

Engineering for Ethics, Privacy, and Fairness in Computer Systems

CMSC 25910 (“CS 259”)

Spring 2024

The University of Chicago



Course Overview

Course Description

This course takes a **technical approach to understanding ethical issues** in the design and implementation of **computer systems**. Tensions often arise between a computer system's utility and its privacy-invasiveness, between its robustness and its flexibility, and between its ability to leverage existing data and existing data's tendency to encode biases. The course will demonstrate how computer systems can violate individuals' privacy and agency, impact sub-populations in disparate ways, and harm both society and the environment. It will also introduce algorithmic approaches to fairness, privacy, transparency, and explainability in machine learning systems. Through hands-on programming assignments and projects, students will design and implement computer systems that reflect both ethics and privacy by design. They will also wrestle with fundamental questions about who bears responsibility for a system's shortcomings, how to balance different stakeholders' goals, and what societal values computer systems should embed.

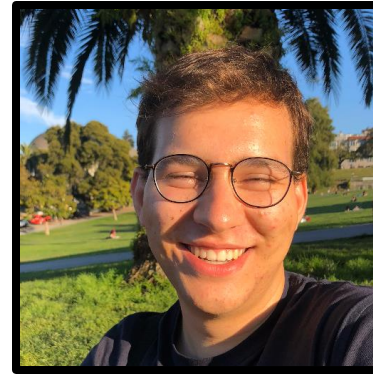
Course Staff



Blase
Ur



Alison
Gan



Arthur
Borém



Madison
Pickering

Office Hours

- Mondays 4:30pm – 5:30pm (Blase)
 - Tuesdays 1:00pm – 2:00pm (TA)
 - Tuesdays 3:00pm – 4:00pm (TA)
 - Wednesdays 4:30pm – 5:30pm (TA)
 - ...or by appointment
-
- All office hours will be held in **JCL 356**

Coursework / Grades

- **Lectures:** not recorded; **attendance required**
- **Programming assignments:** $8 \times 8\% = 64\%$
- **Reading responses:** $8 \times 1.5\% = 12\%$
- **Project:** 19%
- **Participation in lecture:** 5%

Communication

- **Course website:** Syllabus, schedule, readings
- **Canvas:** Distribution of assignments and reading responses
- **Canvas:** Submission of code/data for assignments
- **Gradescope:** Submission of prose (single PDF) for assignments and reading responses (paste into Gradescope)
- **Ed:** Questions, announcements, requests; **don't email us**
- **Gradescope:** Grades, regrade requests

Course Policies

- **P/F grading:** Post on Ed with “P/F request” by end of ninth week
- **Late submission:** Up to 24 hours late with -15 point penalty
 - Applies to both assignments and reading responses
 - Not permitted for project due to grading deadlines
- **Extensions:** Granted only for medical/personal emergencies
 - Not granted for other classes, interviews, travel, RSOs/sports
- **Wellness:** Reach out in a private Ed post to the course staff
 - Emails to Blase are also ok

Academic Integrity

- Detailed on syllabus
- All work submitted must be your own
- You may speak in general terms about approach, but not share code; **do not look at each other's code**
- Encouraged to talk to classmates and form study groups
- On each Gradescope submission, you **must document everyone in the class you spoke to, as well as every major resource you consulted** other than what we provide
- For **LLMs** (e.g., ChatGPT), you must include the **prompt** you used

Academic Integrity

- Example for the top of your Gradescope submission:
 - “I discussed the whole assignment with Jane Smith. We also discussed Part 3 with John Doe. I consulted:
<https://www.helpfuldomain.com/helpfulpage.html> to understand the fetch() API and I used two lines from *<https://www.other.com/page.html>* in Part 3.”
- Code reuse from websites, Stack Overflow, LLMs, and published resources only allowed if **all** of the following apply:
 - Around 4 lines of code or fewer
 - Doesn't solve the **intellectual point** of that part of the assignment
 - Documented at top (see above) or as comments
 - Exception: If the assignment itself points you directly to a tutorial, it's ok to reuse larger code fragments (with attribution)

Course Topics and Schedule

- <https://www.classes.cs.uchicago.edu/archive/2024/spring/25910-1/schedule.html>

CMSC 25910 vs. DATA 25900

CMSC 25910 (Blase): <https://classes.cs.uchicago.edu/archive/2024/spring/25910-1/>

