



**KRESSIN**  
LAW GROUP

# Digital Platforms' Crawling and Use of Publisher Content

December 2023

# Crawling & Indexing: Core Concepts

## Crawling

Software programs known as crawlers or spiders systematically explore and copy the content from web pages across the internet, following links from page to page to discover and retrieve content

## Indexing

The process of organizing the vast amount of content collected during crawling, categorizing and cataloging web pages' textual, visual, and multimedia elements, so search engines (and other technologies) can quickly retrieve and utilize the information

## Googlebot

Google's web crawler, a sophisticated computer program that follows links on the web, visits web pages, and copies their content for inclusion in Google's search index (and for use in other Google products)



# Robots.txt: How Publishers Control Access

## Robots.txt

A text file placed in a website's root directory that provides instructions to web crawlers, specifying which areas of the site they are allowed or disallowed to crawl and index. It helps website owners control how search engines access and interact with their site's content.

The example to the right gives access to Googlebot but blocks all other crawlers

**Important:** Robots.txt lets publishers control *whether* their content is scraped, not *how* their content is *used*

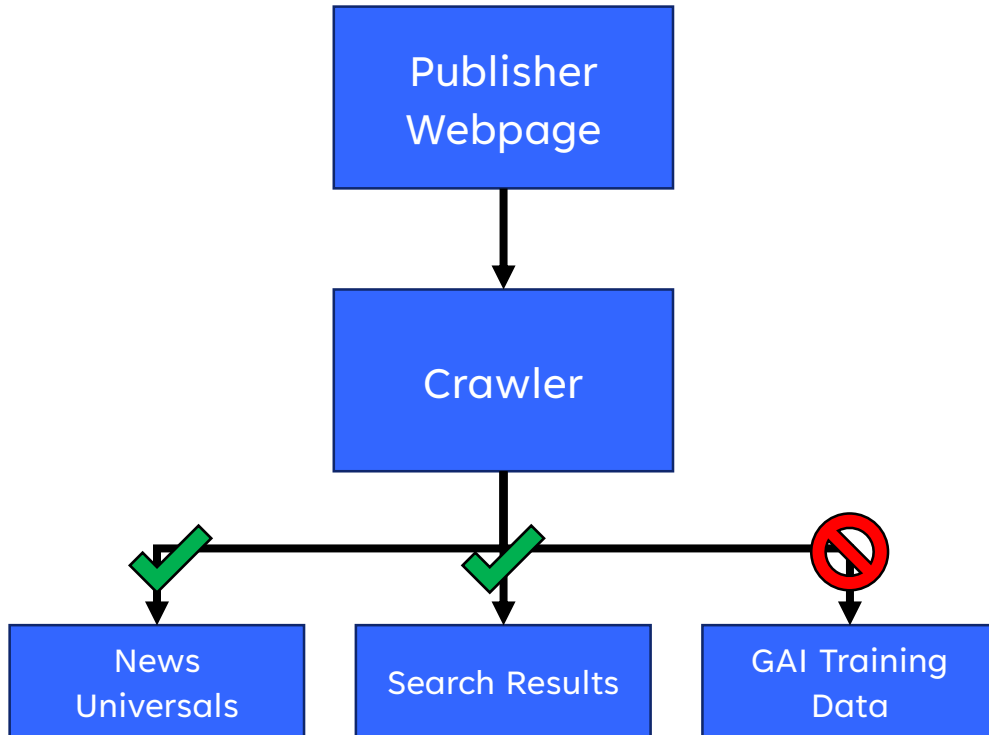
```
#robots.txt
```

```
User-agent: Googlebot  
Allow: /
```

