

Data

CS 579 Online Social Network Analysis

Dr. Cindy Hood
8/26/25

Exams and Final Project Poster Presentation

- ▶ Exam 1 - Oct 9 in class
- ▶ Exam 2 - Dec 2 in class
- ▶ Final Project Poster Session - Dec 4 in class
- ▶ Online students (sections 2 and 3) will have remote options


Teaching Assistants

- ▶ Siva Krishna Golla
 - ▶ sgolla2@hawk.illinoistech.edu
- ▶ Khush Dhiren Patel
 - ▶ kpatel210@hawk.illinoistech.edu
- ▶ Aswith Sama
 - ▶ asama@hawk.illinoistech.edu
- ▶ Office hours tbd

Logistics - Extra Credit

- ▶ Occasionally I will make extra credit assignments available during class
 - ▶ For main campus student, these assignment will be in-class
 - ▶ Please bring paper and a pen/pencil to each class
 - ▶ Online students will be given an opportunity to complete these assignments after watching the lecture
- ▶ These assignments will count as extra credit for up to 5% of your grade (I will decide depending on how many assignments I give out)
- ▶ These assignments will not be posted in the assignment section of Canvas
 - ▶ In lecture only
- ▶ No deadline extensions possible

Data



DictionaryThesaurus

data

×

Q

Games & QuizzesWord of the Day

Dictionary

Definition

SynonymsExample SentencesWord HistoryPhrases ContainingRelated ArticlesEntries NearShow More

Save Word

data

noun

da·ta 'dā-tə 'da- also 'dä-

plural in form but singular or plural in construction

often attributive

[Synonyms of data](#)

1

: factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation

the *data* is plentiful and easily available

— H. A. Gleason, Jr.

comprehensive *data* on economic growth have been published

— N. H. Jacoby

2

: information in digital form that can be transmitted or processed

3

: information output by a sensing device or organ that includes both useful and irrelevant or [redundant](#) information and must be processed to be meaningful

Let's say we are going to build a social network of everyone in the class

- ▶ What information would be useful?
- ▶ How do you feel about providing information for this social network?
- ▶ Did the fact that you will be providing information change how you thought about the info provided?

The Point of Collection by

**MIMI
ONUOHA**

- ▶ <https://datasociety.net/points/the-point-of-collection/>

The conceptual, practical, and ethical issues surrounding “big data” and data in general begin at the very moment of data collection. Particularly when the data concern people, not enough attention is paid to the realities entangled within that significant moment and spreading out from it.

The Point of Collection by

MIMI
ONUOHA

1. Data sets are the results of their means of collection.

It's easy to forget that the people collecting a data set, and how they choose to do it, directly determines the data set.

An illustrative example can be found in the statistics for how many hate crimes were committed in the United States in 2012. According to the FBI Uniform Crime Reporting Program (UCR), the number was 5,796. However, the Department of Justice's Bureau of Statistics reported 293,800 hate crimes.

The reason for the variation was simple. The UCR gathers data that is voluntarily reported by law enforcement agencies across the country. The Bureau of Statistics, on the other hand, distributes the National Crime Victimization Survey, which collects data from the victims of hate crimes. The result is a more transparent and inclusive surveying.

Same data set, two different means of collection, two wildly different results. What they show is an important fact we must keep in mind: There's no pure objectivity encoded into data sets. Each one is the result of a number of human processes and decisions that affect, in a variety of ways, the data that they aim to report. In this sense, the moment of data collection starts before any data is actually produced.

The Point of Collection by

MIMI
ONUOHA

2. As we collect more data, we prioritize things that fit patterns of collection.

Or as Rob Kitchin and Martin Dodge say in Code/Space, “The effect of abstracting the world is that the world starts to structure itself in the image of the capta and the code.” Data emerges from a world that is increasingly software-mediated, and software thrives on abstraction. It flattens out individual variations in favor of types and models.