Focus on computer systems / algorithms / engineering / implementation

8 programming assignments; 8 reading responses; project (19% of grade)

Multiple languages, primarily Python

(Highly similar to CMSC 25900 from Spring 2021)

DATA 25900 (Raul): <https://raulcastrofernandez.com/DATA-25900-Spring24/>

Focus on data science / data analysis

5 programming assignments; 7 reading responses; project (50% of grade)

Python using Jupyter Notebooks

Class Introduction

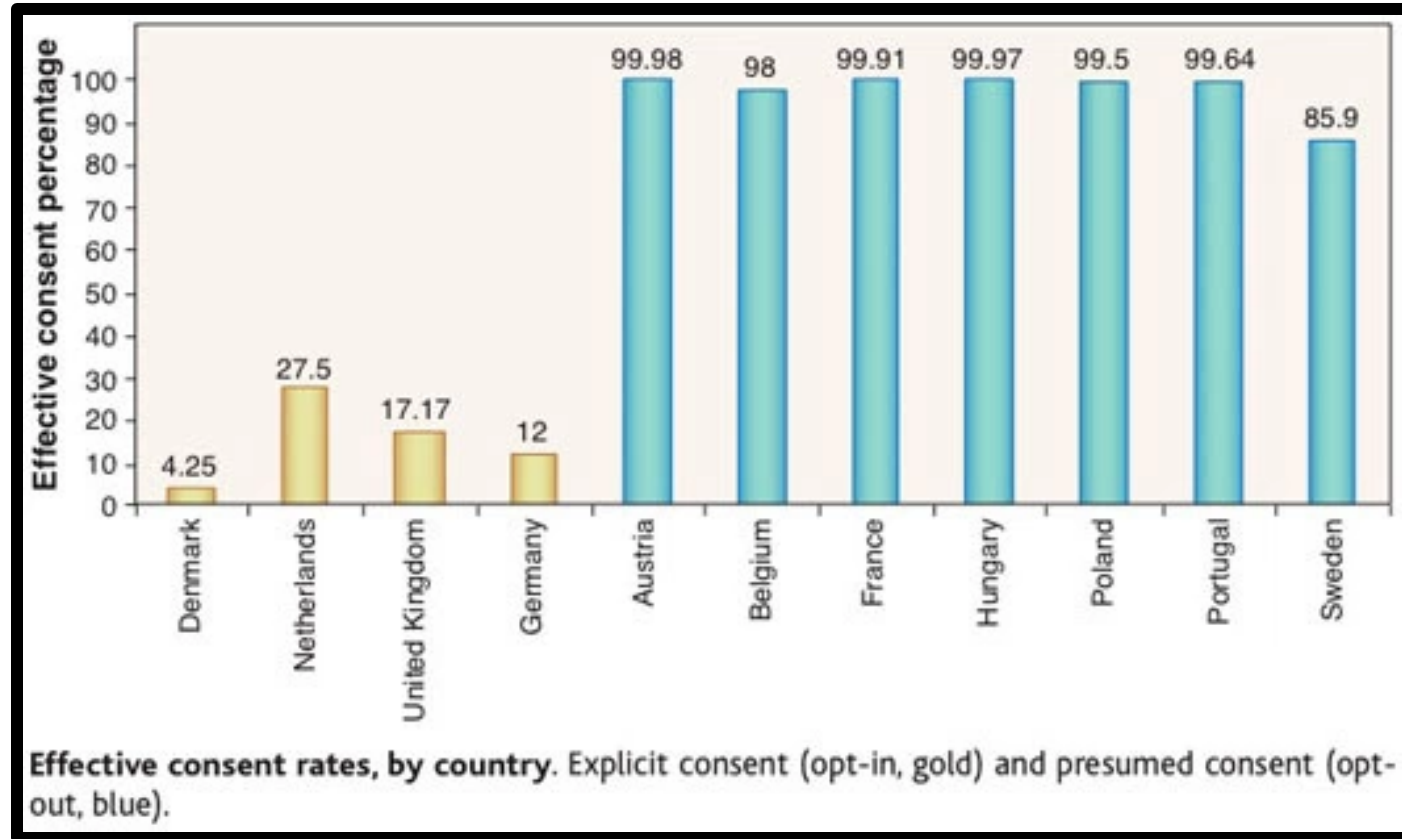
- 1) First name
- 2) Class year and major (e.g., third-year CS major)
- 3) Topic you are most interested in learning about

