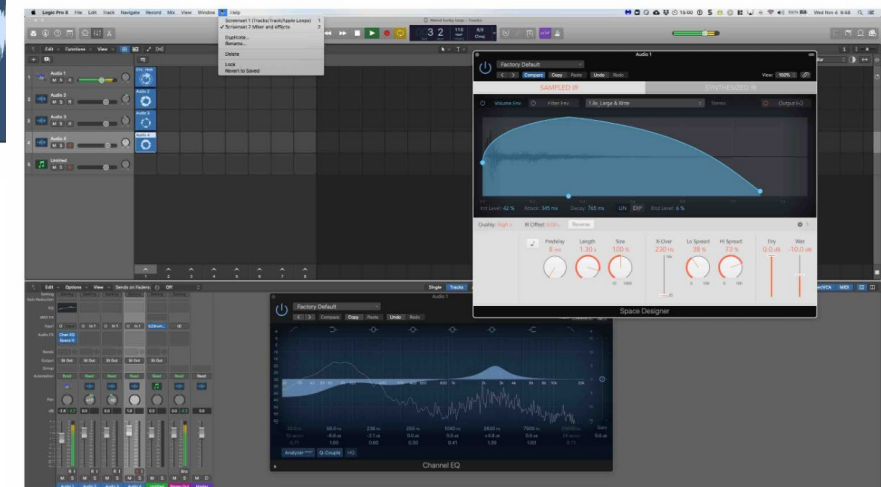
Screenset 2

8. Aesthetic and minimalist design (DESIGN)

Dialogues should not contain information, which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility..
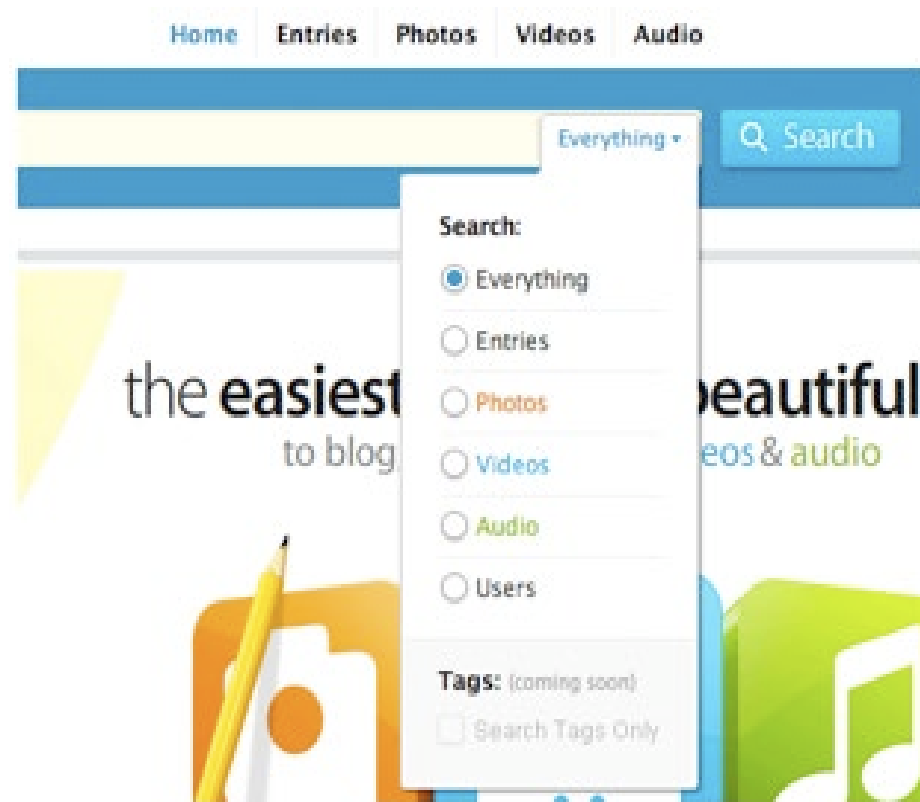
Kontain' search menu exemplifies the four principles of visual design:

