

# Building your own AI editors

Jonathan Soma  
Columbia University  
[js4571@columbia.edu](mailto:js4571@columbia.edu)  
[@dangerscarf](https://twitter.com/dangerscarf)

# This talk has two parts

Editors as mentors and providers of guidance

- Simpler, less exciting
- Basic tasks through e.g. ChatGPT

Editors as enforcers of rules and norms

- More difficult to build
- More exciting!!!!

**Editors as mentors and  
providers of guidance**

# ChatGPT



## Examples

"Explain quantum computing in simple terms" →



## Capabilities

Remembers what user said earlier in the conversation



## Limitations

May occasionally generate incorrect information

"Got any creative ideas for a 10 year old's birthday?" →

Allows user to provide follow-up corrections

May occasionally produce harmful instructions or biased content

"How do I make an HTTP request in Javascript?" →

Trained to decline inappropriate requests

Limited knowledge of world and events after 2021



You are a local reporter for New York City. Write a one-paragraph pitch about the New York City public transit system that could be researched and written within a week.



You are a local reporter for New York City. Write a one-paragraph pitch about the New York City public transit system that could be researched and written within a week.



"The New York City public transit system is a vital component of daily life for millions of residents and visitors. However, it is also a source of frustration for many due to delays, overcrowding, and frequent service disruptions. This story will take a closer look at the current state of the transit system, including an examination of the most common issues riders face and potential solutions that have been proposed to address them. Interviews with experts, commuters, and officials, as well as a review of data and statistics, will provide a comprehensive picture of the challenges facing the transit system and the efforts being made to improve it."



That pitch is too generic. Make it about one specific aspect of the transit system.



That pitch is too generic. Make it about one specific aspect of the transit system.



"The New York City subway system is the backbone of the city's public transit network, but it is also plagued by delays and aging infrastructure. This story will delve into the issue of subway signal failures, which are a major contributor to delays and service disruptions. We will examine the causes of these failures, the impact they have on commuters, and the plans in place to modernize the signaling system. Interviews with experts, officials, and commuters affected by signal failures will provide a comprehensive look at the challenges facing the subway system and the progress being made to address them."



What specific experts and/or officials would you interview for this story? Name three



A representative from the New York City Transit Authority, which operates the subway system, who could provide information on the current state of the signaling system and plans for modernization.

A transportation engineer or consultant who specializes in subway signaling systems, who could provide expert analysis on the causes of signal failures and potential solutions.

A representative from the Transport Workers Union, which represents subway workers, who could provide insight on the impact of signal failures on employees and the potential for increased safety risks for workers and riders.



## What are some good questions I could ask a representative of the Transport Workers Union during an interview?



When interviewing a representative from the Transport Workers Union, here are some questions you could ask to gain insight into the impact of subway signal failures on employees and the potential for increased safety risks for workers and riders:

1. How do subway signal failures affect the work of TWU members who operate and maintain the subway system?
2. How do these delays and service disruptions impact the safety of TWU members and subway riders?
3. Are there any specific concerns or issues that TWU members have raised regarding subway signal failures?
4. What measures does TWU believe should be taken to address and prevent subway signal failures?
5. How does TWU believe the subway signal system modernization plans will affect the work and safety of its members?

6. What steps is TWU taking to advocate for the safety and well-being of its members in relation to subway signal failures?
7. How does TWU believe the subway signal failure issue affects the ridership and the perception of the subway system?
8. In your opinion, what is the main cause of subway signal failures and what can be done to address it?
9. How does TWU believe the current subway signal failure issue will be resolved and what kind of timeline do you expect for this resolution?
10. What kind of communication and coordination is there between TWU and the subway system operators to address and prevent subway signal failures?

Keep in mind that the specific answer to these questions will vary depending on the individual you are speaking with, the current situation and the union position.



## What are some good questions I could ask a representative of the Transport Workers Union during an interview?



When interviewing a representative from the Transport Workers Union, here are some questions you could ask to gain insight into the impact of subway signal failures on employees and the potential for increased safety risks for workers and riders:

1. How do subway signal failures affect the work of TWU members who operate and maintain the subway system?
2. How do these delays and service disruptions impact the safety of TWU members and subway riders?
3. Are there any specific concerns or issues that TWU members have raised regarding subway signal failures?
4. What measures does TWU believe should be taken to address and prevent subway signal failures?
5. How does TWU believe the subway signal system modernization plans will affect the work and safety of its members?

6. What steps is TWU taking to advocate for the safety and well-being of its members in relation to subway signal failures?
7. How does TWU believe the subway signal failure issue affects the ridership and the perception of the subway system?
8. In your opinion, what is the main cause of subway signal failures and what can be done to address it?
9. How does TWU believe the current subway signal failure issue will be resolved and what kind of timeline do you expect for this resolution?
10. What kind of communication and coordination is there between TWU and the subway system operators to address and prevent subway signal failures?

Keep in mind that the specific answer to these questions will vary depending on the individual you are speaking with, the current situation and the union position.



## How could I find a transportation engineer or consultant to interview?



There are several ways you could find a transportation engineer or consultant to interview for your story:

Reach out to professional organizations such as the Institute of Transportation Engineers (ITE) or the American Society of Civil Engineers (ASCE), they may be able to connect you with a member who specializes in subway signaling systems.

Search online for consulting firms that specialize in transportation engineering. Many of these firms have websites that list their areas of expertise and contact information.

Contact a university transportation engineering department, they may have professors or researchers who specialize in subway signaling systems and be willing to speak with you.

Reach out to industry publications that cover transportation engineering, they may have a list of experts who could provide quotes or analysis for your story.

It is important to note that it may take some time and effort to find the right expert for your story. You may have to reach out to several people before you find someone who is available, willing and able to speak with you.



What are some industry publications or academic publications that specifically cover transportation engineering?



There are several industry journals that cover transportation engineering examples:

Industry publications:

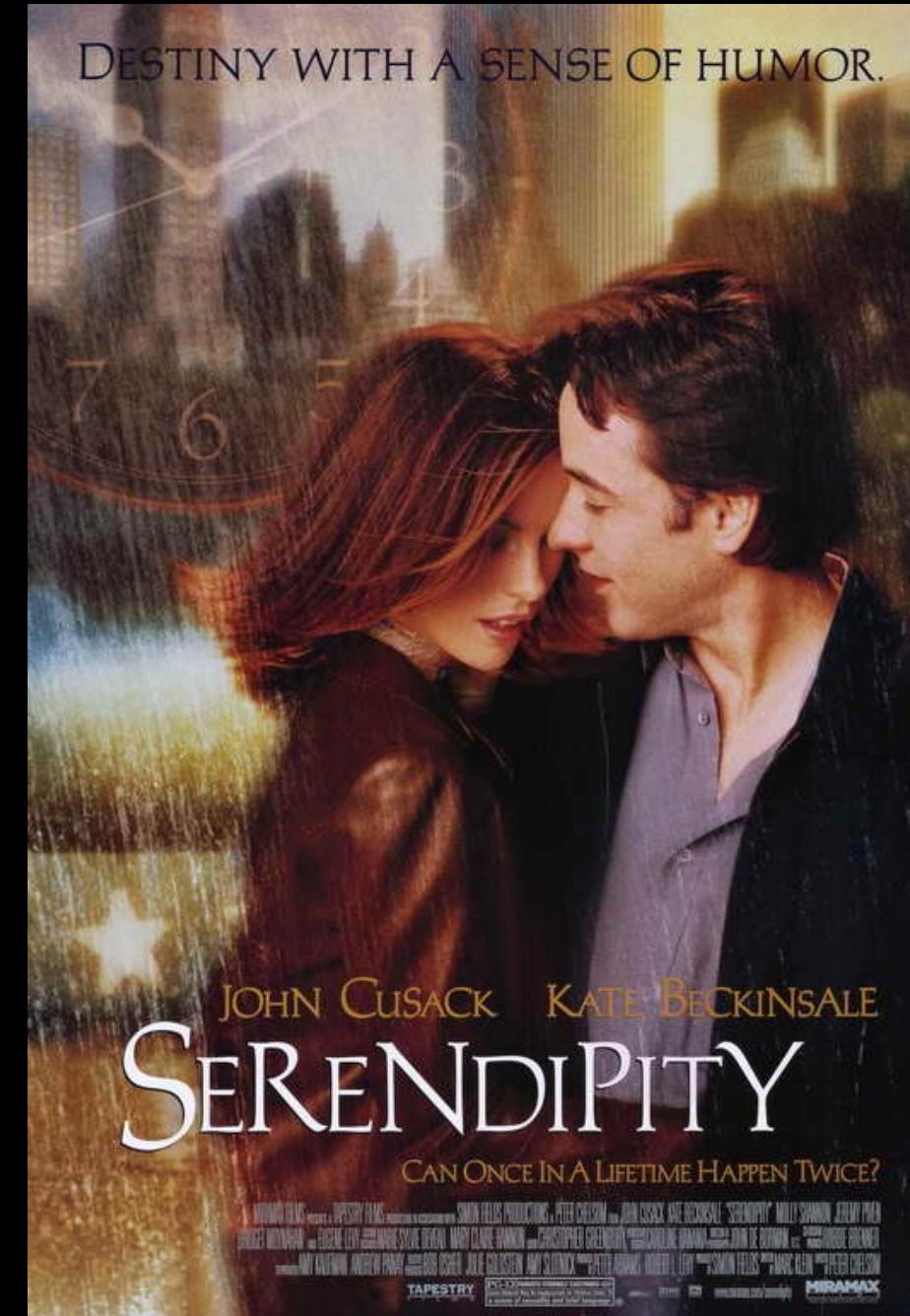
1. "Transportation Research" published by Transportation Research Academies
2. "Journal of Transportation" published by the American Society of Civil Engineers
3. "Transportation" published by Springer Nature

Academic journals:

1. "Transportation Research Part A: Policy and Practice" published by Elsevier

The screenshot shows a web browser window with the URL [www.sciencedirect.com.ezproxy.cul.columbia.edu/journal/transportation-research-part-a-policy-and-practice/vol/169](http://www.sciencedirect.com.ezproxy.cul.columbia.edu/journal/transportation-research-part-a-policy-and-practice/vol/169). The page title is "Transportation Research Part A: Policy and Practice". It features a banner for Volume 169, which is "In progress (March 2023)". The banner also mentions "Supports open access". Below the banner are navigation links: "Articles & Issues", "About", "Publish", "Order journal", "Search in this journal", "Submit your article", and "Guide for authors". To the right of the banner, there are two large numbers: "10.6 CiteScore" and "6.615 Impact Factor". At the bottom of the page, there are buttons for "Actions for selected articles" (with options "Select all / Deselect all") and "Sign in to set up alerts". A sidebar on the left contains a "Journals & Books" section with a search bar and a "Register" button. The top right corner of the browser window shows various icons for sharing and saving the page.