Responsible and Respectful User Interfaces

***Should a product have
an opinion?***

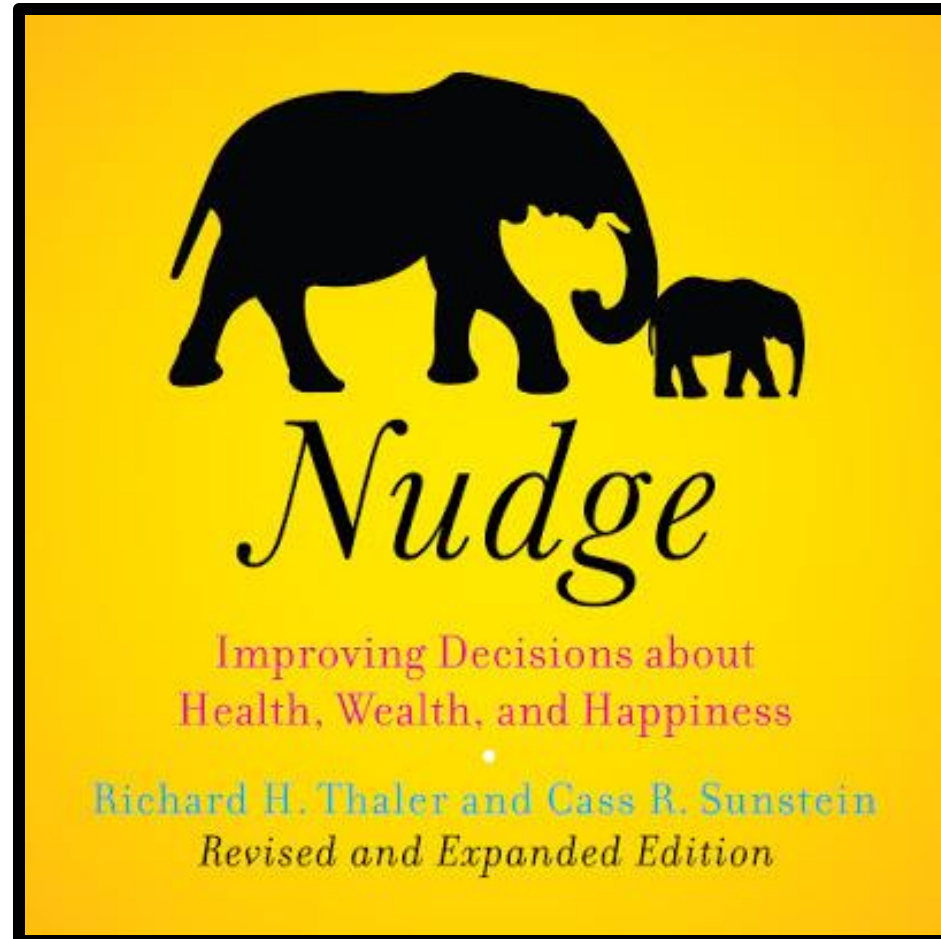
The power of defaults

The power of the default (organ donation)



Interfaces that influence

Behavioral nudging



Dark patterns (AKA Deceptive patterns)

Taxonomy from www.darkpatterns.org

- Trick questions
- Sneak into basket
- Roach motel
- Privacy Zuckering
- Price comparison prevention
- Misdirection
- Hidden costs
- Bait and switch
- Confirmshaming
- Disguised ads
- Forced continuity
- Friend spam

<https://twitter.com/darkpatterns> has many examples

Examples from www.darkpatterns.org



Steve Bromley
@Steve_Bromley

"No, I don't like savings"

	50% OFF 3 Months
Pro Monthly £11.99/mo <i>£143.88 billed annually</i>	Pro Monthly Plan + Discount £6.0/mo <i>£101.92 billed annually / per license</i> Save £ 17.99 per license
No, I don't like savings	Yes, Take Offer

Examples from www.darkpatterns.org



Benedict Evans ✓

@benedictevans

You could write a book about all the dark patterns Amazon uses to promote Prime.

amazon.co.uk

WELCOME ADDRESS ITEMS WRAP DISPATCH **PAY** CONFIRM

Benedict Evans, we're giving you immediate access to all Prime benefits with a 30-day FREE trial

Try 30 days of Prime for free. You can cancel anytime.