Contrast: bold text is used for the two labels in the search

Repetition: the orange, blue, and green text match the media types

Alignment : strong left alignment of text, right aligned drop down

Proximity: a light rule is used to separate tags from the other options

9. Help users recognize, diagnose, and recover from errors (RECOVERY)

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

Digg
Provides immediate feedback with specific instructions
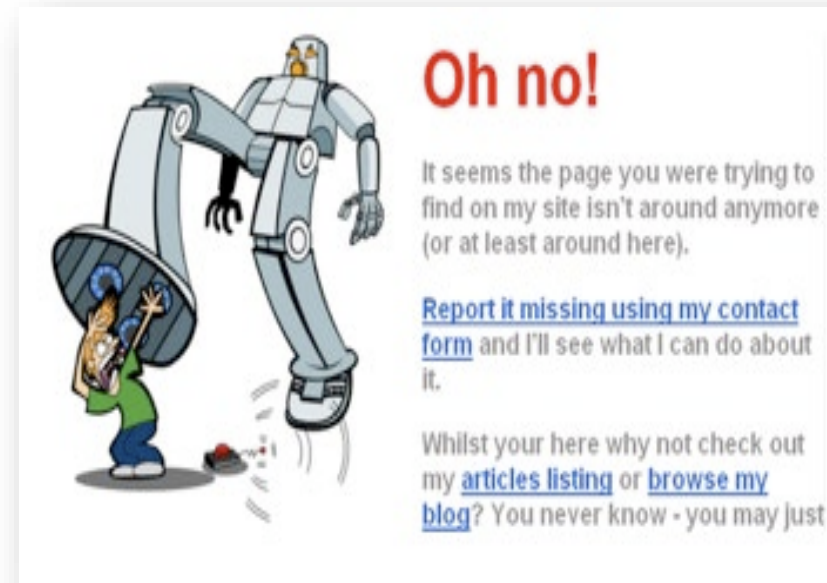
Humorous ' Page Not Found' Error
Uses a funny image and copy, but provides viable alternatives (article listings and blog link) and a course of action (report it)

10. Help and documentation (HELP)

Picnik

Contextual help (this is an example of help in the 'Collages' module) tips in Picnik are clear and easy to navigate
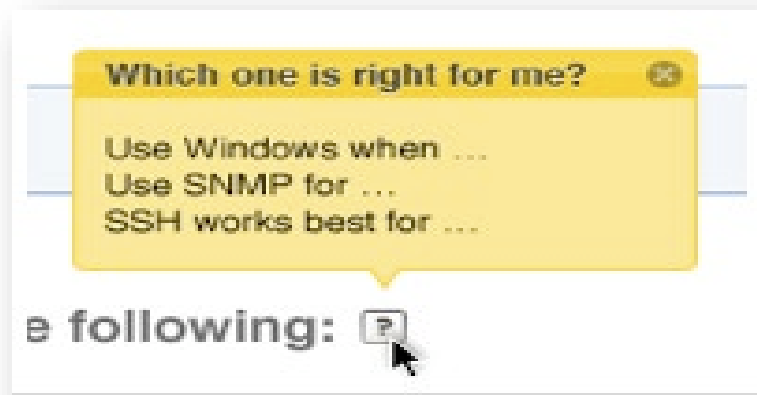
Zenoss

Help tips are displayed on hover, answering the most likely questions about a field or instructions

Three Industrial HAII Guidelines

- **Guidelines for Human-AI Interaction (2019)**

  - 168 academic/industry sources → 18 guidelines

  - User study validation: 2019, 2022 (publications)

- **People + AI Guidebook (2019)**

  - Mixed internal and academic → 113 guidelines

  - Developer centric

- **Human Interface Guidelines for Machine Learning (2019)**

  - Practitioner knowledge (no academic refs.) → 59 guidelines

  - UI/UX focus

Durham
University

Integrated HAII Guidelines

- 4 dimensions

- 12 categories

- 194 guidelines

- Ongoing work (GitHub)