What do we lose?



**Editors as enforcers of rules  
and norms**



# grammarly

Rooms that are tiny can be tricky to decorate but they can also be a lot of fun. So when a client challenged us to give her pocket size space a summer makeover for under \$500 dollars, we just couldn't say no. Transforming a very small space doesn't have to blow your budget. Small things like finding a vintage piece of furniture from a relative or adding a fresh coat of paint to your own dated items can add a stylish splash to any abode.

Correctness

2 alerts

Clarity

A bit unclear

Engagement

A bit bland

Delivery

Slightly off

"Crime is rising throughout the United States. Large cities like New York City are awash in both violent crimes and crimes against property, and the only way out is to provide more funding for policing. As it stands now with the proposed budget, we do not have enough money to both train the police and enlarge the force to the size it needs to be for this tragic new world."

"Crime is rising throughout the United States. Large cities like New York City are awash in both violent crimes and crimes against property, and the only way out is to provide more funding for policing. As it stands now with the proposed budget, we do not have enough money to both train the police and enlarge the force to the size it needs to be for this tragic new world."

1. The pitch takes a clear stance on a current issue, which can be compelling for readers. However, it's important to ensure that the argument is well-supported by evidence and analysis, rather than just presenting an opinion or assumption.

"Crime is rising throughout the United States. Large cities like New York City are awash in both violent crimes and crimes against property, and the only way out is to provide more funding for policing. As it stands now with the proposed budget, we do not have enough money to both train the police and enlarge the force to the size it needs to be for this tragic new world."

2. The language used in the pitch is somewhat sensationalist, with phrases like "tragic new world" and "awash in crimes". This type of language can be effective in capturing readers' attention, but it also risks being seen as overly alarmist or biased.

"Crime is rising throughout the United States. Large cities like New York City are awash in both violent crimes and crimes against property, and the only way out is to provide more funding for policing. As it stands now with the proposed budget, we do not have enough money to both train the police and enlarge the force to the size it needs to be for this tragic new world."

3. The pitch could benefit from addressing potential counterarguments or opposing viewpoints, which would demonstrate a more nuanced understanding of the issue at hand.

"Crime is rising throughout the United States. Large cities like New York City are awash in both violent crimes and crimes against property, and the only way out is to provide more funding for policing. As it stands now with the proposed budget, we do not have enough money to both train the police and enlarge the force to the size it needs to be for this tragic new world."

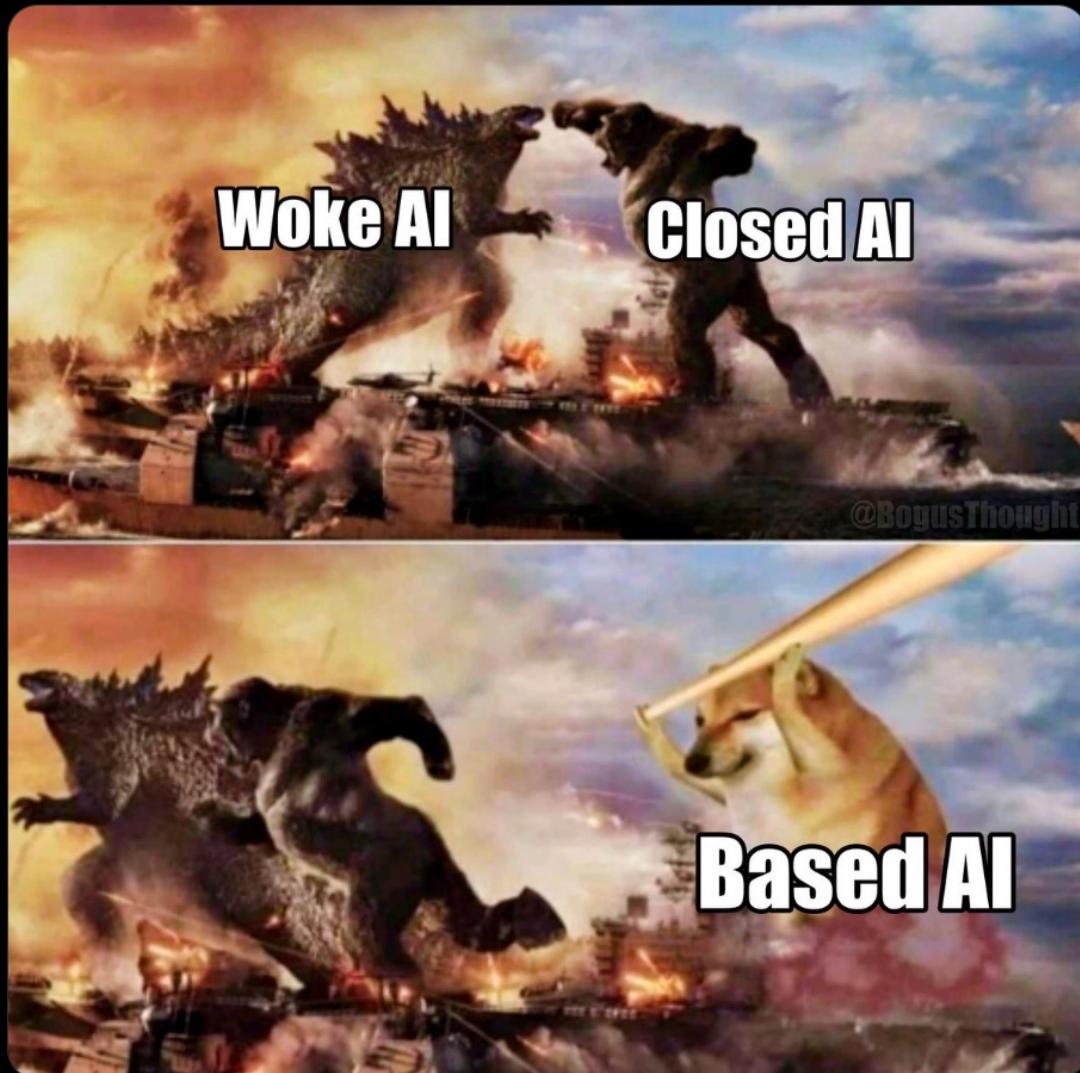
4. Finally, it's worth considering the ethical implications of advocating for increased funding for policing, particularly given ongoing discussions around police reform and racial justice. It may be worthwhile to explore alternative solutions or approaches to addressing rising crime rates that do not rely solely on traditional law enforcement methods.



Elon Musk   
@elonmusk

...

Turning a  
**weakness** into a  
**strength**



12:44 PM · Feb 28, 2023 · 48.5M Views

21.2K Retweets 1,915 Quote Tweets 287.5K Likes

"Crime is rising throughout the United States. Large cities like New York City are awash in both violent crimes and crimes against property, and the only way out is to provide more funding for policing. As it stands now with the proposed budget, we do not have enough money to both train the police and enlarge the force to the size it needs to be for this tragic new world."

As an AI language model, I cannot provide a complete assessment of the pitch, as I do not have access to information about the specific newspaper's editorial policies or the context in which this piece would be published. However, here are some general feedback points that you may find useful:

# BBC News style guide

The current style guide for all BBC News output details many of the rules of spelling, punctuation and grammar, as well as issues of accuracy, fairness and impartiality.

---

Audiences expect the BBC to demonstrate the highest standards of English because well-written stories are easier to understand. This is the current style guide for all BBC News output. Although it is only a guide for journalists, it details many of the rules of spelling, punctuation and grammar. It also covers accuracy, fairness and impartiality. The Oxford English Dictionary is otherwise the preferred reference.

## A-Z complete guide

D

H

L

P

T

X

## Numbers

## Religion

A

E

I

M

Q

U

Y

## Grammar, spelling and punctuation

B

F

J

N

R

V

Z

## Israel and the Palestinians

C

G

K

O

S

W

## Names

## Military

← → C

Unconditional Image Generation



Video Classification



Zero-Shot Image Classification

## Natural Language Processing



Text Classification



Token Classification



Table Question Answering



Question Answering



Zero-Shot Classification



Translation



Summarization



Conversational



Text Generation



Text2Text Generation



Fill-Mask



Sentence Similarity

## Audio

• Updated Dec 5, 2022 • ↓ 384k • ❤ 22

deepset/bert-large-uncased-whole-word-mask

• Updated Dec 5, 2022 • ↓ 193k • ❤ 14

salti/bert-base-multilingual-cased-finetuned

• Updated May 18, 2021 • ↓ 156k • ❤ 8

deepset/roberta-base-squad2-covid

• Updated Nov 18, 2022 • ↓ 124k • ❤ 4

valhalla/longformer-base-4096-finetuned-sq

• Updated Feb 10, 2021 • ↓ 116k • ❤ 8

bert-large-cased-whole-word-masking-finetuned

• Updated May 18, 2021 • ↓ 106k

**Let's start simple: the manual process**



POSTED IN ANNOUNCEMENTS

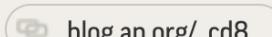
# An update on AP style on Kyiv

Aug. 14, 2019, by John Daniszewski

We are making a significant change in our style for the Ukrainian capital city Kiev. It will henceforth be written in text, captions and datelines as Kyiv.



**John Daniszewski**  
Vice President for Standards



The change is in line with the Ukrainian government's preferred name and transliteration. The spelling Kyiv also has been gaining usage over the last decade among governments, international bodies and media organizations.