		Without Prime
Prime Video	Included	Unavailable
Amazon Music	Included	Unavailable
Fast delivery on Prime eligible items	FREE	From £4.99

[Order without Prime. Decline free trial](#)

Order Now with Prime

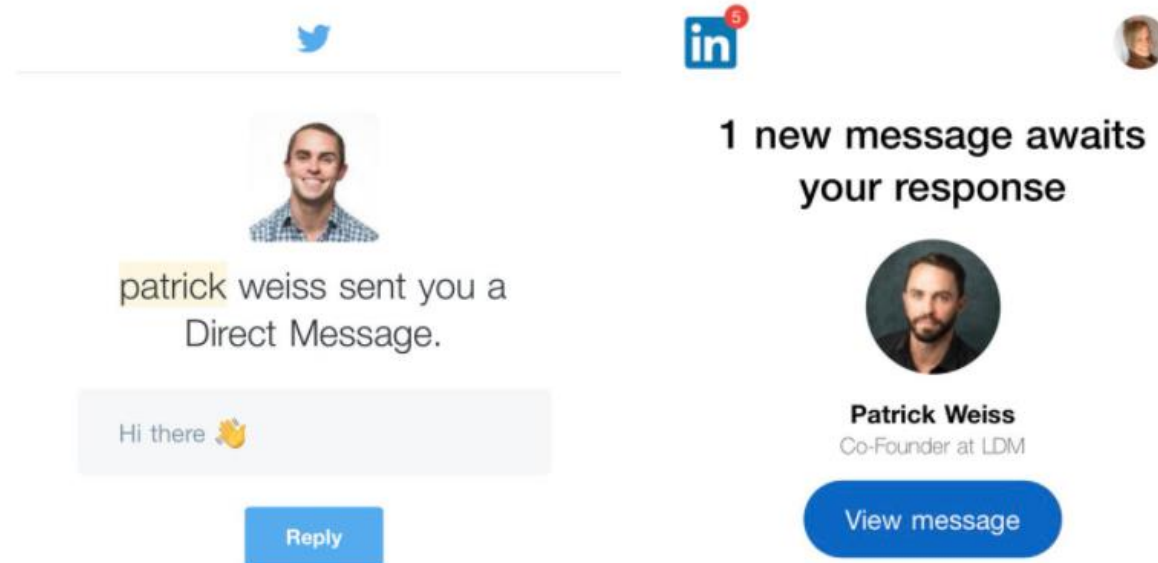
[Pay later](#)

Examples from www.darkpatterns.org



Katie Dill
@lil_dill

When engagement metrics drive the decision..
On the left, Twitter's email with the direct message text included.
On the right, LinkedIn's email forcing me to open the app to see the message. Drives me 🍌s every time!
Twitter, I appreciate you for putting the user first on this.



Examples from www.darkpatterns.org

The image is a composite of two screenshots. The left screenshot shows a 'CHRISTIE'S' privacy preference center. It has a title 'Privacy preference centre' and explanatory text about cookies. Below the text is an 'Allow all' button. Underneath is a section 'Manage consent preferences' with three rows: 'Functional cookies' with a toggle switch, 'Strictly necessary cookies' with the text 'Always active', and 'Performance cookies' with a toggle switch. At the bottom is a 'Confirm my choices' button and a 'Powered by OneTrust' logo. The right screenshot is a tweet from Adam Smith-Kipnis (@AdamSmithKipnis) dated 6:55 PM · Mar 18, 2021. The tweet text reads: '@darkpatterns Cookie consent preferences from @ChristiesInc on christies.com does not say whether grey on grey, or black on white, means approve or deny. No info means tracking isn't informed or consensual. This UI gives the illusion of consent. #darkpatterns'. The tweet has 2 retweets and 11 likes. Below the tweet is a reply from 'Dark Patterns' (@da...) dated Mar 19, replying to @AdamSmithKipnis and @ChristiesInc.

CHRISTIE'S

Privacy preference centre

When you visit any website, it may store or retrieve information on your browser, mostly in the form of cookies. This information might be about you, your preferences or your device and is mostly used to make the site work as you expect it to. The information does not usually directly identify you, but it can give you a more personalised web experience.