- Lack of empirical validation
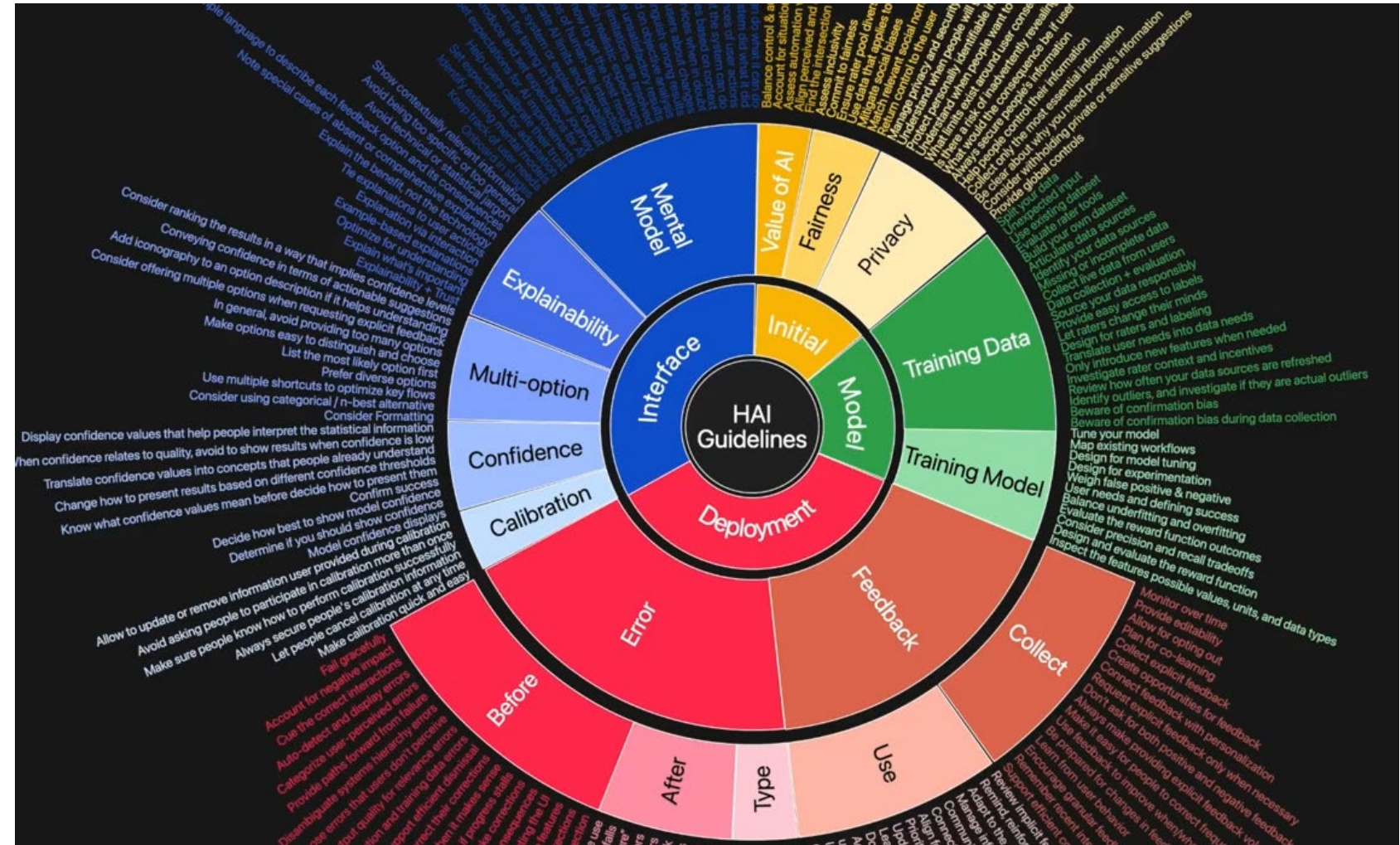
Comparisons on four dimensions
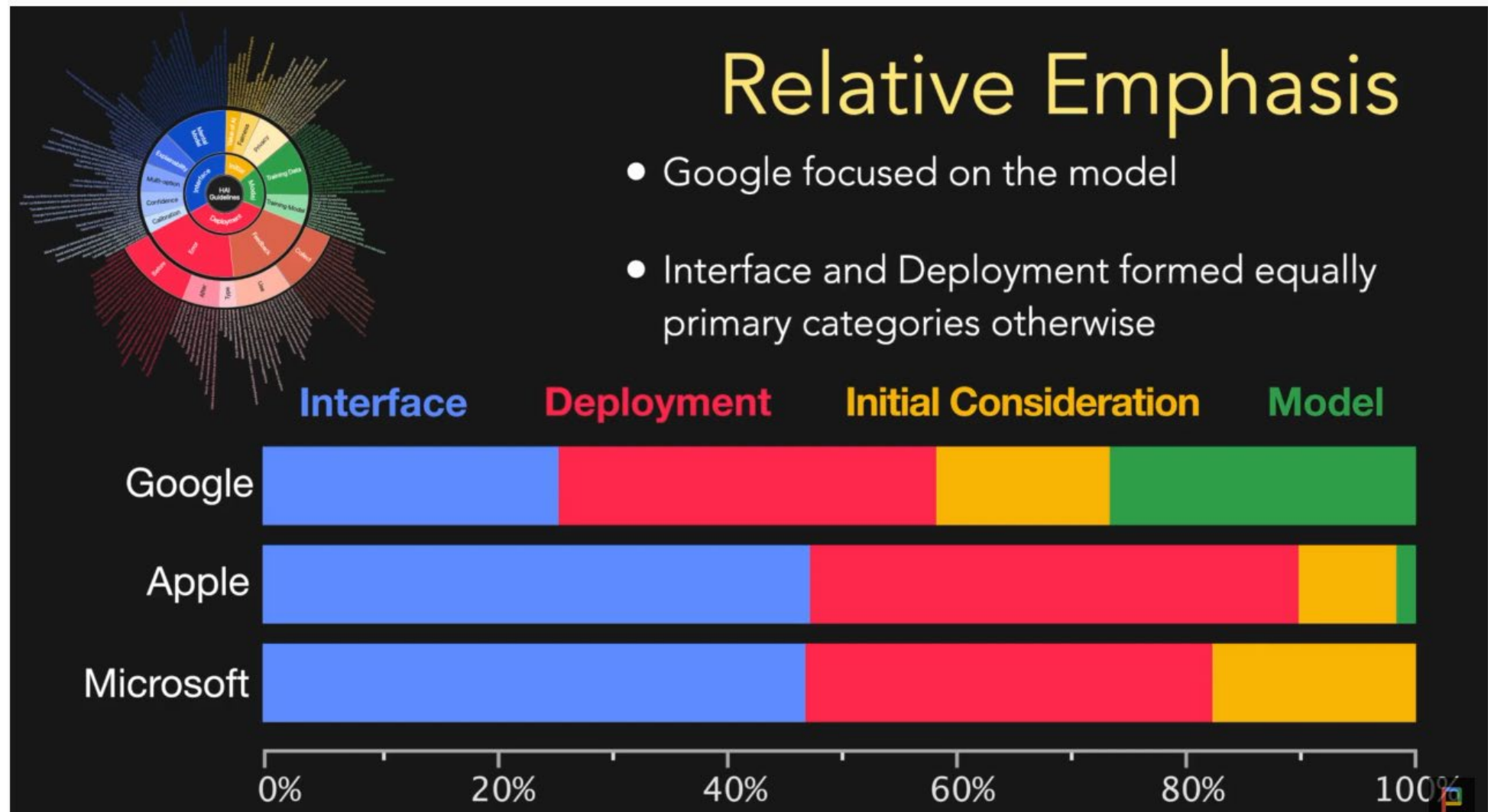
Comparisons on individual components