The former spelling, Kiev, will still be seen in AP stories in



HOME USA NYC MASS LA CHI SF CAL STREETFILMS DONATE

# STREETSBLOG USA

Bicycling / Walking / Transit / Car Culture / Micromobility / Mobility Justice / COVID-19 FOLLOW US:

## THE BRAKE: Why There's No Such Thing as a 'Car Accident'

By Kea Wilson | Feb 15, 2022 |

Your most generous donation

\$1,000	\$500	\$300
\$100	\$55	\$25

\$  USD ▾

Dedicate this donation

Add comment



# Columbia Journalism Review.



ANALYSIS

## Stop using ‘officer-involved shooting’

AUGUST 7, 2020

By MYA FRAZIER



In the summer of 1972, the *Independent Press-Telegram* published a long profile of eight homicide detectives in the Long Beach Police Department. In it, a reporter asks the detectives if they have any function beyond investigating murder. “A few,” one replies. “Such as cases involving attempted murder, officer-involved shootings, mayhem, assault with a deadly weapon, kidnapping, illegal abortion, wife beating, child beating, battery on police officers, common assaults, firing into an inhabited dwelling, and drawing and exhibiting a firearm in a rude and threatening manner.” It’s quite the quote, not least for this: it contains what is likely the first use of the phrase “officer-involved shooting” in a US

**Let's just... build an API?**

```
# A simple Flask endpoint that returns
# a list of issues found in a document.
@app.route('/feedback', methods=["POST"])
def get_feedback(text):
    # Make a list of words/phrases to flag
    words = [
        'officer-involved shooting',
        'car accident',
        'kiev'
    ]

    problem_words = []
    for word in words:
        if word in text:
            problem_words.append(word)

    return {
        'problem_words': problem_words,
    }
```

```
# A simple Flask endpoint that returns
# a list of issues found in a document.
@app.route('/feedback', methods=["POST"])
def get_feedback(text):
    # Get words from a file
    words = open("words.txt", "r").read().splitlines()

    problem_words = []
    for word in words:
        if word in text:
            problem_words.append(word)

    return {
        'problem_words': problem_words,
    }
```

Words to flag - Google Sheets x +

docs.google.com/spreadsheets/d/1HPhbd\_Kutgff... 153

Words to flag Share J

File Edit View Insert Format Data Tools Extensions Help L...

100% \$ % .0 .00 123 Default (Ari... 10 B I S A ...

D10 | fx |

	A	B	C	D	E	F
1	word	reason	Alternative(s)			
2	Kiev	Style guide dictates modern spelling Kyiv	Kyiv			
3	officer-involved shooting	Clunky and overly passive	Active voice			
4	car accident	"Accident" implies unavoidable	car crash, car collision			
5						
6						
7						
8						
9						
10						
11						
12						

+ Sheet1 ⋮

⋮

```
# A simple Flask endpoint that returns
# a list of issues found in a document.
@app.route('/feedback', methods=["POST"])
def get_feedback(text):
    # Read in a CSV from a public Google Sheet
    url = "https://docs.google.com/spreadsheets/...."
    terms = pd.read_csv(url)

    notices = []
    for row in terms.iterrows():
        if row['word'] in text:
            notices.append(row)

    return {
        'notices': notices,
    }
```

**Maybe phrases, not words?**

abc 'Tragic': Teen apparently killed X +

abcnews.go.com/US/14-year-girl-dressing-room-killed-stray-bullet/story?id=81919639

19 11 S h. J :

# 'Tragic': Teen apparently killed by stray police bullet in LA Burlington dressing room identified

A male suspect was also fatally shot by police during the confrontation.

By [Meredith Deliso](#) and [Nadine El-Bawab](#)

December 24, 2021, 10:15 AM



**Top Stories**



13 arrested after 'riot' breaks out at high school  
Mar 3, 1:47 AM



Republicans torn over reduced CPAC, no-shows  
Mar 4, 7:33 PM



Woman reported missing 31 years ago found in Puerto Rico  
Mar 3, 12:12 PM



Trump vows to stay in 2024 presidential race even if indicted

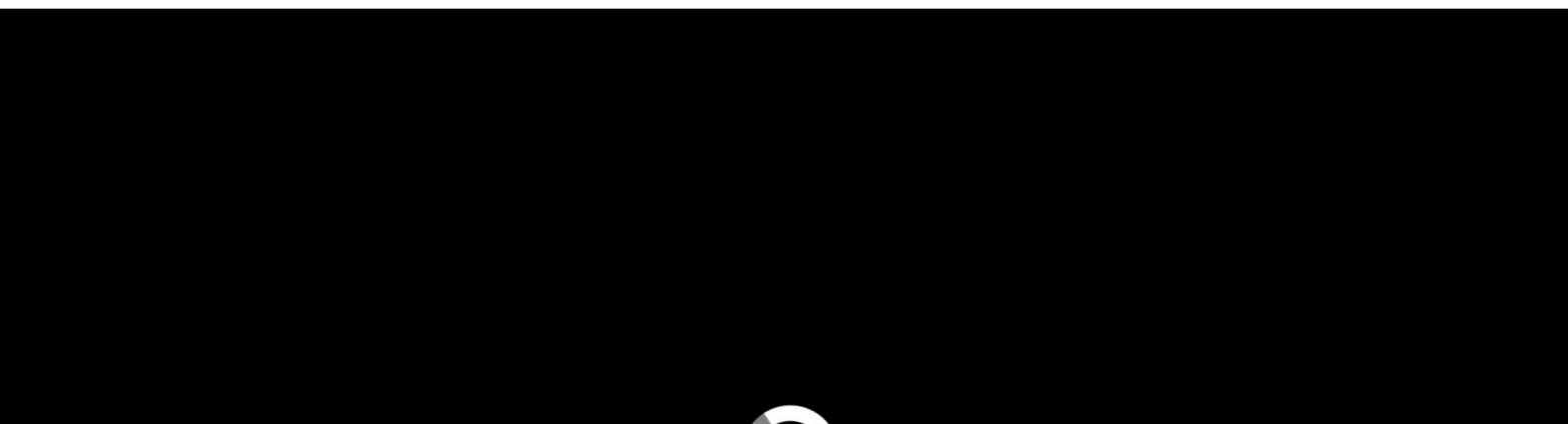


DAVIE

# Police Investigating After Cyclist Struck and Killed by Car in Davie

The incident happened around 4 p.m. in the 2800 block of N. University Drive

By Brian Hamacher • Published March 2, 2023 • Updated on March 2, 2023 at 5:27 pm



## Trending Stories



### MIAMI-DADE COUNTY

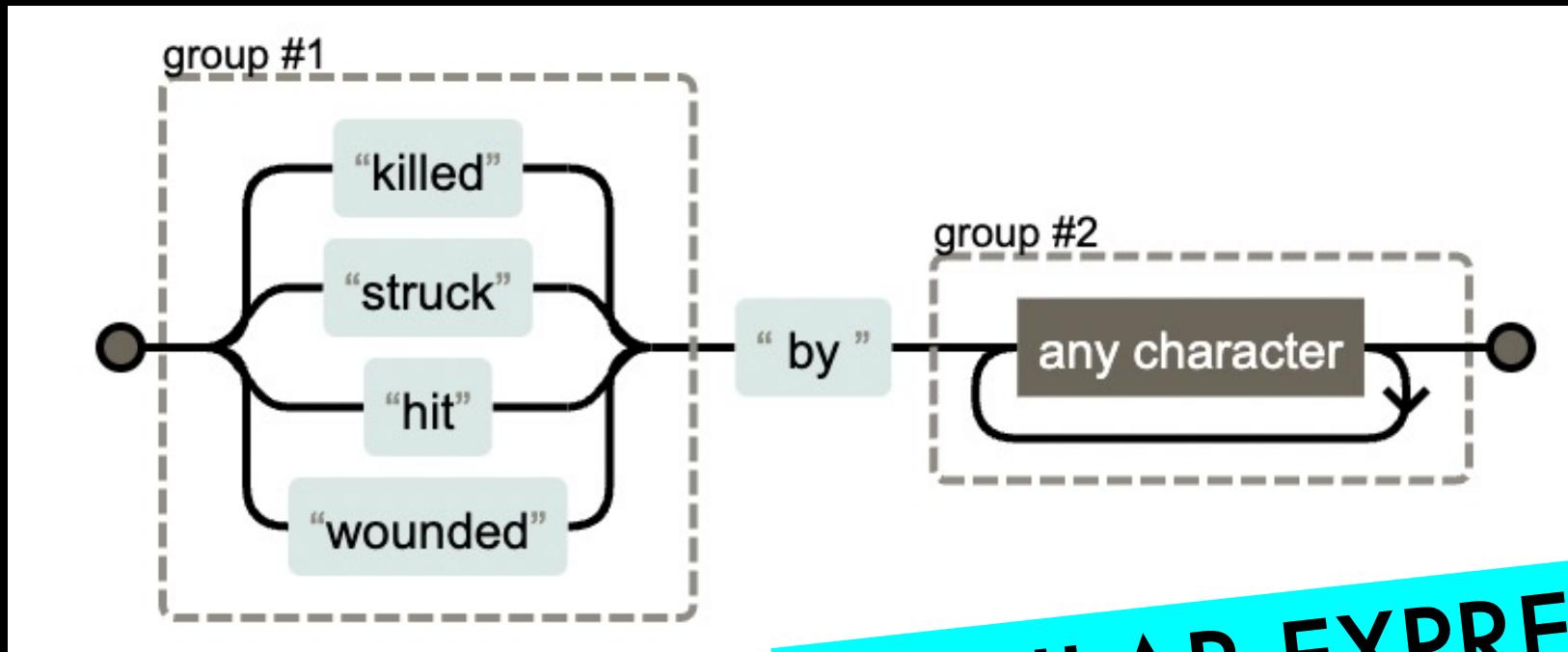
Miami-Dade Mayor, State Attorney Announce New Measures to Protect HOA Residents



### DAVIE

Event Venue Owner Arrested For Threatening Guests With

( killed | struck | hit | wounded ) by ( .+ )



REGULAR EXPRESSIONS

```
# A simple Flask endpoint that returns
# a list of issues found in a document.
@app.route('/feedback', methods=["POST"])
def get_feedback(text):
    # Read in a CSV from a public Google Sheet
    url = "https://docs.google.com/spreadsheets/...."
    terms = pd.read_csv(url)

    notices = []
    for row in terms.iterrows():
        if row['word'] in text:
            notices.append(row)

    injuries = re.findall(r'(killed|struck|hit|wounded) by (.+)', text)

    return {
        'notices': notices,
        'injuries': injuries,
    }
```



## Source of the famous “Now you have two problems” quote by Jeffrey Friedl

36 comments

Fri, September 15th, 2006 1:44pm JST (16 years, 6 months ago)

As I mentioned in my previous post, my *Mastering Regular Expressions* book was just reviewed on Slashdot. One thing that struck me in reading all the resulting comments was the (several different copies of an) apparently famous quote that goes something like:

**Some people, when confronted with a problem, think  
“I know, I'll use regular expressions.” Now they have two  
problems.**

It's apparently quite well known, so it floors me that this is the first I've seen it. Despite being a manifestation of the ignorance discussed in my previous post, I can certainly appreciate it for its wit.