As we abstract the world, we prioritize abstractions of the world. The more we look to data to answer our big questions (in areas like policing, safety, and security), the more incentives we have to shape the world into an input that fits into an algorithm. Our need to generate things that feed a model rings true even in cases where the messy bounds of experiences can’t be neatly categorized into bits and bytes, or easily retrieved from tables through queries.

Biometric data is a great example of this. Fingerprint authentication technologies and iris scanners point to a system where individuals are uniquely identified through metrics and data. In order for this to work, people themselves have to be conceptualized more and more as machine-readable.

The Point of Collection by

MIMI
ONUOHA

3. Data sets outlive the rationale for their collection.

Spotify can come up with a list of reasons why having access to users' photos, locations, microphones, and contact lists can improve the music streaming experience. But the reasons why they decide these forms of data might be useful can be less important than the fact that they have the data itself. This is because the needs or desires influencing the decisions to collect some type of data often eventually disappear, while the data produced as a result of those decisions have the potential to live for much longer. The data are capable of shifting and changing according to specific cultural contexts and to play different roles than what they might have initially been intended for.

Ultimately, the question of intention behind the collection or generation of a data set can be rendered irrelevant. Thinking through the moment of collection can reveal the distance between it and the data's use. And it's often far more critical to think about the potentials and possibilities surrounding what can be done with collected data.

The Point of Collection by

MIMI
ONUOHA

4. Corollary: Especially combined, data sets reveal far more than intended.

We sometimes fail to realize that data sets, both on their own and combined with others, can be used to do far more than what they were originally intended for. You can make inferences from one data set that result in conclusions in completely different realms. Facebook, by having huge amounts of data on people and their networks, could make reasonable hypotheses regarding people's sexual orientations.

People who work with data know this intimately, but it can often be difficult to see the connections between the collection of one thing and the inference of something else. Unfortunately, the effects of these connections can become very strongly felt. As Bruce Schneier puts it, “data we’re willing to share can imply conclusions that we don’t want to share.”

The Point of Collection by MIMI ONUOHA

5. Data collection is a transaction that is the result of an invisible relationship.

This is a frame — connected to my first point — useful for understanding how to think about data collection on the whole:

Every data set involving people implies subjects and objects, those who collect and those who make up the collected. It is imperative to remember that on both sides we have human beings. I point this out not for any fluffy reasons related to humanism or human-centered design, but because power arises out of hierarchies, interactions, and dynamics. The below-the-surface work of a particular data set is joined to the reasons and means that created it and the relationships running through those reasons and means. If we can keep that in mind, we're better positioned to see data as an intermediate result, one piece in a larger process, something that is as much human-oriented as it is systematic. The challenge is for us to keep in mind both aspects of data collection, to see systematic as well as human tensions and biases.

The Point of Collection by

**MIMI
ONUOHA**

The point of data collection is a unique site for unpacking change, abuse, unfairness, bias, and potential. We can't talk about responsible data without talking about the moment when data becomes data.

MIMI ỌNỤỌHA

work

about

contact

writing

Archives Say The
Subtext Out Loud
(ongoing)

The Hair in the Cable
(2021 - ongoing)

Computational
Violences (2017-19)

Us, Aggregated
(2017-19)

Missing Data
(2015-2022)

Older



OFFICIAL

Nigerian-American artist Mimi Ọnụọha's work questions and exposes the contradictory logics of technological progress. Through print, code, data, video, installation, and archival media, Ọnụọha offers new orientations for making sense of the seeming absences that define systems of labor, ecology and relations.

Ọnụọha's work has been featured at the Whitney Museum of Art (USA), the Australian Centre for Contemporary Art (AUS), Mao Jihong Arts Foundation (China), La Gaitê Lyrique (France), Gropius Bau (Germany), The Photographers Gallery (UK), and the Atlanta Contemporary (USA) among others. Her public art engagements have been supported by Akademie der Kunst (Germany), the Royal College of Art (UK), the Rockefeller Foundation (USA), and Princeton University (USA). Her work is in the collections of the Victoria and Albert Museum (UK).

Ọnụọha is a Creative Capital and Fulbright-National Geographic grantee. She is also the Co-Founder of A People's Guide To Tech, an artist-led organization that makes educational guides and workshops about emerging technology.