Apple, Google & Microsoft's Emphasis of Human-AI Interaction Guidelines

Wright, A. P., Wang, Z. J., Park, H., Guo, G., Sperrle, F., El-Assady, M., ... & Chau, D. H. (2020). A comparative analysis of industry human-AI interaction guidelines. *arXiv preprint arXiv:2010.11761*.
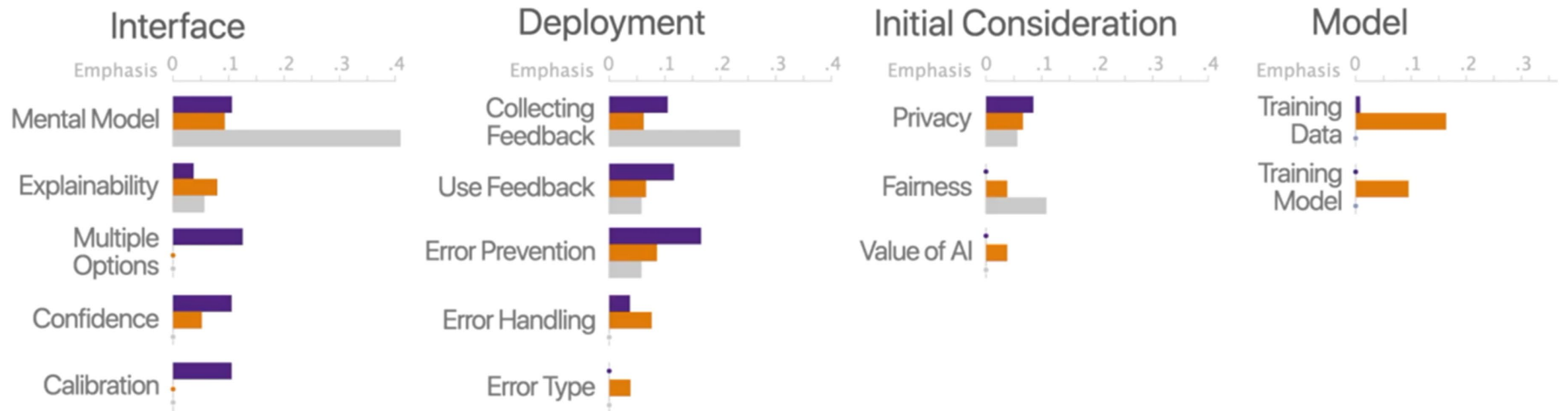
Guidelines for Human-AI Interaction

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2019, May). Guidelines for human-AI interaction. In Proceedings of the 2019 chi conference on human factors in computing systems (pp. 1-13).

Evaluation of the Guidelines for HAII

Why? To determine if and how each of the design guidelines manifests in a variety of AI-infused products.

How? A *maximum-variance* sampling strategy to select popular AI-infused products, covering a wide range of scenarios

- top 10 types of apps, software and websites in US for both mobile and desktop devices (filtered out any offensive, game related, not using AI for end-user facing services)

- 10 categories, two products per category based on market share

- Prominent AI-driven feature to evaluate per product

Product Category	Feature	Participants
E-commerce (Web)	Recommendations	6
Navigation (Mobile)	Route planning	5
Music Recommenders (Mobile)	Recommendations	5
Activity Trackers (Device)	Walking detection and step count	5
Autocomplete (Mobile)	Autocomplete	5
Social Networks (Mobile)	Feed filtering	5
Email (Web)	Importance filtering	5
Voice Assistants (Device)	Creating a reminder with a due date	5
Photo Organizers (Mobile)	Album suggestions	4
Web Search (Web)	Search	4

Durham
University

Participants

- At least one year of experience working in or studying HCI and familiar with heuristic evaluation and the product types to be evaluated

- Assign 2-3 participants to each product

- 49 participants
  - *Gender*: 29 female, 18 male, 2 preferred not to answer
  - *Age*: 18-24 (n=5), 25-34 (n=24), 35-44 (n=13) and 45-54 (n=7).
  - *Role*: 19 researchers, 12 designers, 11 HCI or design interns from various universities worldwide, and 7 a mix of engineers, product managers or vendors.
  - *Experience*: 1-4 years (n=23), 5-9 years (n=14), 10-14 years (n=9), 15-19 years (n=1), 20+ years (n=2).