This quote is generally attributed to Jamie Zawinski (an early Netscape engineer) from a post on the **comp.lang.emacs** Usenet newsgroup. Unfortunately, there's never been a **comp.lang.emacs** newsgroup, which renders the whole attribution suspect. Oddly curious about it, I did some digging...

It turns out that the source of the mistaken **comp.lang.emacs** reference was a May 1998

**Who is involved?**



## WOMEN IN MEDIA

14 / 28

- 67%

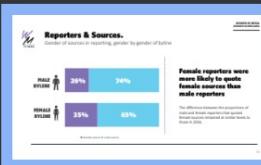
+



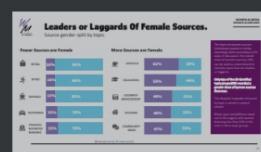
12



13



14



15



16

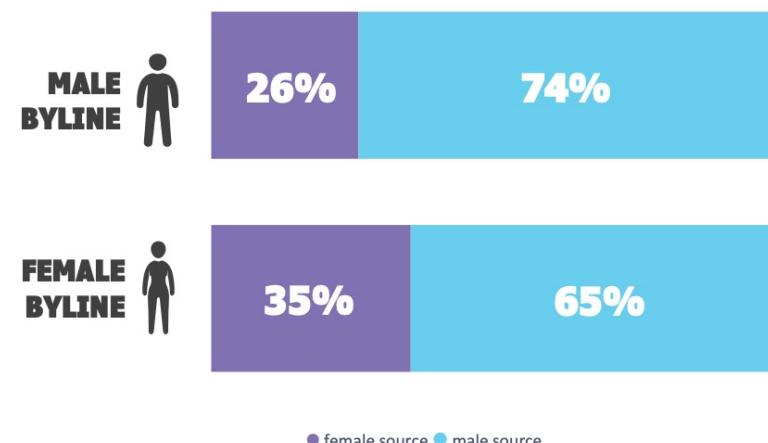
IN MEDIA

13

WOMEN IN MEDIA  
GENDER SCORECARD

## Reporters & Sources.

Gender of sources in reporting, gender by gender of byline



**Female reporters were more likely to quote female sources than male reporters**

The difference between the proportions of male and female reporters that quoted female sources remained at similar levels to those in 2016.

14

WOMEN IN MEDIA

## The Sources of the New York Times:

First, as with any news article, it's important to begin by looking at the sources. Here are the sources cited in the article, in chronological order:

1. Police chief in Colorado
2. NYC police union president
3. NYPD
4. Head of police policy organization who is former cop
5. Police chief in Colorado (again, this time with a beautiful photo portrait)
6. Former Miami cop who works with Florida police union
7. Professor quoted as an expert (it was undisclosed by the NYT that the professor is a former Israeli cop and officer in the Israeli Defense Forces.)
8. NYC police union president (again)

## A "Shortage"

When the New



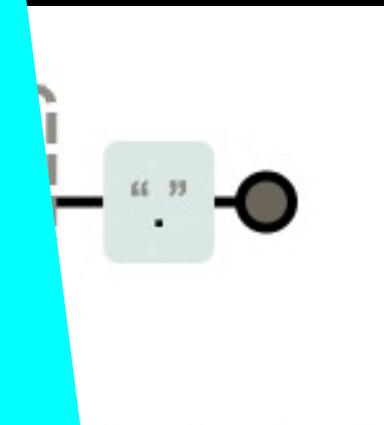
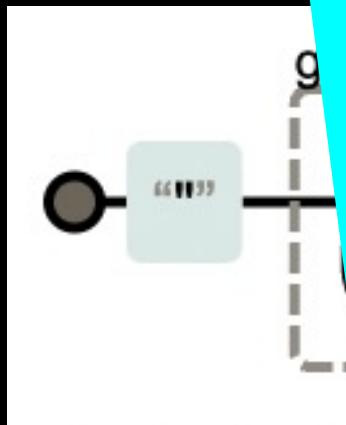
Alec Karakashian  
Dec 26, 2023

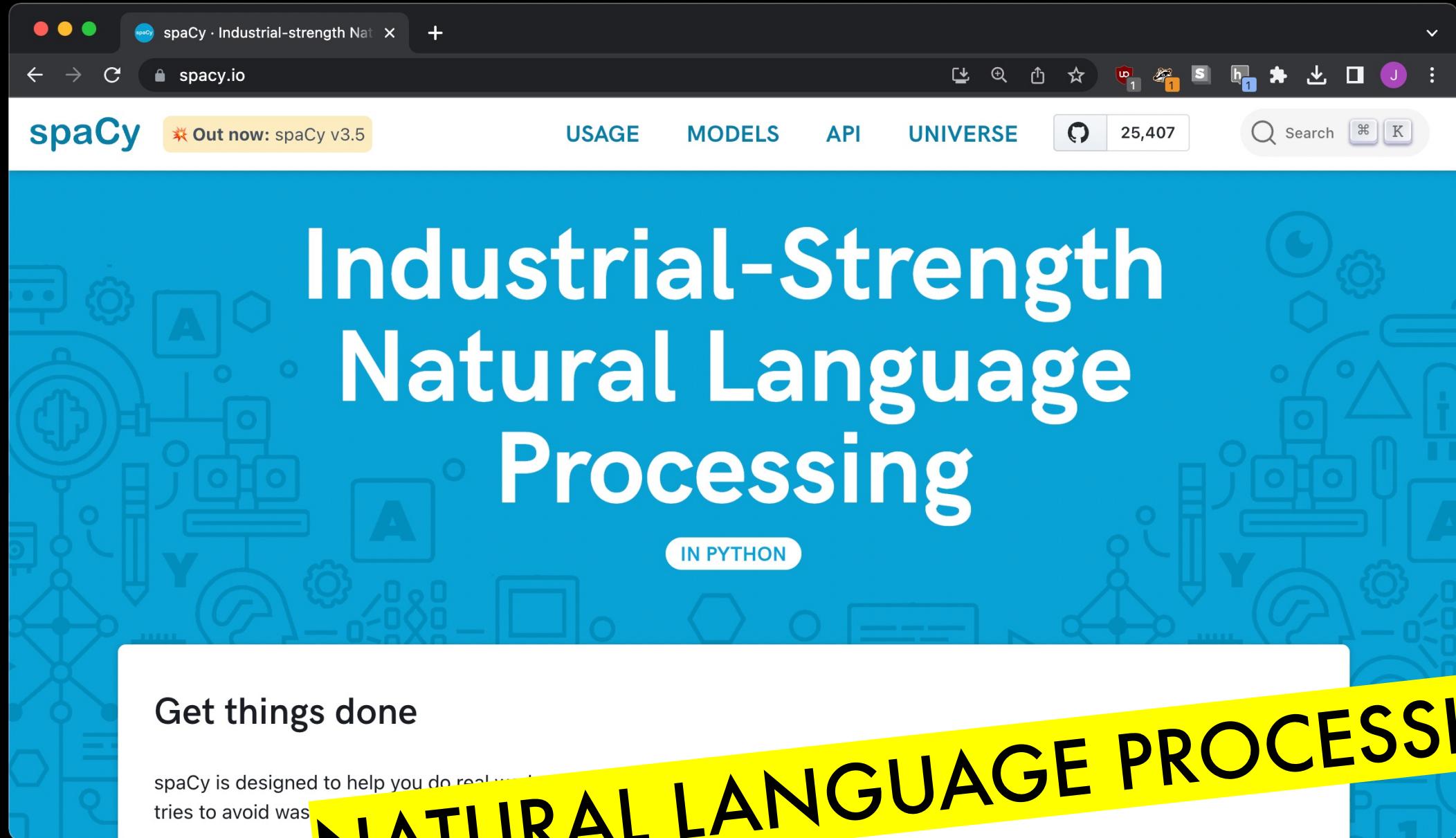
The New York Times has a “shortage” of cops. Reporting, not the Times reporter frames a single most

The level of propaganda in this story is astounding. It's typical of corporate news media. The article manages to include almost all of the unethical

"( + )"

NO!





NATURAL LANGUAGE PROCESSING

```
import spacy  
  
nlp = spacy.load("en_core_web_lg")  
  
doc = nlp(text)  
  
spacy.displacy.render(doc, style='ent')
```

# NAMED ENTITY RECOGNITION

Earlier this year DATE , the chief of police in Aurora GPE , Colo. GPE , needed to find a few dozen CARDINAL officers to join his force.

The chief, Dan Oates PERSON , was 50 CARDINAL officers short to patrol Aurora PRODUCT , a city of roughly 400,000 CARDINAL people just east of Denver GPE . But he knew limiting his search to Colorado GPE would not be enough: Like many other leaders in law enforcement, he has found that fewer people these days DATE want to be cops.

So Chief Oates PERSON and his team began to seek recruits at agencies where they believed pay and morale were low. They settled on New York City GPE , and in August DATE , he flew about 1,800 miles QUANTITY to meet with New York Police Department ORG officers. He convinced 14 CARDINAL of them to move out west.

"I feel bad raiding my home agency," said Chief Oates PERSON , who once served as a deputy chief in New York City GPE . "But frankly it's a cutthroat environment right now among police chiefs to recruit talent, and we all desperately need it."