Because we respect your right to privacy, you can choose not to allow some types of cookies. Click on the different category headings to find out more and change our default settings. However, blocking some types of cookies may impact your experience of the site and the services we are able to offer. [More information](#)

Allow all

Manage consent preferences

+ Functional cookies

+ Strictly necessary cookies Always active

+ Performance cookies

Confirm my choices

Powered by OneTrust

Adam Smith-Kipnis
@AdamSmithKipnis

@darkpatterns Cookie consent preferences from @ChristiesInc on christies.com does not say whether grey on grey, or black on white, means approve or deny. No info means tracking isn't informed or consensual. This UI gives the illusion of consent. #darkpatterns

6:55 PM · Mar 18, 2021 · Twitter Web App

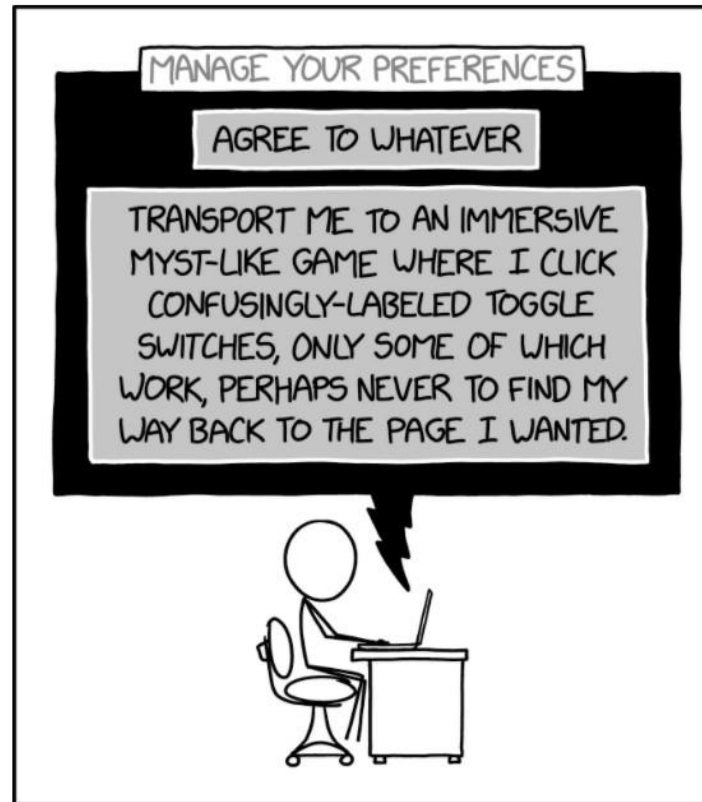
2 Retweets 11 Likes

Dark Patterns @da... · Mar 19
Replying to @AdamSmithKipnis and @ChristiesInc

Meta-example from xkcd

MANAGE YOUR PREFERENCES

|< < PREV RANDOM NEXT > >|



|< < PREV RANDOM NEXT > >|

Are dark patterns bad?

Security Dark Pattern



This Connection is Untrusted

You have asked Firefox to connect securely to **grey-dev.ece.cmu.edu**, but we can't confirm that your connection is secure.

Normally, when you try to connect securely, sites will present trusted identification to prove that you are going to the right place. However, this site's identity can't be verified.

What Should I Do?

If you usually connect to this site without problems, this error could mean that someone is trying to impersonate the site, and you shouldn't continue.

Get me out of here!

- ▶ **Technical Details**
- ▶ **I Understand the Risks**

Security Dark Pattern

Normally, when you try to connect securely, sites will present trusted identification to prove that you are going to the right place. However, this site's identity can't be verified.

What Should I Do?

If you usually connect to this site without problems, this error could mean that someone is trying to impersonate the site, and you shouldn't continue.

Get me out of here!

▼ Technical Details

grey-dev.ece.cmu.edu uses an invalid security certificate.

The certificate is not trusted because it is self-signed.

(Error code: sec_error_untrusted_issuer)

▼ I Understand the Risks

If you understand what's going on, you can tell Firefox to start trusting this site's identification. **Even if you trust the site, this error could mean that someone is tampering with your connection.**

Don't add an exception unless you know there's a good reason why this site doesn't use trusted identification.