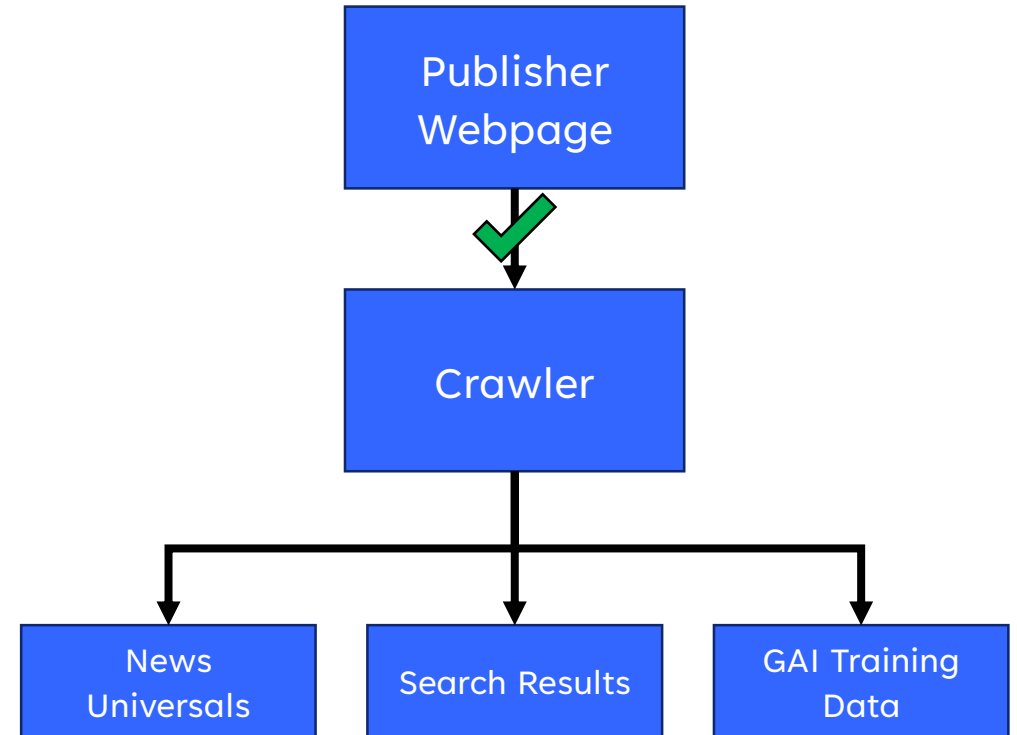
```
User-agent: *  
Disallow: /
```

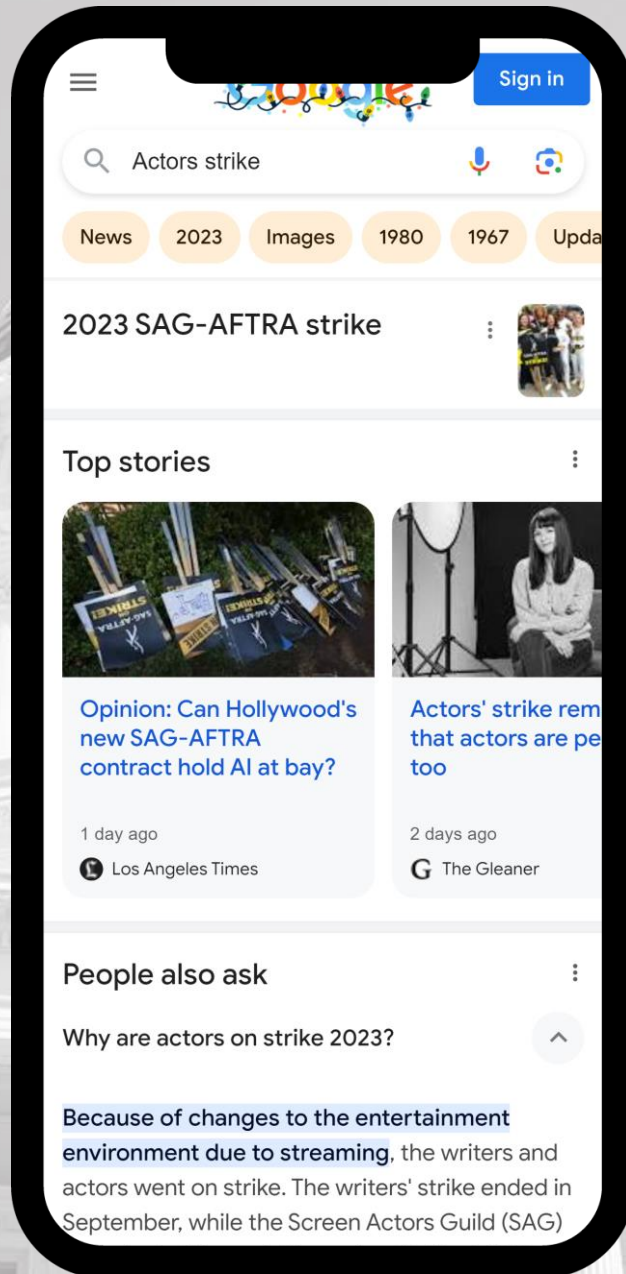
# Publishers Lack Control Over Use of Crawled/Indexed Content