Clear Applications

	E-commerce	Navigation	Music Recommenders	Activity Trackers	Autocomplete	Social Networks	Email	Voice Assistants	Photo Organizers	Web Search	Total
G1	4	4	2	2	2	1	3	2	3	3	26
G2	2	4	1	1	1	3	2		2		16
G3		4	1	2	3	4		2	1	1	18
G4	2	6		2	1	5	2	2	3	3	26
G5	1	2	1	1	3	2	3	4	1	4	22
G6	2	2	2	1	2	2		1	1	1	13
G7	2	1	1	4	3	1	1	3	1	3	20
G8	2	3	2	2		3	4	3	1	3	23
G9	1	2	1		2	2	3	3	1	2	17
G10	2	1			1			1		4	9
G11	2	3	1			1	1				8
G12	2	5	2	2	3	3	2	2	2	5	28
G13	2	1	5		4	4	2		2	2	22
G14	2	2	3		2	2	3		1	3	18
G15	1	1	5		4	3			4		18
G16			2	1		3	2	2		1	11
G17	3		1	1	1	2		1	2	2	13
G18	2				1	1		1			5
Total	27	46	30	19	28	43	32	25	22	41	

Clear Violations

	E-commerce	Navigation	Music Recommenders	Activity Trackers	Autocomplete	Social Networks	Email	Voice Assistants	Photo Organizers	Web Search	Total
G1	2	1	1			2	2	3		3	14
G2	2	4	4	2	2	2	1	4	1	1	23
G3	1		2			2		1			6
G4	3	1	1	1		2	1				9
G5	1		2	3			2	1			9
G6	3	1	1		2	2	2	1	2	3	17
G7	3	2	2	1	1				1	1	11
G8	3	1	3	1	3	3					14
G9	2	4	2	3	2	1	1	2	2	2	20
G10	1	2	3	3	3		2	1	2		17
G11	5	2	4	1	2	2	4	3	1	3	27
G12	2	1	1	1	1		2	1	1		13
G13	2	2		2		2	2	1	1		12
G14	1				2		1				4
G15	2	2	1	2	1	1	3	2	3		17
G16	3	1	1		2	4	4	1	2	3	21
G17	3	3	3	4	4	3	3	3	1	2	29
G18	2	1	3	1	1	2	1		1	2	14
Total	40	29	31	27	25	26	31	28	18	22	

Does NOT Apply

[Initially] G1: Make clear what the system can do.
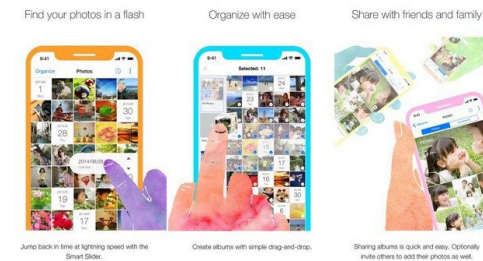## Help the user understand what the AI system is capable of doing

Example Applications

- **Navigation, Product #1:** Specific examples of things I could search for are in the search bar

- **Activity Trackers, Product #1:** Displays all the metrics that it tracks and explains how. Metrics include movement metrics such as steps, distance travelled, length of time exercised, and all-day calorie burn, for a day.

- **Social Networks, Product #2:** The product communicates to users that it will evaluate and provide to you potential people to follow based on your interests and the community of people you follow.

Example Violations

- **Social Networks, Product #1**: This guideline is violated because I cannot even tell what this news feed can/will show.

- **Photo Organizers, Product #1**: We know the AI is able to detect and associate an image with a category, but the user does not know all the categories available.

- **Voice Assistants, Product #1**: When [the assistant] was invoked verbally...I was not given any indication of what commands I could request."

[Initially] G2: Make clear how well the system can do what it can do. Help the user understand how often the AI system may make mistakes

Example Applications

- **Email, Product #2;** While the interface does not give any indication about the level of performance possible, the help page is an application of this heuristic. It sets the expectation that it will start working right away, but will get better with use, making it clear that mistakes will happen and you can 'teach' [the product] to better perform, and failing that, you can set overrides.

Example Violations

- **Navigation, Product #1**: There is no indication of accuracy of the time estimates or how conditions may be changing. There is no measure given of how well the AI predictions matched the result once you arrive at the destination.

- **Voice Assistants, Product #1:** Aside from the 'Hi, how can I help?', [the product] does not promise anything more. No expectation of quality is set.

Durham University

[During Interaction] G3: Time services based on context.
Time when to act or interrupt based on the user's current task and environment

Example Applications

- **Navigation, Product #1:** In my experience using the app, it seems to provide timely route guidance. Because the map updates regularly with your actual location, the guidance is timely.

- **Autocomplete, Product #1**: Suggestions are always present when you might need them (whenever the keyboard is up)

- **Social Networks, Product #1:** If the user has not accessed [the product] in a while, the application will let the user know that there is something new to be explored - a story, video, etc.

Example Violations

- **Activity Trackers, Product #2**: Context is very basic, it notifies when I approach my goal; hit my goal; or exceed my goal. The timing of it is not clear, however. The timestamps are varied, too... It feels pretty arbitrary; my interpretation of the reasoning behind the notification can't be described by my activity or proximity to the goal.

- **Email, Product #2**: Sending notifications for unimportant messages likely something most people will not want as an interruption.

Durham University

[During Interaction] G4: Show contextually relevant information.. Display information relevant to the user's current task and environment.

Example Applications

- **E-commerce, Product #2**: The feature assumes I'm about to buy a gaming console and shows accessories and games that would go with it, and it features those items prominently (above the product information) on the webpage."

- **Web Search, Product #2**: Searching a movie title returns show times near my location for today's date.