<https://mimionuoha.com/>



Archived

In/visible: Missing Datasets



Watch the event

About the Event:

We are living in a time of unprecedented levels of data collection. Everything around us is being tracked, recorded, and documented constantly, and in ways we may not even fully grasp. But in the giant “data ecosystem” where this information goes to live, there exist certain spaces that are curiously empty. We’ve gotten so accustomed to the ubiquity of data that these “missing datasets” leave us scratching our heads. Enter: **Mimi Onuoha**. There’s always a reason behind it, she says, when certain information is being collected and saved while other information is not. Join Onuoha as she reveals common patterns behind these data absences; ways missing datasets have been used systematically to disadvantage certain groups; and ways in which data can be used as a measure for change.

Event Details

festivals

Science and Technology

In/visible: Missing Datasets

Saturday, Nov 10, 2018 1:00 pm - 2:00 pm
CT

[Venue SIX10](#)

--- -- -- -- --

Spotify data - 2015

W I R E D

NEWSLETTERS

GORDON GOTTSEGEN

SECURITY AUG 20, 2015 4:56 PM

You Can't Do Squat About Spotify's Eerie New Privacy Policy (UPDATED)

A new Spotify privacy policy shows that you don't have any.

Spotify Wants to Go Through Your Phone

"With your permission, we may collect information stored on your mobile device, such as contacts, photos, or media files. Local law may require that you seek the consent of your contacts to provide their personal information to Spotify, which may use that information for the purposes specified in this Privacy Policy." – Spotify

Spotify Wants to Know Where You're Going

"Depending on the type of device that you use to interact with the Service and your settings, we may also collect information about your location based on, for example, your phone's GPS location or other forms of locating mobile devices (e.g., Bluetooth). We may also collect sensor data (e.g., data about the speed of your movements, such as whether you are running, walking, or in transit)." – Spotify

Spotify Wants to Be Your Facebook Friend

"You may integrate your Spotify account with Third Party Applications. If you do, we may receive similar information related to your interactions with the Service on the Third Party Application, as well as information about your publicly available activity on the Third Party Application. This includes, for example, your "Like"s and posts on Facebook." – Spotify

<https://www.wired.com/2015/08/cant-squat-spotifys-eerie-new-privacy-policy/>

SEARCH

FORTUNE

SUBSCRIBE NOW

SIGN IN

Spotify updates privacy policy to clarify how it collects your data



BY ROBERT HACKETT

September 4, 2015 at 9:49 AM EDT



<https://fortune.com/2015/09/04/spotify-privacy-policy-update/>

Spotify Data - 2025

The New York Times *We Are Tech Privacy Reporters. Our Music Habits Got Doxxed.*

The “Panama Playlists” exposed the Spotify listening habits of some famous people — and two journalists who didn’t know as much about protecting their privacy as they had thought.



By Mike Isaac and Kashmir Hill

Mike Isaac is a metalhead who covers Silicon Valley. Kashmir Hill is a technology reporter who alternates between listening to ambient tunes and that one Taylor Swift song.

Published Aug. 24, 2025 Updated Aug. 25, 2025, 12:03 p.m. ET

https://www.nytimes.com/2025/08/24/technology/spotify-panama-playlists-privacy.html?unlocked_article_code=1.g08.31sl.9oCKU7XT2EZH&smid=em-share

Doxing

Article [Talk](#)

From Wikipedia, the free encyclopedia

"Doxes" redirects here. For the singular of doxes, see [Dox \(disambiguation\)](#).

Doxing or **doxxing** is the act of publicly providing [personally identifiable information](#) about an individual or organization, usually via the [Internet](#) and without their consent.^[1]^[2]^[3] Historically, the term has been used to refer to both the aggregation of this information from public databases and [social media](#) websites (like [Facebook](#)), and the publication of previously private information obtained through criminal or otherwise fraudulent means (such as [hacking](#) and [social engineering](#)).^[citation needed]

The aggregation and provision of previously published material is generally legal, though it may be subject to laws concerning [stalking](#) and [intimidation](#).^[4] Doxing may be carried out for reasons such as [online shaming](#), [extortion](#), and [vigilante](#) aid to law enforcement.^[5]^[6]

<https://en.wikipedia.org/wiki/Doxing>

PANAMA PLAYLISTS

I found the real Spotify accounts of celebrities, politicians, and journalists. Many use their real names. With a little investigating, I could say with near-certainty: yep, this is that person.