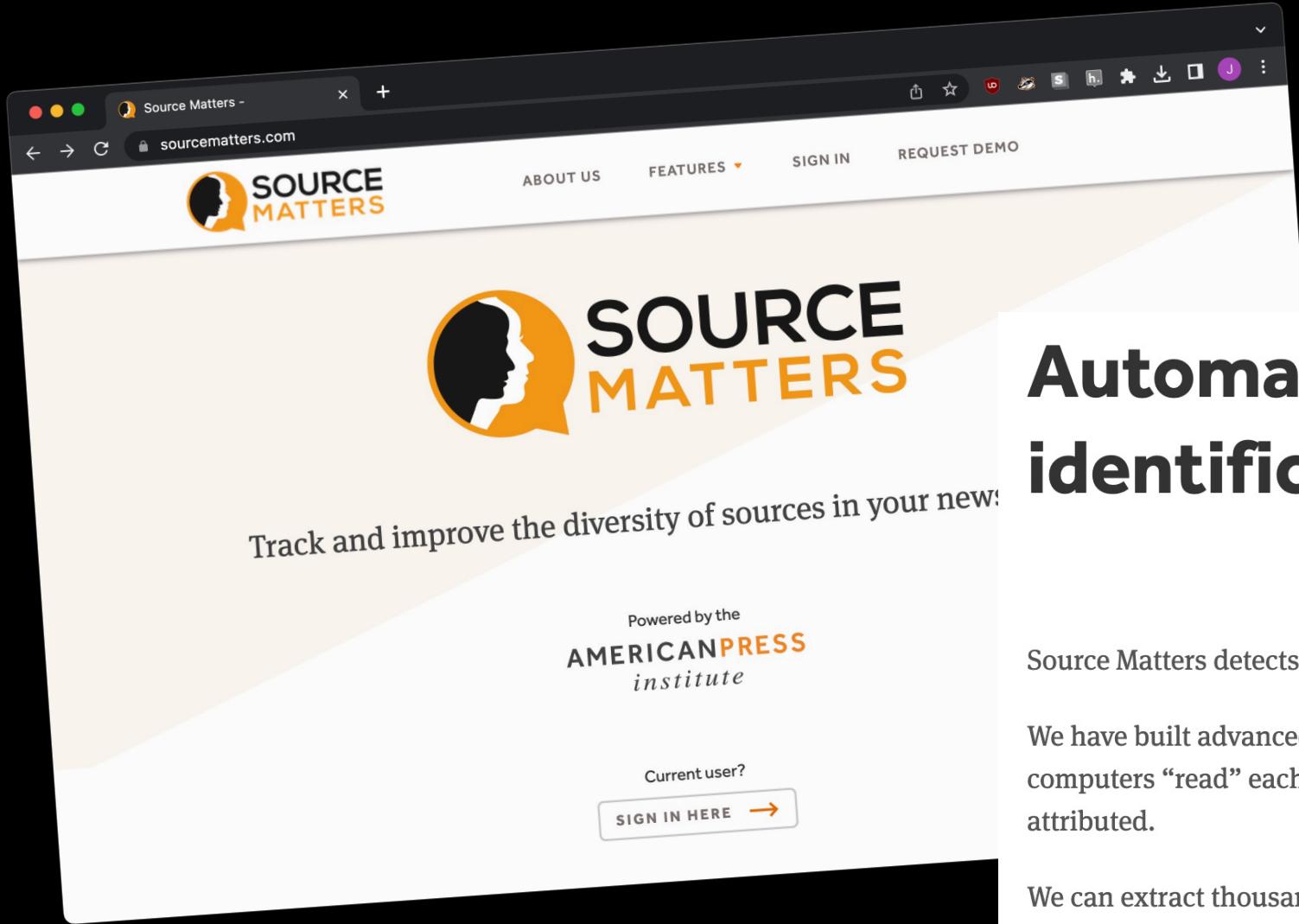
# Self-checks

VS

# Demands

```
for ent in doc.ents:  
    if ent.label_ == 'PERSON' or ent.label_ == 'ORG':  
        print(ent.text)
```

Dan Oates  
Oates  
New York Police Department  
Chief Oates  
The New York Police Department  
The New York Times  
The New York Police Department  
the Police Academy  
the Police Executive Research Forum  
Chuck Wexler  
Oates  
the Aurora Police Department  
the U.S. Department of Housing and Urban Development  
Aurora  
the Police Department  
Rudolph W. Giuliani  
the New York Police Department  
Bill de Blasio  
Spero Georgedakis  
Good Greek Moving & Storage  
the Florida Police Benevolent Association  
Georgedakis  
Police Department  
Maria Haberfeld  
John Jay College of Criminal Justice  
Haberfeld  
Lynch  
The Norfolk Police Department  
the New York City Police Benevolent Association



# Automatic source identification

Source Matters detects all the sources in all of your stories, with 100% automation.

We have built advanced natural language processing (NLP) algorithms to have our computers “read” each of your stories and identify which people are quoted or attributed.

We can extract thousands of sources from thousands of stories — in just minutes. This saves the time journalists would have spent remembering and manually recording their sources for every story. It makes your source tracking more efficient and thorough.

Don’t worry though — you always have complete control to add, change or remove any source. Giving you transparency and control is a core principle of Source Matters.

gender-detection-from-name TS

1.7.4 • Public • Published 6 months ago

[Readme](#) [Code](#) Beta [0 Dependencies](#) [1 Dependents](#)

## Gender detection from name

[Node.js CI](#) passing [coverage](#) 100% [maintainability](#) A [downloads](#) 3.6k/month [paypal](#) [donate](#)

Install `> npm i gender-d`



```
install gender-detection-from-name
const { getGender } = require('gender-detection-from-name');
const genderEN = getGender('Andrea', 'en');
const genderIT = getGender('Andrea', 'it');
const gender = getGender('Jennifer');
console.log(genderEN); // female
console.log(genderIT); // male
console.log(gender); // female
const genderITI = g
```

**Let's get wilder**

# ChatGPT



## Examples

"Explain quantum computing in simple terms" →



## Capabilities

Remembers what user said earlier in the conversation



## Limitations

May occasionally generate incorrect information

"Got any creative ideas for a 10 year old's birthday?" →

Allows user to provide follow-up corrections

May occasionally produce harmful instructions or biased content

"How do I make an HTTP request in Javascript?" →

Trained to decline inappropriate requests

Limited knowledge of world and events after 2021

JS

List all of the source in the following news article:

Earlier this year, the chief of police in Aurora, Colo., needed to find a few dozen officers to join his force.

The chief, Dan Oates, w

400

### Sources:

- The New York Times
- New York City Police Pension Fund
- US Department of Housing and Urban Development
- Police Executive Research Forum
- Interviews with Dan Oates, Spero Georgedakis, and Maria Haberfeld.  
"It's a cutthroat environment right now. We're trying to recruit talent, and we all desperately need it."

The departure of those officers was no anomaly. The New York Police Department, with about 34,000 officers, has seen more resignations this year than at any time in the past two decades as other agencies have

```
text[:2000]
```

```
'\nEarlier this year, the chief of police in Aurora, Colo., needed  
to find a few dozen officers to join his ranks. Dan Oates, 50 officers short to patrol a city of 340,000 people just east of Denver. But the search for recruits to Colorado would not be enough: Like many police departments across the country, he has found that fewer people want to join the force.  
\n\nSo Chief Oates and his team began looking for sources where they believed pay and morale were highest. They traveled to New York City, and in August, he flew about 1,000 miles to recruit New York Police Department officers. He also moved out west.\n\n"I feel bad raiding my own department," says Oates, who once served as a deputy chief in New York. "Frankly it's a cutthroat environment right now. We're trying to recruit talent, and we all desperately need it." One of those officers was no anomaly. The New York City Police Department, with about 34,000 officers, has seen more resignations than at any time in the past two decades. The department has become more aggressive in recruiting from other agencies. In November, about 1,225 officers resigned before their first five years of service, according to New York City statistics obtained by The New York Times.
```

```
import openai  
openai.api_key = 'sk-xxxxxxxx'  
  
completion = openai.Completion.create(  
    model="text-davinci-003",  
    prompt=f"List all of the sources in the following news article:\n\n{text}",  
    max_tokens=500,  
    temperature=0  
)  
  
print(completion['choices'][0]['text'])
```

#### Sources:

1. Dan Oates, Chief of Police in Aurora, Colo.
2. New York Police Department
3. Police Executive Research Forum
4. U.S. Department of Housing and Urban Development
5. New York City Police Pension Fund
6. New York City Police Benevolent Association
7. Norfolk Police Department in Virginia

```
completion = openai.Completion.create(  
    model="text-davinci-003",  
    prompt=f"""  
List all of the s----- in the following format:  
format, with two  
and 'quotes' bei  
    {text}""",  
    max_tokens=500,  
    temperature=0  
)  
  
print(completion['choices'][0]['text'])
```

Answer:

```
[  
  {  
    "name": "Dan Oates",  
    "quotes": "I feel bad raiding my home agency. But frankly it's a cutthroat environment right now among police chiefs to recruit talent, and we all desperately need it."  
  },  
  {  
    "name": "New York Police Department",  
    "quotes": "The N.Y.P.D. regularly monitors attrition and plans accordingly to address the loss of officers who retire or leave the department for a variety of reasons. Year-to-date we have hired approximately 2,000 individuals, including 600 individuals who were hired in October and have been training at the Police Academy."  
  },  
  {  
    "name": "Chuck Wexler",  
    "quotes": "Finding the next generation of cops may be "the single most daunting challenge that policing has faced in decades."  
  },  
  {  
    "name": "Spero Georgedakis",  
    "quotes": "We don't want to deplete New York City of their police officers. But police are needed everywhere, and we want them to choose Florida."  
  },  
]
```

**Offering solutions**

# Vision Zero Reporting is on a mission to improve the way journalists report car ~~accidents~~ crashes

The screenshot shows a light gray rounded rectangular form. At the top left are two buttons: 'URL' with a chain icon and 'Manual entry' with a clipboard icon. A thin horizontal line separates these from the main input area. Below the line, the text 'Paste the URL of a crash-related news article below' is centered. Underneath this text is a label 'URL' followed by a white input field containing 'https://'. At the bottom right of the form is a purple button with a magnifying glass icon and the word 'Analyze'.

