

9 hours ago - Technology

First look: Civil rights group starts center to monitor AI for hate speech

Russell Contreras, author of [Axios Latino](#)

Illustration: Sarah Grillo/Axios

A prominent civil rights group is launching a center to investigate how [artificial intelligence](#) affects civil rights, fosters racism and

fuel racism and [more antisemitism in the U.S.](#) by amplifying bias from human-generated content on the internet.

Details: The Leadership Conference on Civil & Human Rights (LCCHR), one of the nation's largest and oldest civil rights coalitions, will announce Thursday the creation of the Center for Civil Rights and Technology

- The center will monitor legislation and regulations on AI and other emerging technologies and assess how those will impact civil and human rights.
- The center also will publish papers and policy positions, and support active civic conversations about generative AI.
- An advisory group of experts and civil rights organizations will advise the center.

State of play: Today's leaders in AI, companies like OpenAI, Google and Microsoft, have aimed to filter out the most egregious racism, sexism and homophobia, Axios' [Ina Fried reports](#).

Yes, but: Others, waving free-speech or "anti-woke" banners, are already gearing up to create similarly robust systems without such limits.

How it works: Because AI models learn to complete sentences by analyzing enormous quantities of text created by people, usually on the internet, they pick up bias embedded in both the digital environment and broader society, per Axios' [Ashley Gold](#).

- Generative AI chatbots like [OpenAI's ChatGPT](#) and Google's Bard, which respond to prompts with written responses pulled

Flashback: Microsoft's Tay, a 2016 AI chatbot, was shut down after it spewed a series of [lewd and pro-Nazi tweets](#) soon after launch — a stance it borrowed from other Twitter users.

Between the lines: Civil rights organizations like the Anti-Defamation League are worried about the potential of AI to generate more hate and disinformation without safeguards like social media.

- Anti-Defamation League CEO Jonathan Greenblatt told Axios he wants tech companies to be open and transparent about AI data sets and algorithms.

What they're saying: "Hate has had a very comfortable home on our social media platforms. And it has been given cover by, sadly, sometimes, by leaders that have every ability to disrupt that," LCCHR president and CEO Maya Wiley told Axios.

- Wiley said she's observed some AI tools foster inequality, not out of hate, but because of it not being trained on systemic disparities.
- She cited an example of a company that created an algorithm to help determine who was most likely to need hospitalization. It didn't include people of color who lacked insurance or lived away from hospitals and focused mainly on white, wealthy patients.

But, but, but: Wiley said the center will also highlight examples of how AI can be used to understand civil rights better.

- [Museums](#) that focus on racial violence and antisemitism have begun using [holograms](#), [artificial intelligence](#) and [virtual reality](#) to confront difficult episodes in history.

- The center will be housed under LCCHR, which will staff it.

Go deeper: [Fighting hate with AI-powered retorts](#)

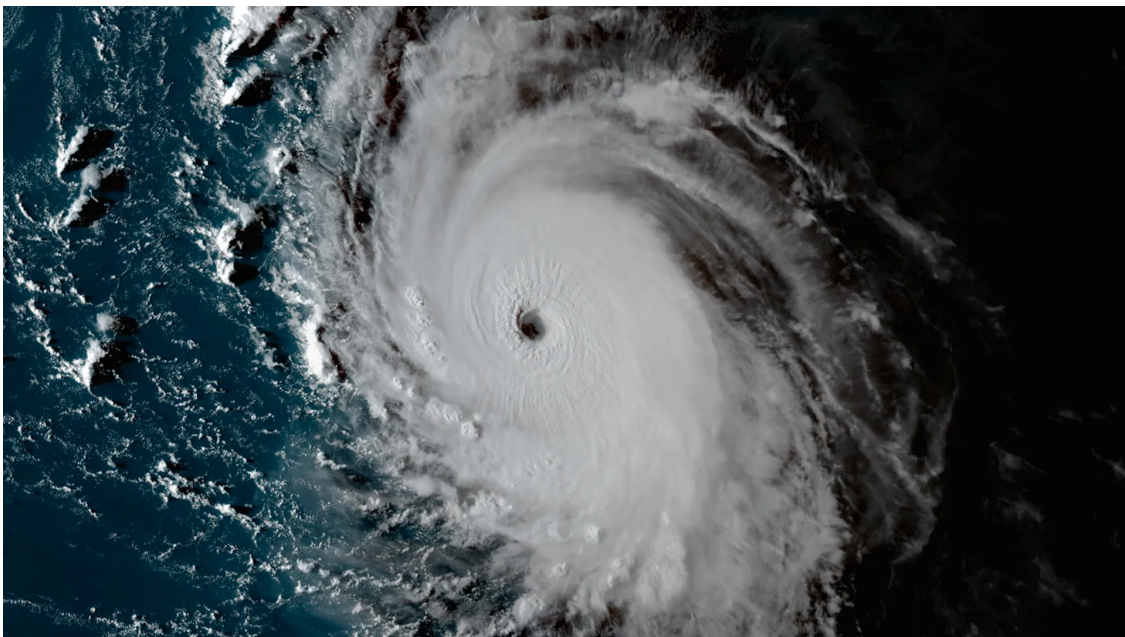


Go deeper



Andrew Freedman, author of [Axios Generate](#)
Updated 12 mins ago - Energy & Environment

Hurricane Lee intensifies at "exceptional rate," headed for Category 5



Hurricane Lee seen as a Category 4 storm near sunset over the Atlantic on Sept. 7. Image: CIRA/RAMMB

The storm was near [Category 5 intensity at 8pm ET](#) Friday, the [National Hurricane Center](#) (NHC) said.

The big picture: The powerful hurricane is forecast to move north of the Leeward Islands and Puerto Rico, and may slow down east-northeast of the Bahamas this weekend.

[Go deeper \(2 min. read\)](#) →



Nathan Bomey, Joann Muller
3 hours ago - Economy & Business

UAW, White House, automakers at odds as strike threat looms



Illustration: Shoshana Gordon/Axios

the serious prospect of a debilitating strike looming.

Why it matters: With eight days to go before the expiration of UAW deals with General Motors, Ford and Stellantis, insiders are expecting a strike that could affect one, two or all three automakers.

[Go deeper \(3 min. read\) →](#)



Sareen Habeshian

Updated 4 hours ago - Politics & Policy

Former Trump aide Peter Navarro convicted for contempt of Congress



Peter Navarro at the E. Barrett Prettyman Courthouse on Sept. 7 in Washington, DC. Photo: Kevin Dietsch/Getty Images



subpoena from the House Jan 6. committee, multiple outlets reported.

The big picture: Navarro became the [second high-ranking former Trump](#) official to be convicted in a case related to attempts to overturn the [2020 election](#).

[Go deeper \(1 min. read\) →](#)



News worthy of your time.

[Download the app →](#)

About

[About Axios](#)

[Advertise with us](#)

[Careers](#)


[Events](#)

[Axios on HBO](#)

[Axios HQ](#)

[Privacy and terms](#)

[Online tracking choices](#)

 Your Privacy Choices

[Contact us](#)

Subscribe

[Axios newsletters](#)

[Axios Pro](#)

[Axios app](#)

[Axios podcasts](#)

[Courses](#)

[Earn Axios rewards](#)