I've been scraping their playlists for over a year. Some individuals even have a setting enabled that displays their last played song. I scraped this continuously, so I know what songs they played, how many times, and when.

The Panama Papers revealed hidden bank accounts. This reveals hidden tastes.

Not affiliated with Spotify or any of the featured individuals. Album art and song previews from iTunes.
Published July 29, 2025.

All

Tech

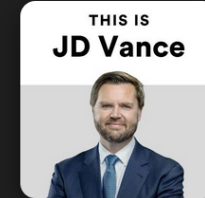
Politics

Media

Journalists

JD Vance

U.S. Vice President



From his playlist, **Making Dinner**



I Want It That Way
Backstreet Boys



One Time
Justin Bieber

From his playlist, **Gold On The Ceiling**



What Makes You Beautiful
One Direction



San Francisco
The Mowgli's



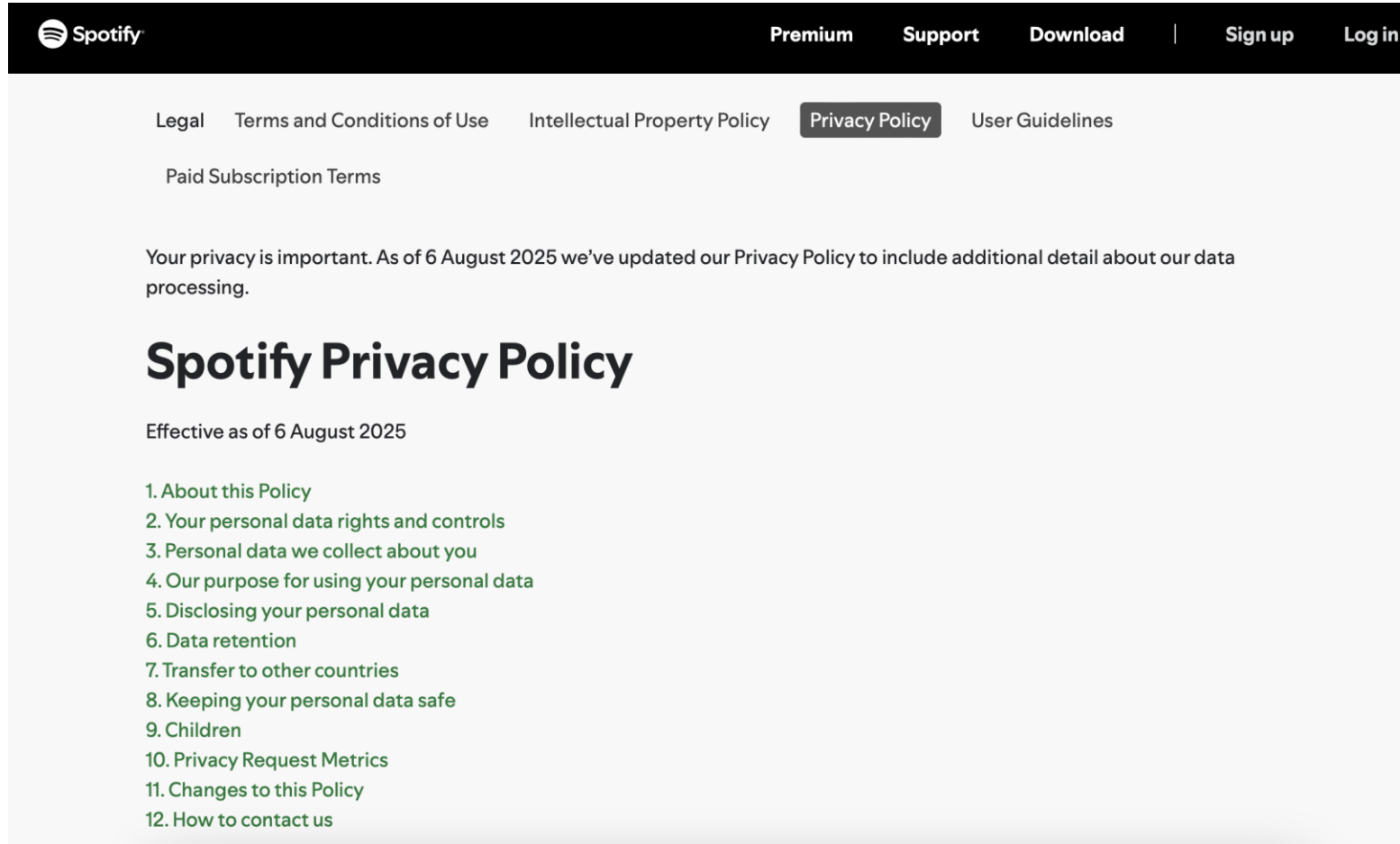
Get Lucky
Daft Punk, Pharrell Williams, Nile Ro...