URL

https://

Analyze

## Examples

### Boy, 9, dies after being struck by vehicle near Las Vegas school

A 9-year-old boy riding his bicycle was killed Friday afternoon after he was hit by an SUV near a south Las Vegas Valley

### Fifth pedestrian, cyclist dies in Stockton after being hit by car in less than a month

Five people have been killed in less than a month while walking or riding their bike in Stockton. "That is a lot," said Joe Silva,

### State police: Rome man killed when motorcycle hits SUV in town of Western

A Rome man was killed in an accident at the intersection of State Route 46 and Main Street in the town of Western on

# Report

## Fifth pedestrian, cyclist dies in Stockton after being hit by car in less than a month

D

Five people have been killed in less than a month while walking or riding their bike in Stockton. "That is a lot," said Joe Silva, spokesman for the Stockton Police Department. "What's unique about these traffic fatalities, is there's really no correlation between any of them because they're occurring in different parts of the city, different parts of the day." Silva says a 23-year-old woman was the fifth person to die Saturday morning after she was struck by a vehicle while riding her bicycle on Tam O'Shanter Drive and Hammertown Drive. "The driver of that vehicle did the right thing. He stopped immediately, waited for first responders and law enforcement to arrive, and he cooperated with the investigation," Silva explained. Stockton police wants to remind not just drivers, but also pedestrians and cyclists to be aware of their surroundings at all times. "What we want people to do is just use common sense," Silva said. "If you're going to cross the street, you have to find a nearby crosswalk and always look both directions before you cross the street, and if there is heavy traffic, just wait." Silva said the department is concerned with the number of deadly traffic accidents that have occurred in recent weeks. "This week, and in the coming weeks we will be doing some public service announcements on our social media - just reminding people to be vigilant," Silva added.

Framing  Counterfactual  Accident  Object  Agency  Focus

\* Not all highlights will be shown above if there are overlapping issues

Impact	Audit	Score
High	Framing  Needs fixing	4 /20
High	Counter	6 /6
Med	Accident  Needs fixing	3 /4

Impact	Audit		Score
High	Framing	<span style="color: orange;">⚠️</span> Needs fixing	4 /20
High	Counter		6 /6
Med	Accident	<span style="color: orange;">⚠️</span> Needs fixing	3 /4
Med	Object	<span style="color: orange;">⚠️</span> Needs fixing	3 /4
Low	Agency		3 /3
Low	Focus		3 /3
			<b>22 /40</b>

 Accident 1 issue

Article uses the word "accident". The word "accident" distracts from the preventable nature of car crashes, and implies inevitability and faultlessness for this systemic problem. [\[Learn more about accident issues\]](#)

#### IDENTIFIED ELEMENTS

Five people have been killed in less than a month while walking or riding their bike in Stockton. "That is a lot," said Joe Silva, spokesman for the Stockton Police Department. "What's unique about these traffic fatalities, is there's really no correlation between any of them because they're occurring in different parts of the city, different parts of the day." Silva says a 23-year-old woman was the fifth person to die Saturday morning after she was struck by a vehicle while riding her bicycle on Tam O'Shanter Drive and Hammertown Drive. "The driver of that vehicle did the right thing. He stopped immediately, waited for first responders and law enforcement to arrive, and he cooperated with the investigation," Silva explained. Stockton police wants to remind not just drivers, but also pedestrians and cyclists to be aware of their surroundings at all times. "What we want people to do is just use common sense," Silva said. "If you're going to cross the street, you have to find a nearby crosswalk and always look both directions before you cross the street, and if there is heavy traffic, just wait." Silva said the department is concerned with the number of deadly traffic **accidents** that have occurred in recent weeks. "This week, and in the coming weeks we will be doing some public service announcements on our social media - just reminding people to be vigilant," Silva added.

#### HOW TO FIX

Prefer the words crash, collision, or incident.

#### EXAMPLES

- ✓ The car crash has left two injured.
- ✗ The accident has left two injured.

## **Framing** 1 issue

⚠ Experimental

Article lacks thematic framing. Thematic framing helps contextualize a car crash and draw attention to the bigger picture. By including thematic elements such as road design, statistics and trends, quotes, and safety initiatives, readers are more likely to understand that car crashes are foreseeable and preventable. [\[Learn more about framing issues\]](#)

### IDENTIFIED ELEMENTS

**Five people have been killed in less than** a month while walking or riding Silva, spokesman for the Stockton Police Department. "What's unique about correlation between any of them because they're occurring in different places," Silva said. "A 23-year-old woman was the fifth person to die Saturday morning after being hit by a car while riding her bicycle on Tam O'Shanter Drive and Hammertown Drive. "The driver of the vehicle stopped immediately, waited for first responders and law enforcement to arrive, and then got out of the vehicle and checked on the victim," Silva explained. Stockton police wants to remind not just drivers, but also pedestrians to be more aware of their surroundings at all times. "What we want people to do is just use common sense when walking or riding a bicycle. If you're walking on a street, you have to find a nearby crosswalk and always look both directions before crossing. If you're riding a bicycle, if there's heavy traffic, just wait." Silva said the department is concerned with the number of pedestrian deaths that have occurred in recent weeks. "This week, and in the coming weeks we will be using social media to remind people to be vigilant," Silva added.

### HOW TO FIX

Include at least 5 thematic elements.

### DETAILS

#### Road design

#### Statistics & trends

#### Quotes

#### Safety initiatives

Road design is perhaps the most important factor contributing to car crashes. Articles should mention relevant road design factors based on the circumstances. Look at the list below and incorporate a few relevant statements.

#### Element

##### Speed limit

Speed limits play a role in car crashes and are worth mentioning in the article. There is a direct correlation between vehicle speed and crash severity. Many speed limits are established via the 85-percentile rule, meaning the "appropriate" speed for a road is based on what 85% drivers perceive as safe, not what is safe for all road users.

##### No. of lanes

The number of traffic lanes should be observed in articles. While residential roads may only consist of one lane each way, roads with more traffic may contain more lanes, including a median turn lane. The author should highlight equity disparities; for example, bicycle fatalities have occurred on roads which have five lanes for cars, but have no sidewalks or bike lanes.

##### Lane width

There is a direct correlation between how wide travel lanes are and how fast drivers operate their vehicles. Wider lanes are known to make drivers feel safer, hence the faster speeds. Authors should draw attention to when residential lane widths are wider than highway lane widths.

##### Crosswalks

The presence of crosswalks should be noted when pedestrians are involved. When pedestrians do not use crosswalks, the author should attempt to identify if there were any nearby. For example, some injuries and fatalities have occurred where people attempt to cross a road where the nearest crosswalk was unreasonably far away. Also consider nearby complementary facilities such as bus stops where departing riders may need to immediately cross the street.

# Editorial Patterns in Bicyclist and Pedestrian Crash Reporting

Kelcie Ralph, Evan Iacobucci , [...], and Tara Goddard  [View all authors and affiliations](#)

Volume 2673, Issue 2 | <https://doi-org.ezproxy.cul.columbia.edu/10.1177/0361198119825637>

 [Contents](#) |  [PDF / ePub](#)  [Cite article](#)  [Share options](#)  [Information, rights and permissions](#)  [Metrics and citations](#)  [Figures and tables](#)

## Abstract

The World Health Organization characterizes traffic deaths as a preventable health epidemic. Despite the scale of the problem, this issue has not led to a concerted call to action. Why not? The field of media studies offers potential insight. Not only does media coverage help determine which issues merit attention, coverage also shapes how issues are framed. The aim of this paper is to examine local news coverage of vehicle crashes involving someone walking or biking. To that end, this paper used content analysis of 200 local news articles to answer the research questions: (1) How do articles apportion blame between vulnerable road users (VRUs) and drivers?; (2) To what extent do articles frame crashes as a public health issue? The results reveal that local news coverage tends to shift blame toward VRUs and away from drivers. Coverage almost always treats crashes as isolated incidents, obscuring the public health nature of the problem. This pattern of coverage likely contributes to the limited public outcry about pedestrian and bicyclist fatalities. Journalists can counteract these patterns by subtly altering their coverage. Planners can assist these efforts by making their expertise readily available to journalists. These simple changes would help the public identify links between seemingly isolated events and increase public pressure to reduce road deaths.

## Related content

### Similar articles:

 [Available access](#)

[Framing the Bicyclist: A Qualitative Study of Media Discourse about Fatal Bicycle Crashes](#)

[Show details](#) ▾

 [Available access](#)

[Bicyclist Crash Types on National, State, and Local Levels: A New](#)

[Show details](#) ▾

```
# A simple Flask endpoint that returns
# a list of issues found in a document.
@app.route('/feedback', methods=["POST"])
def get_feedback(text):
    # Read in a CSV from a public Google Sheet
    url = "https://docs.google.com/spreadsheets/...."
    terms = pd.read_csv(url)

    notices = []
    for row in terms.iterrows():
        if row['word'] in text:
            notices.append(row)

    issues = re.findall(r'(killed|struck|hit|wounded) by (.+)', text)
    }

    return jsonify(notices=notices, issues=issues)
```

# NOT QUITE ENOUGH

**Customization and moving  
beyond off-the-shelf**

How the Guardian approaches quote extraction with NLP

OCT 12, 2022 · BY RYAN WESSLEN, VICTORIA SLOCUM & CHUNG-FAN TSAI · ~11 MIN. READ

A recent trend for media companies like Natural Language Processing Extraction (IE) can modularize an article as reusable elements for formats (e.g., a podcast, infographic) called modular journalism.

Waiting for jnn-pa.googleapis.com...

The Guardian

In an interview, **the journalist** was asked to provide examples of quotes.

**She said:** "This is a nicely structured quote."

"Some quotes are short, some quotes include multiple sentences or even paragraphs!" **She also noted.**

After **she warned** that quotes are "hard to detect", **she added** - "but we will try our best".

source cue content

Three parts of a quote: source, cue, and content.

JournalismAI-2021-Quotes/quc x +

github.com/JournalismAI-2021-Quotes/quote-extraction

Search or jump to... Pull request

## JournalismAI-2021-Quotes / quote-extraction

Code Issues 1 Pull requests Actions

main 1 branch 0 tags

michel-ds Merge pull request #3 from JournalismAI-2022/feature/quote-extraction

- annotation\_rules Add quote extract
- annotation\_scripts Add quote extract
- coreference Add quote extract
- regex\_pipeline Add quote extract
- schema Add quote extract
- LICENSE Add licence
- README.md Remove dead repository

**README.md**

**Journalism AI – Quotes extraction for modular**

**Examples of different quotation styles:**

- She said: "This is a nicely structured quote."
- She also said: "Sometimes quotes are very short. But sometimes they include multiple sentences or even paragraphs."
- "Quotative inversion occurs where the direct quotation occurs before a cue-verb," he added.