Competitive Market



Crawler with Monopoly Power





Google's **Top Stories** box provides a carousel of links to news articles, which often include high-res images and snippets scraped from the publisher

Clicking the links does take the user to publisher's website, but when that publisher monetizes through advertising using Google's dominant ad tech platform, Google extracts up to **30-70%** of the ad revenue for itself

## People also ask

Why are actors on strike 2023?

**Because of changes to the entertainment environment due to streaming**, the writers and actors went on strike. The writers' strike ended in September, while the Screen Actors Guild (SAG) strike lasted through November 9. Nov 9, 2023

<https://www.vulture.com> › article

[The 2023 Hollywood Strike for Dummies - Vulture](#)

[MORE RESULTS](#)

Are the actors still on strike?

Hollywood's strikes are both now over as actors reach deal with studios and return to work with writers. LOS ANGELES (AP) — On Thursday, for the first time in more than six months, **neither Hollywood's actors nor its writers will be on strike**. Nov 9, 2023

[AP](https://apnews.com) <https://apnews.com> › article › actors...

[Actors union reaches deal with studios to end strike | AP News](#)

[MORE RESULTS](#)

How many days was the actors strike?

Google has introduced a **People also ask** section, which displays highlighted excerpts from news articles and other webpages to answer specific questions, obviating the need for users to leave Google's search results page

2023?

## About

From July 14 to November 9, 2023, the American actors' union SAG-AFTRA was on strike over a labor dispute with the Alliance of Motion Picture and Television Producers. [Wikipedia](#)

Feedback

## Quick facts

Period: July 14, 2023 – November 9, 2023

Duration: 118 days – [wrapbook.com](#)



AP News  
apnews.com



### Actors union reaches deal with studios to end strike

Nov 9, 2023 — More than 60,000 SAG-AFTRA members went on strike July 14, joining screenwriters who had walked off the job more than two months earlier. It was ...



Wikipedia  
[https://en.m.wikipedia.org/wiki/2023\\_SAG-AFTRA\\_strike](https://en.m.wikipedia.org/wiki/2023_SAG-AFTRA_strike)



### 2023 SAG-AFTRA strike

From July 14 to November 9, 2023, the American actors' union SAG-AFTRA was on strike over a labor dispute with the Alliance of Motion Picture and Television ...

Background

Unionization in Hollywood

Pre-strike a

Google's **About** and **Quick facts** sections are similarly designed to keep users on the search results page

Wednesday

on a new contract



The New York Times  
www.nytimes.com



### SAG-AFTRA and Hollywood Studios Agree to Deal to End Actors' Strike

Nov 8, 2023 — SAG-AFTRA, the union representing tens of thousands of actors, reached a tentative deal for a new contract with entertainment companies on ...

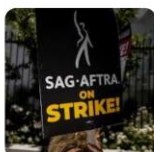


Los Angeles Times  
www.latimes.com



### SAG strike over: What happened with the actors and studios

Nov 10, 2023 — SAG-AFTRA has approved a deal from the studios to end its historic strike. The actors were on strike for more than 100 days.



CNN.com  
www.cnn.com



### Actors union reaches tentative deal with Hollywood and TV studios

Nov 9, 2023 — Hollywood actors have reached a tentative agreement with the major film and television studios to end their historic strike, the actors ...



Even Google's **Organic** search results again include lengthy snippets and images, discouraging users from clicking through to publishers' sites