You Are a Tourist
Death Cab for Cutie

<https://panamaplaylists.com/>

Spotify Privacy Policy

A screenshot of the Spotify Privacy Policy page. The top navigation bar is black with the Spotify logo on the left and links for Premium, Support, Download, Sign up, and Log in on the right. Below this, a secondary navigation bar contains links for Legal, Terms and Conditions of Use, Intellectual Property Policy, Privacy Policy (which is highlighted with a dark background), and User Guidelines. Under the Privacy Policy link, there is a link for Paid Subscription Terms. The main content area has a heading 'Your privacy is important. As of 6 August 2025 we've updated our Privacy Policy to include additional detail about our data processing.' followed by the title 'Spotify Privacy Policy' and the text 'Effective as of 6 August 2025'. A list of 12 items is provided, detailing the policy's structure.

Spotify

Premium Support Download | Sign up Log in

Legal Terms and Conditions of Use Intellectual Property Policy **Privacy Policy** User Guidelines

[Paid Subscription Terms](#)

Your privacy is important. As of 6 August 2025 we've updated our Privacy Policy to include additional detail about our data processing.

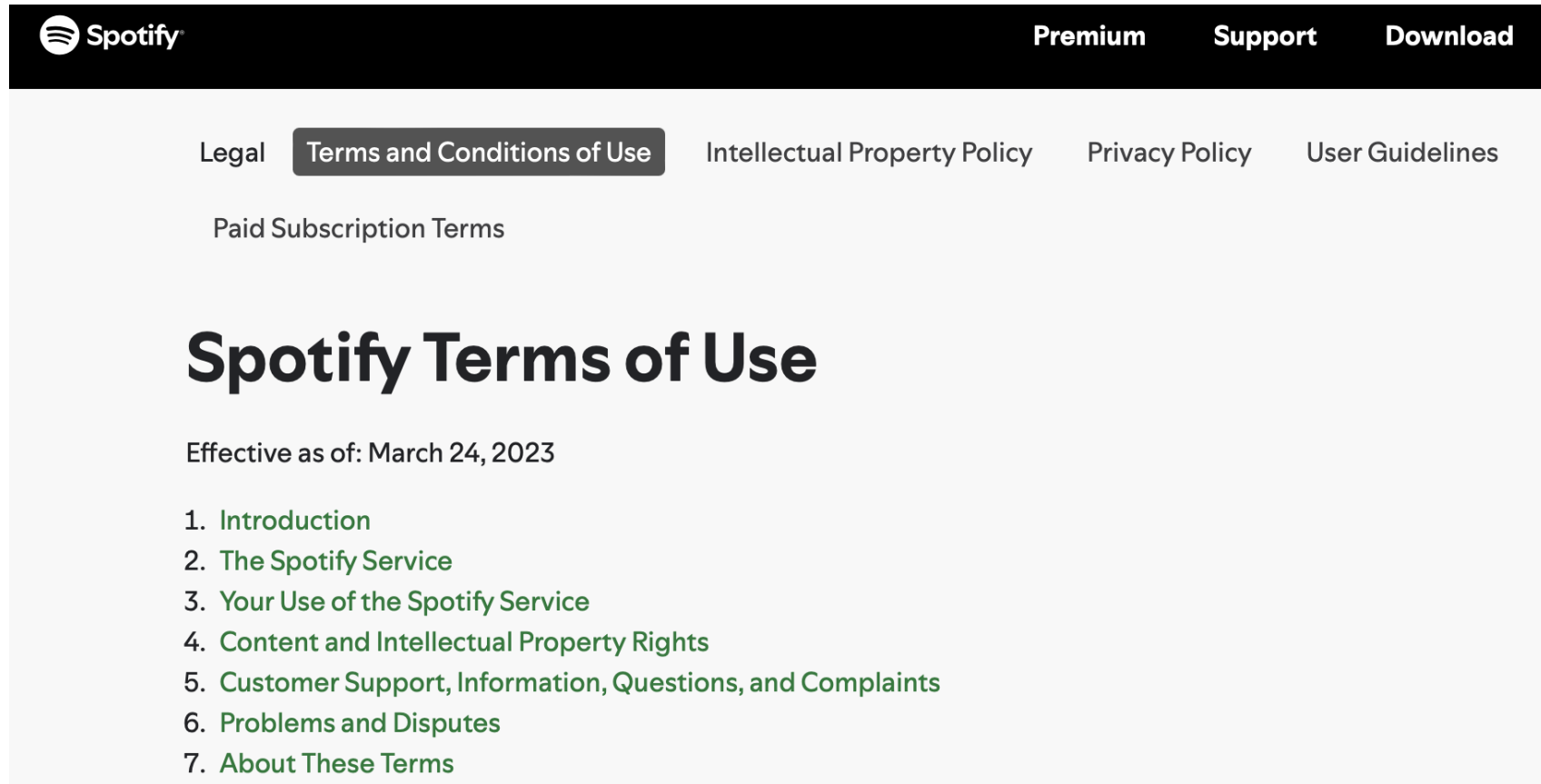
Spotify Privacy Policy

Effective as of 6 August 2025

1. About this Policy
2. Your personal data rights and controls
3. Personal data we collect about you
4. Our purpose for using your personal data
5. Disclosing your personal data
6. Data retention
7. Transfer to other countries
8. Keeping your personal data safe
9. Children
10. Privacy Request Metrics
11. Changes to this Policy
12. How to contact us

<https://www.spotify.com/us/legal/privacy-policy/>

Spotify Terms and Conditions of Use

A screenshot of the Spotify website's legal page. The top navigation bar is black with the Spotify logo on the left and links for Premium, Support, and Download on the right. Below this, a secondary navigation bar contains links for Legal, Terms and Conditions of Use (which is highlighted with a dark background), Intellectual Property Policy, Privacy Policy, and User Guidelines. Under the Legal link, there is a link for Paid Subscription Terms. The main heading is "Spotify Terms of Use" in a large, bold, black font. Below the heading, it states "Effective as of: March 24, 2023". A list of seven sections follows, each preceded by a number and a green dot: 1. Introduction, 2. The Spotify Service, 3. Your Use of the Spotify Service, 4. Content and Intellectual Property Rights, 5. Customer Support, Information, Questions, and Complaints, 6. Problems and Disputes, and 7. About These Terms.