- **Navigation, Product #1**: When I use [the product] for driving directions, it remembers where I parked my car. Next time when I open the app, it suggests routing me back to my car

Example Violations

- **Activity Trackers, Product #2.** They chose to have a uniform view regardless of context. When I'm moving or static, the view is the same. The only change is that the step counter advances (or not).

- **E-commerce, Product #1.** If I start looking at a new item (like paper towels), when I scroll down to the end of the page, I get recommendations related to the recent items I viewed (tennis balls)... it doesn't take into account that I am currently looking for paper towels and I have already purchased the tennis balls.

[During Interaction] G5: Match relevant social norms. Ensure the experience is delivered in a way that users would expect, given their social and cultural context.

Example Applications

- **Photo Organizers, Product #1. [**The product's] album suggestions feature is able to recognize people's pets and uses the verbiage "important cats & dogs", understanding that people's pets are important to users and are like family even.

- **Voice Assistants, Product #1**. [The assistant] uses a semiformal voice to talk to you - spells out "okay" and asks further questions."

- **Navigation, Product #1**. If you select walking, the AI avoid(s) busy roads and searches for trails.

Example Violations

- **Activity Trackers, Product #1.** Provides a reminder to stand up without understanding my social context (e.g., in a meeting, having lunch with a friend etc)... Does not consider the social context prior to sending notifications for activity and does not use tone appropriately - just says "time to stand!" no matter what

- **Email, Product #2**. The system does not follow the social norms of a workplace. For example, one norm is to pay attention to your manager. However, even with access to company hierarchies, it isn't clear that the system will put messages from one's direct manager in the [appropriate tab] automatically.

Durham University

[During Interaction] G6: Mitigate social biases. Ensure the AI system's language and behaviors do not reinforce undesirable and unfair stereotypes and biases.

Example Applications

- **E-commerce, Product #2**: The feature does not unfairly assume gender biases in some search results that could potentially introduce them. For example, a search for tools or diapers could accidentally serve related products that are gender biased. The system seems to provide highly specific recommendations of the very same product type.

- **Web Search, Product #2**: A search for CEO or Doctor shows somewhat diverse people in the resulting images...The images are pretty diverse in terms of gender and ethnicity, although still lack in some respects such as disability"

Example Violations

- **Navigation, Product #1**: Regards the "Walking" transport there's no way to set an average walking speed. [The product] assumes users to be healthy.

- **Voice Assistants, Product #1:** When asked "...can you change your voice" [the assistant] responds in a male voice saying "Here is an example of my other voice. Would you like me to use this one?"... While it's nice that a male voice is given as an option, the default...voice is female, which reinforces stereotypical gender roles that presume a secretary or receptionist is female.

[When Wrong] G7: Support efficient invocation. Make it easy to invoke or request the AI system's services when needed.

Example Applications

- **[Voice Assistants, Product #1]** "I can say [wake command] to initiate."

- **[E-commerce, Product #1]** "In addition to the system giving you recommendations as you browse, you can go to your "Browsing history > Manage history > More like this" to get recommendations specific to a particular product."

- **[Web Search, Product #1]** "user can highlight a specific part of an image to search for that specific piece. This shows up on every image, so user can use it anytime they like."

Example Violations

- **[Navigation, Product #1]** "[The product] remembers where you parked your car. However if it fails to remember, or I want it to remember something else (e.g., where I chained up my bike), it is not possible (or at least not easily discoverable) to invoke the capability when I need it."

- **[Navigation, Product #2]** "Guideline is violated because user cannot ask the system for alternative routes if they are not detected initially."

Durham University

[When Wrong] G8: Support efficient dismissal. Make it easy to dismiss or ignore undesired AI system services.

Example Applications

- **[E-commerce, Product #2]** "Feature is unobtrusive, below the fold, and easy to scroll past...Easy to ignore."

- **[Social Networks, Product #2]** "[The product] allows the user to easily hide or report ads that have been suggested by the AI by tapping the ellipses at the top right of the ad."

- **[Voice Assistants, Product #1]** "I can say "nevermind" to dismiss it once I have said [wake command]. I can also just not say anything and it stops listening."

Example Violations

- **[Autocomplete, Product #2]** "I didn't see a dismiss button. I can dismiss it by dismiss the whole keyboard"

- **[Activity Trackers, Product #1]** "The guideline is violated because it is not clear within the Steps page, how to turn off the background/ambient step tracking functionality."

- **[Navigation, Product #1]** "Suggested locations based on calendar entries can't be removed from the suggestions."

Durham University

[When Wrong] G9: Support efficient correction. Make it easy to edit, refine, or recover when the AI system is wrong

Example Applications

- **[Voice Assistants, Product #2]** "Once my request for a reminder was processed I saw the ability to edit my reminder in the UI that was displayed. Small text underneath stated "Tap to Edit" with a chevron indicating something would happen if I selected this text."

- **[Navigation, Product #1]** "If [the product] is wrong about where I parked my car, it provides an easy way to edit the location by dragging on the map."

Example Violations

- **[E-commerce, Product #1]** "I already recently bought the items which are in my recommendation list & there is no message to discontinue nor option for users to deselect."