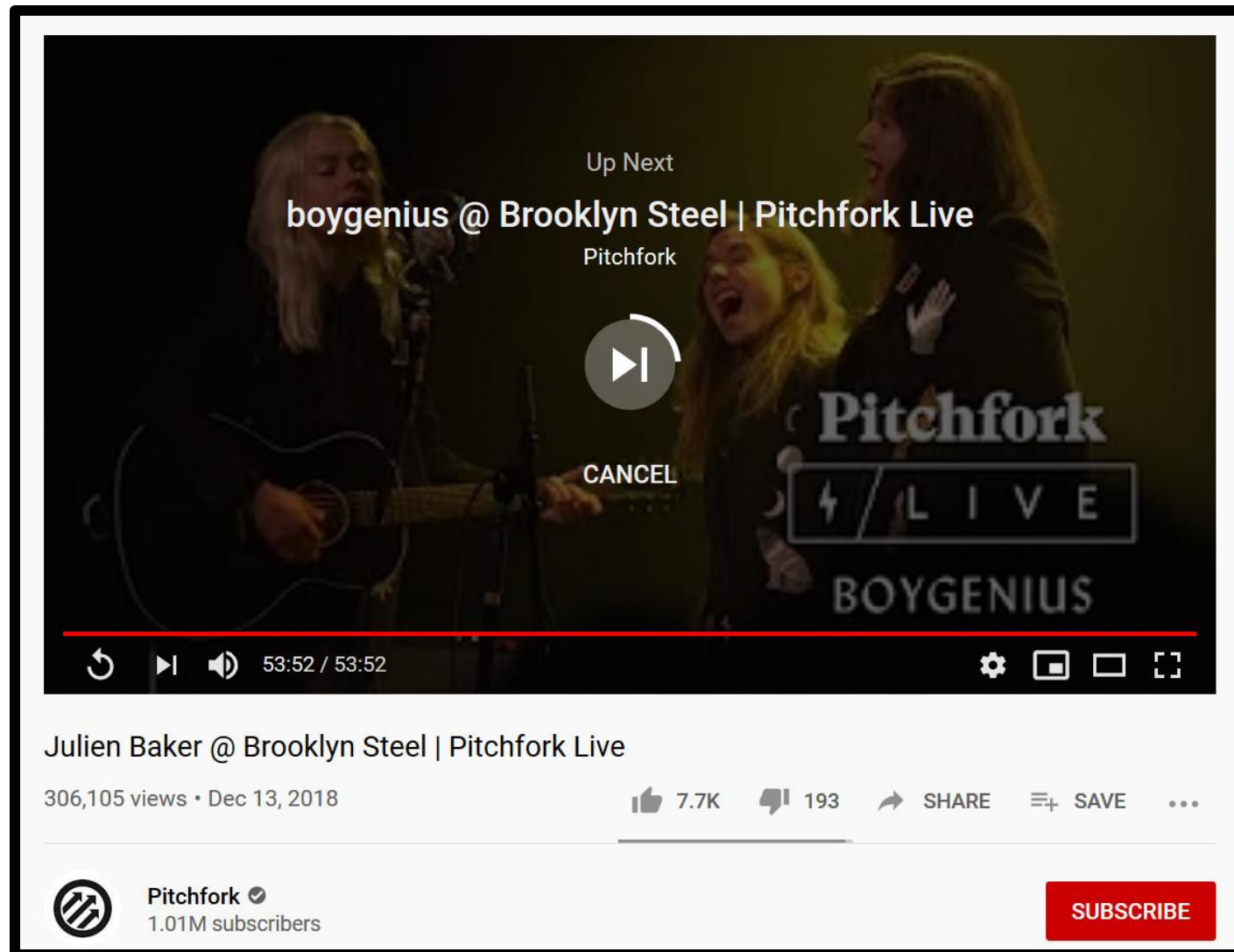
Add Exception...

Interfaces that incentivize behaviors

Ways interfaces incentivize

- Social approval (like counts / view counts / follower counts)
- Karma / leaderboards / Reddit gilding
- Creating a Fear Of Missing Out (FOMO)
- “Only 1 left in stock”
- Forced obsolescence or devices that seem to slow down
- Intentionally showing inaccurate measurements

Does auto-play capture intent? Addiction?