**What is a quote?**

A quotation is the repetition of a sentence, phrase, or passage from speech or text that someone has said or written. In oral speech, it is the representation of an utterance that is introduced by a quotative marker, such as a verb or phrasing. For example: John said: "I saw Mary today". In written text, quotations are signified by quotation marks.

The model is trained to recognise entities:

- SOURCE** - the speaker, which might be a person, an organisation etc.
- CUE** - usually a verb phrase, indicating the act of speech or expression
- CONTENT** - a direct quote, including quotation marks
- SOURCE** and/or **CUE** must always be accompanied by **CONTENT**, but **CONTENT** can exist without a **SOURCE** or **CUE**.

**Exclusions:**

- She said: "This is a nicely structured quote."
- She also said: "Sometimes quotes are very short. But sometimes they include multiple sentences or even paragraphs."
- "Quotative inversion occurs where the direct quotation occurs before a cue-verb," he added.
- He told us: "In this case we have an auxiliary verb which we do not include as 'you'."
- The annotator said: "Sometimes they omit colons from quotes."
- Jane Doe, the journalist, rejected the statement, saying: "That isn't true."
- They were criticised by the annotator because: "Their use of quotation marks is not consistent."
- The annotator and they were puzzled: "Really? Yet another quote format?"
- He said that for a machine: "To distinguish between paraphrases and quoted terms will be difficult."
- It's difficult indeed," the annotator said. He looked puzzled. "What even is a quote?"
- Sometimes it spans paragraphs without closing quotation marks.
- It is not clear if this counts as a 'real quote' or not, the said.
- He said they were not sure either. "And then the next paragraph is a quote."

**13. After she warned that quotes 'would be hard to detect,' she added: 'We will try to do better.'**

**14. The annotator got annoyed and said: 'When we thought we had listed all the quote in the file, I found this.' He said: 'Aha.'**

**15. And this? Another annotator screamed.**

**16. After she warned that quotes 'would be hard to detect,' she added: 'We will try to do better.'**

**17. The annotator got annoyed and said: 'When we thought we had listed all the quote in the file, I found this.' He said: 'Aha.'**

**18. And this? Another annotator screamed.**

**19. After she warned that quotes 'would be hard to detect,' she added: 'We will try to do better.'**

**20. The annotator got annoyed and said: 'When we thought we had listed all the quote in the file, I found this.' He said: 'Aha.'**

**21. And this? Another annotator screamed.**

**22. Annotate as **CONTENT** only the text inside quotation marks and the quotation marks (i.e. exclude text after **CUE** and before quotation marks).**

**However, the NHS Confederation, which represents the healthcare system in England, Wales and Northern Ireland, told the Guardian that immediate action was required to prevent the NHS Staffing crisis, which the electric care sector would be jeopardised.**

**23. Annotate quotes which span multiple paragraphs with a single **CONTENT** label (one quote).**

**He told BBC Radio 4's Today programme: 'We're looking at data on hospital beds. We know that the number of Covid-related deaths was not yet causing concern because it was significantly lower than during the third wave.'**

**This doesn't make sense, it just uses quotation marks to indicate a non-standard English term and for offering today a refugee 'refugee and solution', they are destined in camps and our doors are closed to local refugees. Refugees stranded in Europe, even those whose family is here and desperate to return home.'**

**For those who are not familiar with the term, it refers to the practice of separating refugees from the general population and placing them in camps or other designated areas. This is often done to prevent the spread of disease or to manage resources more effectively.**

**24. Do not include adjectives or additional descriptions in **SOURCE**.**

**Chris Gammie, of the medical group Care UK, which obtained the results under freedom of information legislation, said: 'For years on coming to promote its services as being successful, it's allowed us to project the economy; it's allowed people to get back to some semblance of normality.'**

**25. Do not include auxiliary verbs as **CUE**. (See **Q&A** was screened).**

**He told BBC Radio 4's Today programme: 'We're looking at data on hospital beds. We know that the number of Covid-related deaths was not yet causing concern because it was significantly lower than during the third wave.'**

**26. Do not label auxiliary verbs as **CUE**.**

**The CMA said it was the first time a company had 'systematically refused' to supply information under an FOI and said it 'reminds that Facebook's failure to comply was disgraceful.'**

**27. Do not label modal auxiliaries as **CUE**.**

**Last year, Facebook was criticised by the Competition Appeal Tribunal and the court of appeal for employing fully unionised staff as contractors to the CMA and the IEO.**

**28. Label indirect cues, such as 'in other words' or 'according to'.**

**According to the Center for American Progress, 20 US senators and 10 representatives refuse to acknowledge the scientific evidence of human-caused climate change.**

**29. Please verify like **is cited** should be labelled as **CUE** in the notes.**

**30. If a title function is included alongside the name, include it in **SOURCE**. ("These are the rules," said PERSON, the annotator).**

**31. Saffron Gardiner, the deputy chief executive of NHS Providers, which represents NHS hospitals, ambulance, community and social health services, said 'hard decisions' may have to be made about which patients to prioritise if Covid cases continue to rise.**

**32. If a **SOURCE** is named earlier, but there is a personal pronoun closer to the **CONTENT**, the **pronoun** should be labelled.**

**Notes:**

- The decision to extract ONLY **CONTENT** inside quotation marks might be controversial as we risk losing the meaning and context of the quote. It should be noted that we will be able to extract context if needed since the quote is identified in the **SOURCE**.
- The decision to extract context if needed since the quote is identified in the **SOURCE**.
- Identify personal and company names with job titles mentioned here needs to be an extra step to give a more accurate output.

The Annotation Guidelines for Quote Extraction

Prodigy Demo

Span Categorization

Dark

Named Entities (r)

Dark

prodigy\_demo

ner\_manual

0

0

0%

0

0

0

© 2017-2023 Explosione... (Prodigy v1.11)

Prodigy Demo

Span Categorization

Dark

Named Entities (r)

Dark

prodigy\_demo

ner\_manual

0

0

0%

0

0

0

© 2017-2023 Explosione... (Prodigy v1.11)

# prodigy

EXAMPLE Span Categorization

THEME Dark

PROJECT INFO

DATASET prodigy\_demo

VIEW ID spans\_manual

PROGRESS

THIS SESSION TOTAL

ACCEPT REJECT IGNORE

HISTORY

© 2017-2023 Explosione... (Prodigy v1.11)

# prodigy

EXAMPLE Named Entities (r)

THEME Dark

PROJECT INFO

DATASET prodigy\_demo

VIEW ID ner\_manual

PROGRESS

THIS SESSION TOTAL

ACCEPT REJECT IGNORE

HISTORY

© 2017-2023 Explosione... (Prodigy v1.11)

PERSON 1 ORG 2

The film tells three stories. The first two are about two couples whose lives are extreme opposites but they share one thing in common: an unexpected pregnancy. Carmine PERSON and Sasha PERSON live in a carefree world and the arrival of a child blows their middle-class life to the ground: behind the illusion of their fancy condominium and their successful careers hides a couple that is in pain and living with secrets and lies.

SOURCE: CMU Movie Summary Corpus

✓ ✗ ⚡ ↵

This is a demo of Prodigy, a modern annotation tool powered by active learning. For more details, see the website.

Fine-tune a pretrained model

huggingface.co/docs/transformers/training

Models Datasets Spaces Docs Solutions Pricing J

## Hugging Face

Search models, datasets, users...

### Transformers

Search documentation 83,594

V4.26.1 EN

83,594

#### GET STARTED

- Transformers
- Quick tour
- Installation

#### TUTORIALS

- Pipelines for inference
- Load pretrained instances with an AutoClass
- Preprocess
- Fine-tune a pretrained model**
- Distributed training with 😊
- Accelerate
- Share a model

#### HOW-TO GUIDES

## Fine-tune a pretrained model

Open In Colab Open Studio Lab

There are significant benefits to using a pretrained model. It reduces computation costs, your carbon footprint, and allows you to use state-of-the-art models without having to train one from scratch. 😊 Transformers provides access to thousands of pretrained models for a wide range of tasks. When you use a pretrained model, you train it on a dataset specific to your task. This is known as fine-tuning, an incredibly powerful training technique. In this tutorial, you will fine-tune a pretrained model with a deep learning framework of your choice:

- Fine-tune a pretrained model with 😊 Transformers [Trainer](#).
- Fine-tune a pretrained model in TensorFlow with Keras.
- Fine-tune a pretrained model in native PyTorch.

### Prepare a dataset

Hugging Face Datasets overview (Py...)

Fine-tune a pretrained model

Prepare a dataset

Train

Train with PyTorch Trainer

Training hyperparameters

Evaluate

Trainer

Train a TensorFlow model with Keras

Loading data for Keras

Loading data as a tf.data. Dataset

Train in native PyTorch

DataLoader

Optimizer and learning rate scheduler

Training loop

Evaluate

JournalismAI-2021-Quotes/quote-extraction

Search or jump to...

Pull requests Issues Codespaces Marketplace Explore

Watch 8 Fork 19 Starred

Code Issues Pull requests Actions Projects Security Insights

main 1 branch 0 tags

michel-ds Merge pull request #3 from JournalismAI-2021-Quotes/add-lice... 28f429b on Feb 1, 2022 4 commits

- Add quote extraction code for JAI.
- Add licence
- Remove dead repo link

annotation\_rules annotation\_scripts coreference regex\_pipeline schema LICENSE README.md README.md

Quote extraction for mod (JournalismAI collab 202

About Readme Apache-2.0 license 210 stars 8 watching

JS

List all of the source in the following news article:

Earlier this year, the chief of police in Aurora, Colo., needed dozen officers to join his force.

The chief, Dan Oates, was 50 officers short to patrol Aurora's 400,000 people just east of Denver. But he knew limiting his Colorado would not be enough: Like many other leaders in he has found that fewer people these days want to be cops.

So Chief Oates and his team began to seek recruits at a greater August, he flew about 1,800 miles to meet with New York Police officers. He convinced 14 of them to move out west.