Spotify

Premium Support Download

Legal **Terms and Conditions of Use** Intellectual Property Policy Privacy Policy User Guidelines

Paid Subscription Terms

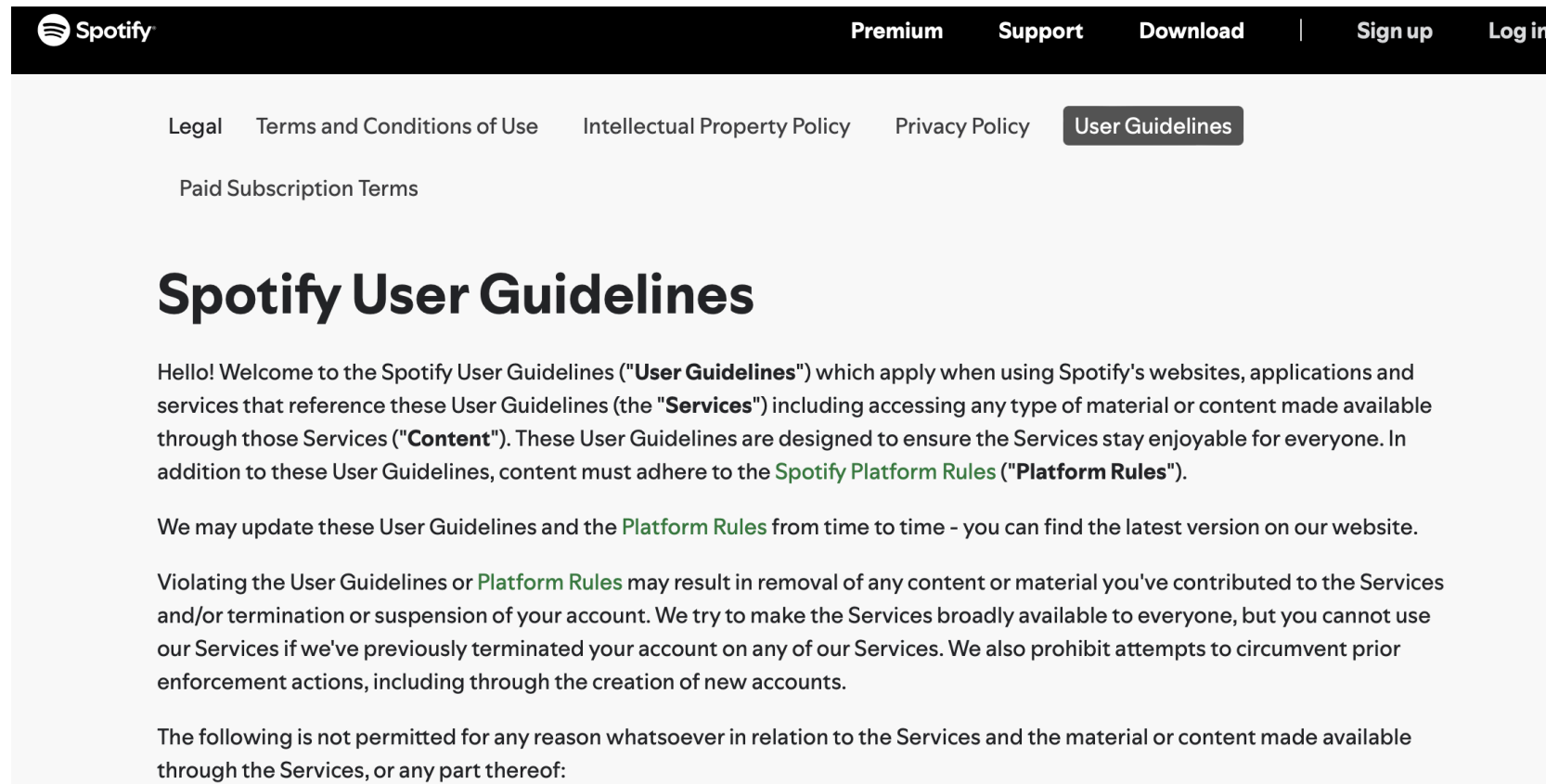
Spotify Terms of Use

Effective as of: March 24, 2023

1. Introduction
2. The Spotify Service
3. Your Use of the Spotify Service
4. Content and Intellectual Property Rights
5. Customer Support, Information, Questions, and Complaints
6. Problems and Disputes
7. About These Terms