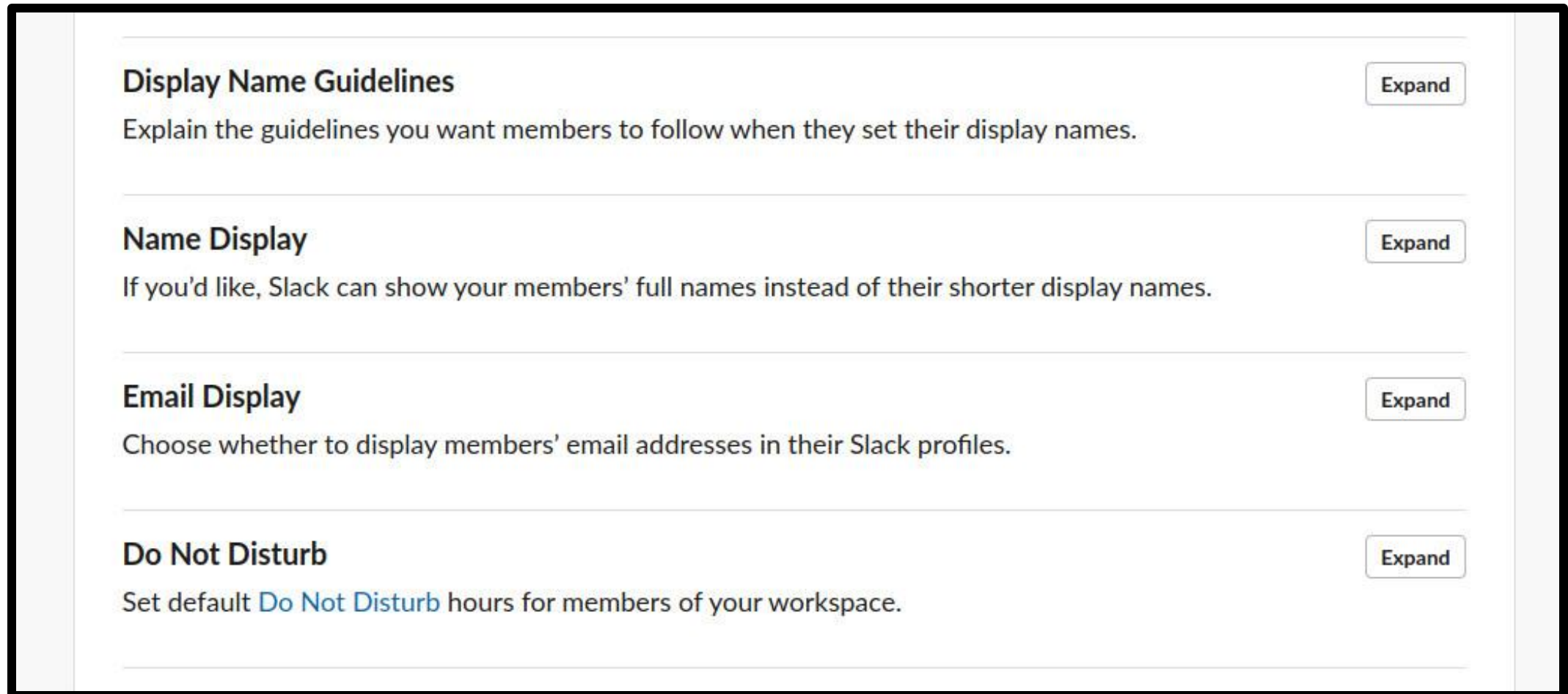
***Can we think about
data-relevant user
interfaces in terms of
power?***

Power in interfaces

- Who chooses what questions are there?
- Who chooses default options / settings / option ordering?
- Who gets to add data?
 - Facebook shadow profiles (for advertising)
 - Tag others or upload photos of others in social media
 - Grubhub, Postmates, and DoorDash add restaurants without permission (see <https://www.eater.com/2020/1/29/21113416/grubhub-seamless-kin-khao-online-delivery-mistake-doordash>)
- Who chooses which data is right?
 - Wikipedia editing wars and controversial articles

Who sets policies?

- Slack organizational policies about names

A screenshot of the Slack organizational policy settings interface. It features four distinct sections, each with a title, a descriptive paragraph, and an 'Expand' button on the right. The sections are: 'Display Name Guidelines' (explaining guidelines for display names), 'Name Display' (offering to show full names instead of shorter display names), 'Email Display' (allowing choice to display email addresses), and 'Do Not Disturb' (allowing setting of default DND hours).

Policy Name	Description	Action
Display Name Guidelines	Explain the guidelines you want members to follow when they set their display names.	Expand
Name Display	If you'd like, Slack can show your members' full names instead of their shorter display names.	Expand
Email Display	Choose whether to display members' email addresses in their Slack profiles.	Expand
Do Not Disturb	Set default Do Not Disturb hours for members of your workspace.	Expand

Respecting time and away-from-screen time

Expectations about types of notifications

- Email vs. text vs. in-app notification
- Are notifications *push* or *pull*?
- Does your device make a sound? Does it buzz?
- Is there an unread number on your screen?