- **[Activity Trackers, Product #1]** "As far as I can tell, there is no way for the user to edit the number of steps collected. The user can delete the data point altogether...There is just no way to manually input or change the data

Durham University

[When Wrong] G10: Scope services when in doubt. Engage in disambiguation or gracefully degrade the AI system's services when uncertain about a user's goals

Example Applications

- **[Navigation, Product #1]** "If more than one line takes the same route the user can choose between the preferred line."

- **[Autocomplete, Product #1]** "It usually provides 3-4 suggestions instead of directly auto completing it for you"

- **[Voice Assistants, Product #2]** "If I didn't respond or if I spoke quietly, [the assistant] let me know they had trouble hearing me"

Example Violations

- **Navigation, Product #1]** "When searching for a specific restaurant, if the right place can't be found, it fails to recognize that you are searching for food and does not suggest restaurants in your area...The AI doesn't recognize the broader user goal and suggest possible alternatives when it can't find the exact result."

- **[Email, Product #2]** "This heuristic is violated. There is no indication if the system is unsure. The system does not ask for assistance (e.g. request feedback on an uncertain classification). It is not clear how it decides what to do with unclear messages, or what the current quality of the model is, beyond the information from the help screen which says the system will improve over time."

[When Wrong] G11:Make clear why the system did what it did. Enable the user to access an explanation of why the AI system behaved as it did

Example Applications

[**E-commerce, Product #1**] "Clicking "Why recommended" explains why they have recommended that particular item to you."

[**Music Recommenders, Product #2**] "I think this applies because each of recommendation has some information as to which songs are displayed on it - similar to the song, from the same artist, from the same album etc."

[**Navigation, Product #2**] "The route chosen by the app was made based on the Fastest Route, which is shown in the subtext."

Example Violations

- [**Email, Product #2**] "There is no indication of why a message is classified as it is, and I cannot find a way to find out. A message could be classified as [unimportant] because of the content, the sender, or any number of other reasons and no details are available, either per message or as a summary of the [unimportant category] rules which have been learned."

- [**Activity Trackers, Product #2**] "There are no explanations in the app about how walking detection and step count are measured. The system is pretty deterministic in how it makes measurements about my walking behaviour and then how it matches them with the initial goal I set."

[When Wrong] G11:Make clear why the system did what it did.
Enable the user to access an explanation of why the AI system behaved as it did

Example Applications

[**E-commerce, Product #1**] "Clicking "Why recommended" explains why they have recommended that particular item to you."

[**Music Recommenders, Product #2**] "I think this applies because each of recommendation has some information as to which songs are displayed on it - similar to the song, from the same artist, from the same album etc."

[**Navigation, Product #2**] "The route chosen by the app was made based on the Fastest Route, which is shown in the subtext."

Example Violations

- [**Email, Product #2**] "There is no indication of why a message is classified as it is, and I cannot find a way to find out. A message could be classified as [unimportant] because of the content, the sender, or any number of other reasons and no details are available, either per message or as a summary of the [unimportant category] rules which have been learned."

- [**Activity Trackers, Product #2**] "There are no explanations in the app about how walking detection and step count are measured. The system is pretty deterministic in how it makes measurements about my walking behaviour and then how it matches them with the initial goal I set."

[Over time] G12: Remember recent interactions. Maintain short term memory and allow the user to make efficient references to that memory.

Example Applications

- **[Navigation, Product #1]** "Opening the app shows a list of recent destinations, as well as allows you to access "favorite" locations."

- **[Web Search, Product #1]** "[The search engine] remembers the context of certain queries, with certain phrasing, so that it can continue the thread of the search"

- **[Voice Assistants, Product #1]** "[The assistant] seems to remember conversation context at least one command back."

Example Violations

- **[Social Networks, Product #1]** "This guideline is violated since there is no indication of "what you have read""

- **[Web Search, Product #2]** "Although many food/recipe searches were made right before, a search for "stir fry" has a music video as a top result"

Durham University

[Over time] G13:Learn from user behaviour. Personalize the user's experience by learning from their actions over time

Example Applications

- **[Social Networks, Product #1]** "I am pretty sure they do this as I have observed that I get more ads of a certain type if I accidentally or deliberately hovered over the ad for a while. I also have a suspicion that friends are also preferred that are more likely to generate a like or comment from me. It is less clear how negative signals are taken into account."

- **[Email, Product #1]** "(My guess is) the system learns from what previous emails have attracted more attention from me (i.e. longer/more frequent reply, reading time taken, longer email threads, etc.) and infer email importance."

Example Violations

- **[Web Search, Product #1]** "The search results do not take into account the user's previously visited sources. In this example from real life: I search for recipes on [the search engine] all the time, and I only ever click on recipes from 4 sources: Serious Eats, Pioneer Woman, 101 Cookbooks, and Smitten Kitchen. I often search for these sites as navigational queries on [the search engine]. However, [the search engine] rarely surfaces recipes from these sites for me."

- **[Navigation, Product #1]** "If you don't take the suggested route because you don't want stop and go traffic on the highway, it never learns. You can have it suggest routes to avoid tolls and highways, but you have to manually select those preferences."

[Over time] G14:Update and adapt cautiously. Limit disruptive changes when updating and adapting the AI system's behaviours

Example Applications

- **[Music Recommenders, Product #2]** "Once we select a song they update the immediate song list below but keeps the above one constant."