<https://www.spotify.com/us/legal/end-user-agreement/>

Spotify User Guidelines

A screenshot of the Spotify User Guidelines page. The top navigation bar is black with the Spotify logo on the left and links for Premium, Support, Download, Sign up, and Log in on the right. Below this, a secondary navigation bar contains links for Legal, Terms and Conditions of Use, Intellectual Property Policy, Privacy Policy, and User Guidelines (which is highlighted with a dark background). Under the User Guidelines link, there is a link for Paid Subscription Terms. The main content area has a large heading 'Spotify User Guidelines' followed by a welcome message. The message explains that the guidelines apply to Spotify's websites, applications, and services, and are designed to ensure the services are enjoyable for everyone. It also mentions that content must adhere to the Spotify Platform Rules. The text continues with a statement about updates to the guidelines and rules, and a warning about the consequences of violating them, including removal of content, account termination, or suspension. Finally, it states that the following actions are not permitted: "crawling" or "scraping" for automated means, using bots or scrapers to collect information, or using the services or content to train a machine learning or AI model.

Spotify Premium Support Download | Sign up Log in

Legal Terms and Conditions of Use Intellectual Property Policy Privacy Policy **User Guidelines**

Paid Subscription Terms

Spotify User Guidelines

Hello! Welcome to the Spotify User Guidelines ("**User Guidelines**") which apply when using Spotify's websites, applications and services that reference these User Guidelines (the "**Services**") including accessing any type of material or content made available through those Services ("**Content**"). These User Guidelines are designed to ensure the Services stay enjoyable for everyone. In addition to these User Guidelines, content must adhere to the [Spotify Platform Rules](#) ("**Platform Rules**").

We may update these User Guidelines and the [Platform Rules](#) from time to time - you can find the latest version on our website.

Violating the User Guidelines or [Platform Rules](#) may result in removal of any content or material you've contributed to the Services and/or termination or suspension of your account. We try to make the Services broadly available to everyone, but you cannot use our Services if we've previously terminated your account on any of our Services. We also prohibit attempts to circumvent prior enforcement actions, including through the creation of new accounts.

The following is not permitted for any reason whatsoever in relation to the Services and the material or content made available through the Services, or any part thereof:

5. "crawling" or "scraping", whether manually or by automated means, or otherwise using any automated means (including bots, scrapers, and spiders), to view, access or collect information, or using any part of the Services or Content to train a machine learning or AI model or otherwise ingesting Spotify Content into a machine learning or AI model;

<https://www.spotify.com/us/legal/user-guidelines/>

Web scraping



← Back to all posts

Law and compliance

Thought leadership

Is web scraping legal?

Web scraping is legal if you scrape data that is publicly available on the internet. But you should be careful when scraping personal data or intellectual property. We cover the confusion surrounding the legality of web scraping and give you tips for compliant and ethical scrapers.

May 26, 2025 by  Ondra Urban

<https://blog.apify.com/is-web-scraping-legal/#>

Extra Credit #2

- ▶ To be done individually or in pairs (max group size = 2)
- ▶ Read all of the articles/posts cited on slides 17-23
- ▶ Write a concise opinion piece (1 page max) **discussing** one or more of the following questions:
 - ▶ Is Spotify's reaction to the Panama Playlists appropriate given their own collection of data?
 - ▶ Given Spotify's policies, is the web scraping that was done to create the Panama Playlists problematic?
 - ▶ Does Spotify have a responsibility to protect user playlists?
 - ▶ What impact might the Panama Playlists have on Spotify and/or Spotify users
- ▶ Submit a pdf of your opinion piece to [Extra Credit #2 - Spotify Opinion Piece](#)
- ▶ For section 1 students submit by midnight Wed 8/27
- ▶ For section 2 and 3 students submit by midnight Thurs 8/28

Chicago Humanities

Fall 2025 Festival

Fall Festival Day Schedule

Pilsen / Little Village Day – Sunday, September 21

Illinois Tech Day – Saturday, October 4

UChicago Arts & Humanities Day – Saturday, October 18

[Explore Our Calendar Of Events](#)

<https://www.chicagohumanities.org/>