Slack pausing notifications

Pause notifications with Do Not Disturb

If you need focus time or time away from work, you can pause your [notifications in Slack](#) with Do Not Disturb (DND). Use DND as needed to help you concentrate, and even set a DND schedule for routine times you'd prefer not to be notified.

How does DND work?

- All notifications and @mentions will be paused. Once you resume notifications, you can review everything you received while in DND.
- Members will see a 🚫 **Do Not Disturb icon** next to your name. When sending a direct message, they can override DND to notify you of urgent messages.

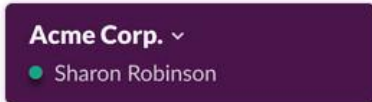
Pause notifications

Desktop

iOS

Android

Turn on DND

- 1 From your desktop, click your workspace name in the top left.

- 2 Hover over **Pause notifications**.
- 3 Choose a time frame from the menu or select **Custom** to set your own.

Slack scheduling notification pauses

Set a DND schedule

For the times you're routinely offline, you can set a Do Not Disturb schedule. With a schedule in place, your notifications will pause every day between the hours you've specified. Keep in mind that Workspace Owners and Admins can [set a default Do Not Disturb schedule](#) for their members.

Desktop

iOS

Android

1

From your desktop, click your workspace name in the top left.

Acme Corp. ▾

● Sharon Robinson

2

Hover over **Pause notifications**, then select **Do Not Disturb schedule** from the menu.

3

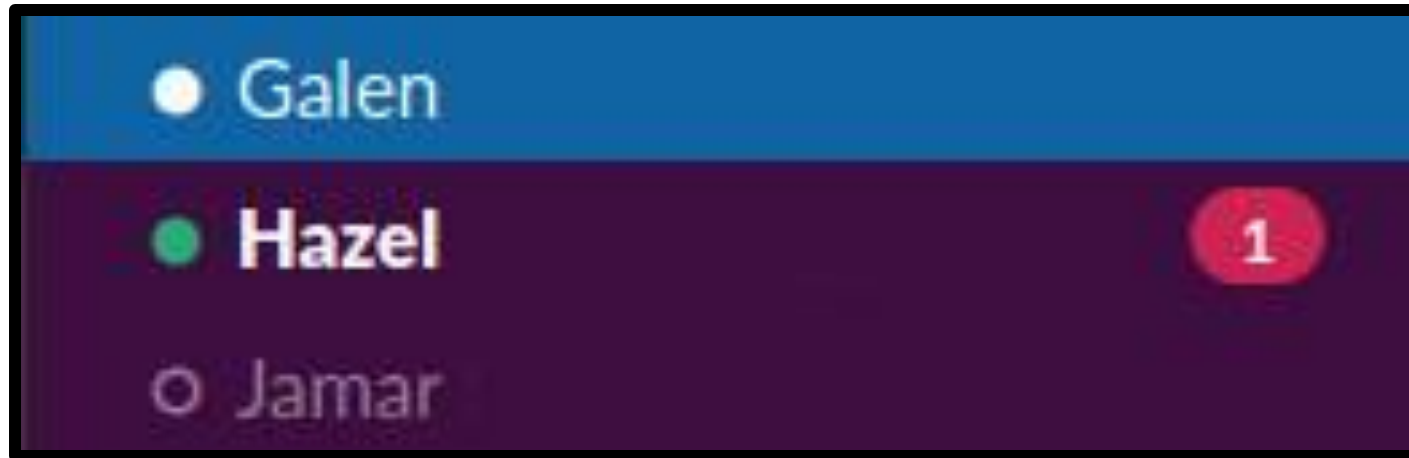
Under **Do Not Disturb**, check the box next to **Automatically disable notifications** and select a start and end time. Your settings will automatically save, so click the **✕ close icon** when you're done.



Tip: To set yourself to DND over the weekend, pause your notifications until the time and date you'll be back at work. There isn't currently a way to automatically pause notifications every weekend.

Online status indicators (OSIs)

- Note that users and Slack organizations can disable OSIs



Visibility of online status / audit logs

- Organizations might want to know who is clocking in/out
- Zoom lets meeting hosts know when it is backgrounded