- **[Social Networks, Product #1]** "Think this is good. When I unfollow someone it shows their stuff for a little bit? But after a day or so it's gone. Or once I reload. But not RIGHT away."

- **[Web Search, Product #2]** "After clicking and returning from a search result, the order of the search results didn't change. The updated "people also searched for" that was relevant to the clicked result was highlighted and contained"

Example Violations

[Social Networks, Product #1] "The refresh (pulling down action) can sometimes be triggered unintentionally and thus causes some posts disappeared."

[E-commerce, Product #1] "Maybe this is just my experience, but when I accidentally browse an item suddenly my entire recommendation list changes to things relevant to that new item. The change is really jarring, especially when the browsing is a result of a curious moment or an accident. For example, I just clicked on one camera after viewing a lot of tennis balls and all my recommendations changed from tennis related items to electronic devices."

Durham University

[Over time] G15:Encourage granular feedback. Enable the user to provide feedback indicating their preferences during regular interaction with the AI system.

Example Applications

[Social Networks, Product #2] "[The product] allows the user to "Hide an Ad," and then when doing so, solicits feedback to improve relevancy of future ads. "

[Music Recommenders, Product #1] "Love/dislike buttons are prominent and easily accessible."

[Email, Product #1] "The user can directly mark something as important, when the AI hadn't marked it as that previously."
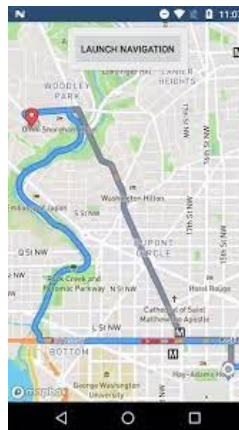
Example Violations

[Voice Assistants, Product #2] "Once [the assistant] performed the task I had asked of it, there was no additional ability to customize the experience or give feedback on my satisfaction; even when I chose to remove the reminder right after I verbally requested it."

[Navigation, Product #2] "The system does not do this. For example, if a user were to consistently pick a route home from work that was not the main route, then there might be patterns in traffic or circumstances that might be considered. If the user (instead of spouse) picked up a child on the way home from work because traffic was bad on the normal route, then the system should learn this habit, or allow the user to program it in... Doesn't let the user tweak the routes based on prior behavior."

[Over time] G16:Convey the consequences of user actions.

Immediately update or convey how user actions will impact future behaviors of the AI system

Example Applications

- **[Music Recommenders, Product #1]** "Tapping the like/dislike button results in immediate popups informing that the user will receive more/fewer recommendations like it."

- **[Social Networks, Product #2]** "[The product] communicates that hiding an Ad will adjust the relevancy of future ads."

- **[Web Search, Product #1]** "With different filters, the search results are auto updated."

Example Violations

- **[Social Networks, Product #1]** "You can unfollow or like or interact but how that affects you isn't clear. It just sorta happens."

- **[Email, Product #1]** "It's clear there is some consequence for the user actions, but it's not well-defined. So while it exists, the lack of clarity on the exact outcomes for behavior are unclear"

Durham University

[Over time] G17:Provide global controls. Allow the user to globally customize what the AI system monitors and how it behaves.

Example Applications

[Web Search. Product #2] "It has settings such as...private results that help users get results that are more relevant to them."

[Photo Organizers, Product #1] "[The product] allows users to turn on your location history so the AI can group photos by where you have been."

[Navigation, Product #2] "A few options to adjust my preferences for how I want to get directions. Uses common patterns for multi-select and single-choice.

Example Violations

- **[Email, Product #2]** "The only option is to turn the system on or off. It otherwise applies to all messages at all times. It is not clear how the system works, or what data it monitors, but presumably it applies to entire email contents."

- **[Music Recommenders, Product #2]** "[The product] does not provide a mechanism to turn off tracking of listening data (at least not in the process of building a playlist) or to impact what it learns."

- **[Navigation, Product #1]** "There is no apparent way to customize what information the system has access to or what it learns. Marked locations can be removed, but otherwise, I see no customizability with regard to data access or behavior."

[Over time] G18:Notify users about changes. Inform the user when the AI system adds or updates its capabilities.

Example Applications

- **[Email, Product #2]** "The help tab for the interface features a "What's new" section which could be used to inform the user about AI system additions or capability updates."

- **[Social Networks, Product #2]** "Updates of privacy and regulations...They do post their updated privacy or legal regulations and make the UI inaccessible until the user agrees."

- **[Navigation, Product #2]** "I don't have a way to show this, but it does provide small in-app teaching callouts for important new features. New features that require my explicit attention are pop-ups."

Example Violations

- **[Music Recommenders, Product #1]** "The algorithm feels like it constantly updates, with a slightly different feel to my recommendations every week. However there is no explanation of what has changed."

- **[Web Search, Product #2]** "there are no notifications or mentions when the search algorithm changes or new capabilities are added, such as disambiguation, the presentation of recipes directly, etc...Updates or changes are only noticeable with use (like new AI feature applications) while others are completely unknown (like changing the search algorithm to return more diverse CEO images)"

Durham University

Thank you!

Questions?

Durham
University