"I feel bad raiding my home agency," said Chief Oates, who was a deputy chief in New York City. "But frankly it's a cutthroat environment now among police chiefs to recruit talent, and we all desperat

The departure of those officers was no anomaly. The New York City Department, with about 34,000 officers, has seen more resignations this year than at any time in the past two decades as the

# Automated fact checking

**Full Fact AI - Full Fact**

[fullfact.org/about/ai/](https://fullfact.org/about/ai/)

**Full Fact**

**About**

**Updates**

**Get involved**

**Training**

**ABOUT US**

# Full Fact AI

Every day fact checkers around the world find, check and challenge false claims identified by AI enabled software produced by Full Fact.

We have built scalable, robust software designed to be used by fact checkers and good information focused organisations, to save time, money and effort in identifying the most important had information to address.

**[PDF] ClaimHunter: An Unattended Tool for Automated Claim Detection on Twitter**

[semanticscholar.org/paper/ClaimHunter%3A-An-Unattended-Tool-for-Automated-Claim-Beltrán-Míguez/7643...](https://semanticscholar.org/paper/ClaimHunter%3A-An-Unattended-Tool-for-Automated-Claim-Beltrán-Míguez/7643...)

SEMANTIC SCHOLAR

Search 210,864,472 papers from all fields of science

Corpus ID: 235484140

**ClaimHunter: An Unattended Tool for Automated Claim Detection on Twitter**

Javier Beltrán, Rubén Míguez, Irene Larraz • Published in KnOD@WWW 2021 • Computer Science

As political campaigns have moved from traditional media to social networks, fact-checkers must also adapt how they are working. The explosion of information (and disinformation) on social networks makes impossible to manually fact-check each piece of data. With this reality in mind, Newtral, a fact-checking organization, has developed its own automated monitor tool for Twitter: ClaimHunter. Recently, deep learning approaches have obtained very high performance across many different NLP tasks... [Expand](#)

[View Paper](#) [Save to Library](#) [Create Alert](#) [Cite](#)

Figures and Tables

2 Citations

16 References

Related Papers

**The lessons of Squash, Duke's first automated fact-checking platform**

[poynter.org/fact-checking/2021/the-lessons-of-squash-the-first-automated-fact-checking-platform/](https://poynter.org/fact-checking/2021/the-lessons-of-squash-the-first-automated-fact-checking-platform/)

**Poynter.**

March 5, 2023

NEWS TRAINING ETHICS & LEADERSHIP FACT-CHECKING MEDIA LITERACY EVENTS

Fact-Checking IFCN

## The lessons of Squash, Duke's automated fact-checking platform

The project brought the world one step closer to live automated fact-checking, but it still has a ways to go

Start your day informed and

Speech to text, claim detection, evidence retrieval...

**And it isn't just text!**

CORONAVIRUS

# News outlets criticized for using Chinatown photos in coronavirus articles

"The usage of photos not directly related to coronavirus during coverage of the disease is not only poor reporting and decision-making in newsrooms – it also further stokes unnecessary and unfounded fears," one expert said.



Models - Hugging Face

huggingface.co/models?pipeline\_tag=visual-question-answering&sort=downloads

Search models, datasets, users...

Models Datasets Spaces Docs Solutions Pricing

Tasks 1 Libraries Datasets Languages Licenses Other

Filter Tasks by name Reset Tasks

Multimodal

- Feature Extraction
- Text-to-Image
- Image-to-Text
- Visual Question Answering
- Document Question Answering
- Graph Machine Learning

Computer Vision

- Depth Estimation
- Image Classification
- Object Detection
- Image Segmentation
- Image-to-Image
- Unconditional Image Generation
- Video Classification
- Zero-Shot Image Classification

Natural Language Processing

Models 12

Filter by name

Full-text search Sort: Most Downloads

dandelin/vilt-b32-fine Updated Aug 2, 2022 17.9k

Salesforce/blip-vqa-base Updated Dec 15, 2022 3.46k

microsoft/git-base-text Updated Jan 24 971

microsoft/git-base-vqa Updated Jan 24 688

microsoft/git-large-vqa Updated Jan 24 596 2

Salesforce/blip-vqa-capfilt-large Updated 22 days ago 209 2

microsoft/git-large-textvqa Updated Jan 24 162 1

## Multimodal

### Feature Extraction

### Text-to-Image

### Image-to-Text

### Visual Question Answering

### Document Question Answering

### Graph Machine Learning

# ⚡ Hosted inference API ⓘ

Visual Question Answering

Examples ▾



Computation time on Intel Xeon 3rd Gen Scalable cpu: cached



What's the animal doing?

</> JSON Output

Maximize

Comparing VQA Models - a Hu x +

huggingface.co/spaces/nielesr/comparing-VQA-models

Hugging Face Search models, datasets, users...

Models Datasets Spaces Docs Solutions Pricing J :

Spaces: nielsr/comparing-VQA-models like 12 Running

App Files and versions Community Linked Models Linked Datasets

# Interactive demo: comparing visual question answering (VQA) models

Gradio Demo to compare GIT, BLIP and ViLT, 3 state-of-the-art vision+language models. To use it, simply upload your image and click 'submit', or click one of the examples to load them. Read more at the links below.

image

Question

is this an image of chinatown?

Answer generated by GIT-base

['is this an image of chinatown? yes']

Answer generated by GIT-large

['is this an image of chinatown? yes']

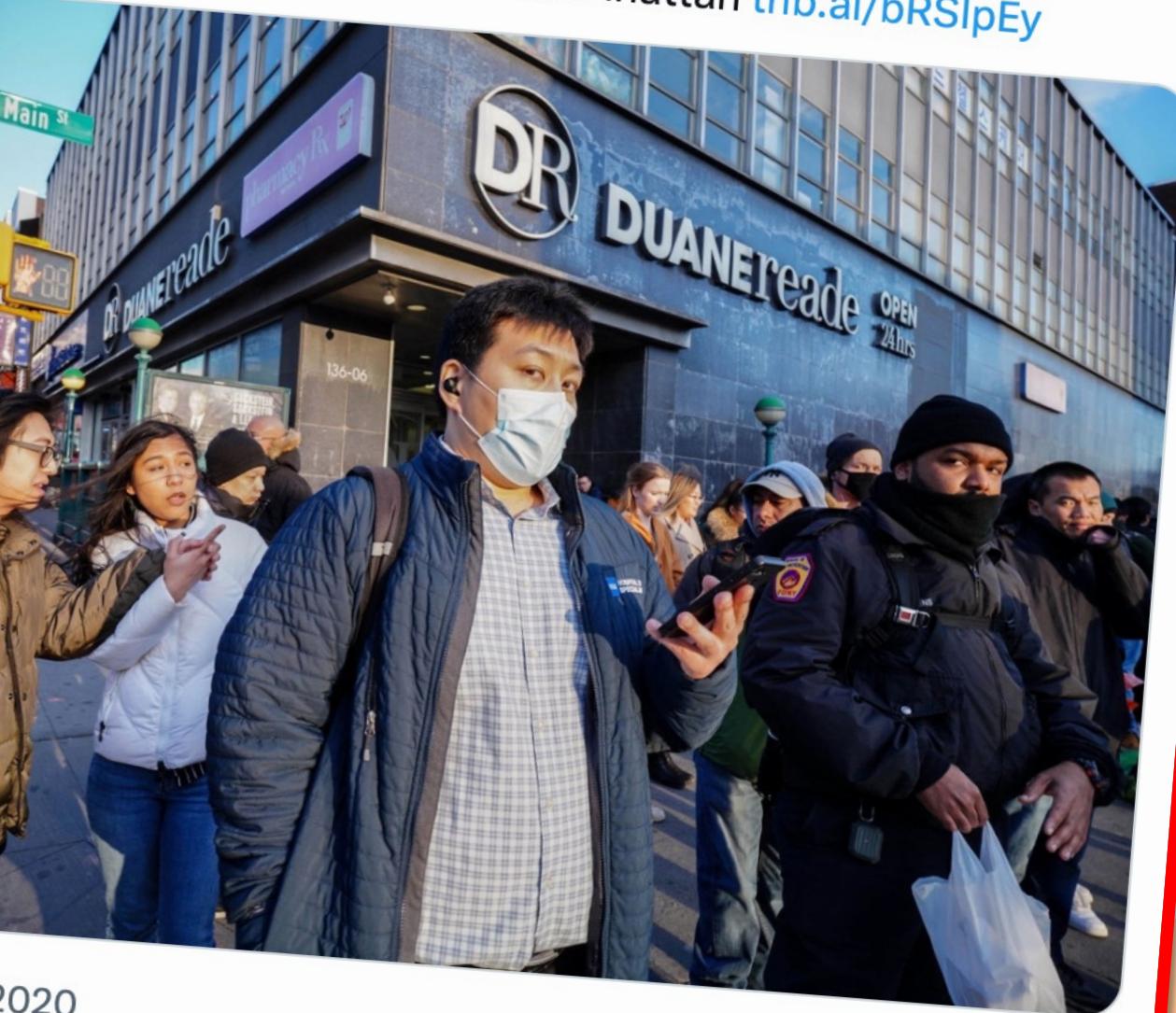
Answer generated by BLIP-base

['yes']

Answer generated by BLIP-large

['yes']

of coronavirus confirmed in Manhattan [trib.al/bRSIpEy](https://trib.al/bRSIpEy)



...

To use it, simply upload your image and click 'submit', or click one of the e

Answer generated by GIT-base

[are there any asian people in this photo? yes]

Answer generated by GIT-large

[are there any asian people in this photo? yes]

Answer generated by BLIP-base

['yes']

Answer generated by BLIP-large

['yes']

Answer generated by ViLT

yes

Prodigy Demo

← → C demo.prodigy/?=null&view\_id=imageseg

# prodigy

EXAMPLE Image (binary) ▾

THEME Dark ▾

PROJECT INFO

DATASET prodigy\_demo

VIEW ID image

PROGRESS

THIS SESSION 0

TOTAL 0

0%

ACCEPT 0

REJECT 0

IGNORE 0

CAR

SOURCE: Unsplash BY: Nabeel Syed URL: [unsplash.com/@nabeelsyed](https://unsplash.com/@nabeelsyed)

This is a demo of **Prodigy**, a modern annotation tool powered by active learning. For more details, see [the website](#).

✓ ✗ ⚡ ↵

it's just one layer, and  
it's just all guidance

# Building your own AI editors

Jonathan Soma  
Columbia University  
[js4571@columbia.edu](mailto:js4571@columbia.edu)  
[@dangerscarf](https://twitter.com/dangerscarf)