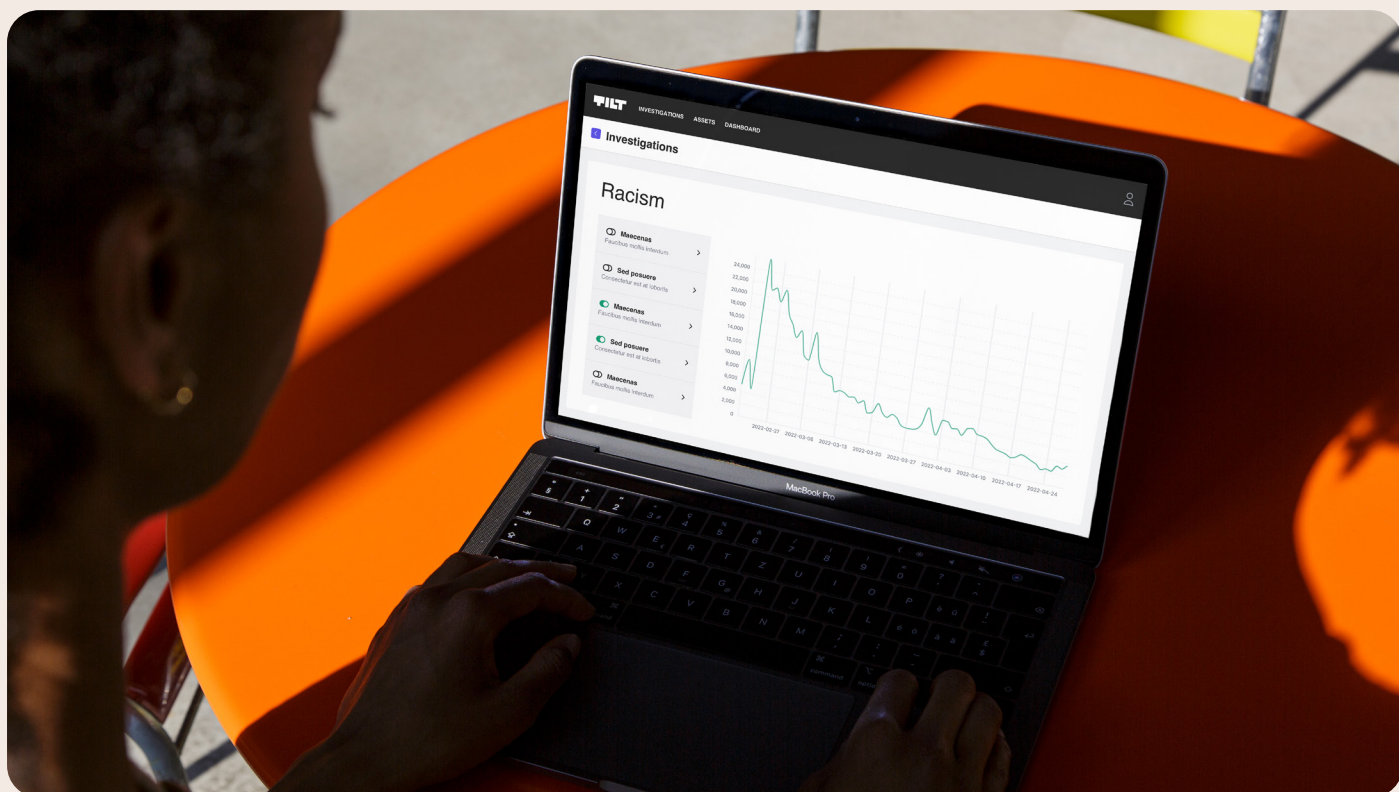


Hate Speech Detection



Tackling hate speech online, reducing real world harm

The internet is a powerful tool to connect people in online public spaces. The anonymity involved creates an atmosphere to share ideas, feelings or opinions in a safe way but simultaneously grants an environment for online hate to flourish. These online nurtured hateful sentiments have an impact in the offline world. Since online hate speech is rising, the offline impact did as well in the form of increased hate crimes and (verbal) assaults or insults.

At Tilt, we believe that the world should be a safe and inclusive place for everyone, regardless of race, religion, gender, or sexual orientation. That's why we're on a mission to make the internet a safer place by helping you to detect, research or combat online hate speech.



Stopping hate speech in its tracks

We understand that hate speech isn't just offensive, it's dangerous because it leads to harmful offline consequences. Detecting hateful content in a fast and efficient way is therefore a must. With Tilt's Hate Speech Detection solution, you can take action against hate speech before it grows online or before it has a chance to escalate into serious offline incidents. We've developed our online Research Tool METIS that enables you to detect hateful messages shortly after it has been posted on open social media platforms such as Twitter, YouTube, Instagram, Reddit or open Telegram groups.

Cut through the noise

Monitoring large amounts of online content is a daunting task. That's where Tilt comes in. Our solution makes it easy for you to detect and analyse hate speech with minimum effort. By speeding up and automating the process, we enable you to focus on taking action against harmful content, making the internet a safer place for everyone.

Some examples

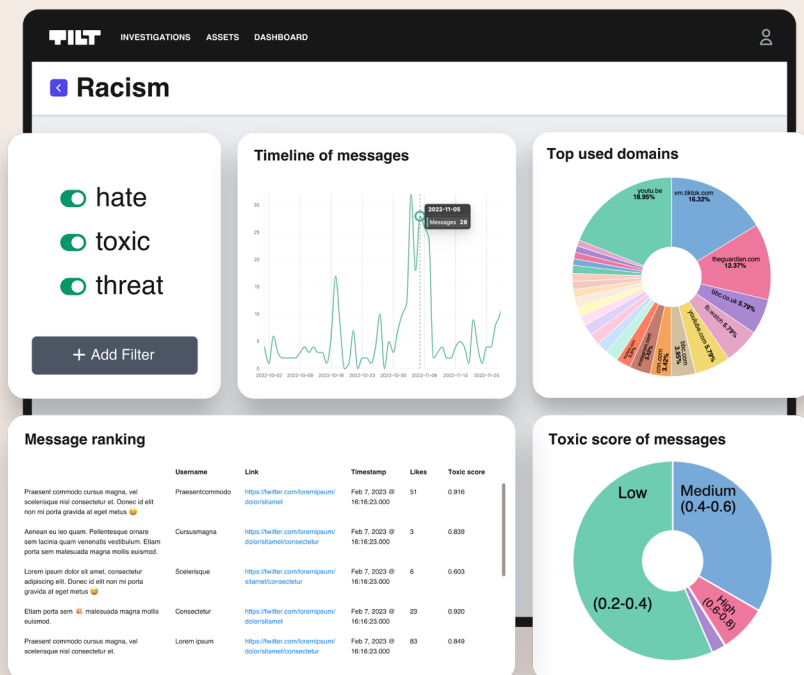
- Protecting free press and democracy by monitoring online hate speech targeting journalists and politicians
- Detecting hate speech and mobilisation language to prevent (escalating) riots
- Detect and analyse messages containing hateful content such as racism, anti-Semitism, homophobia, Islamophobia, misogyny and extremism
- Tracking online indicators of hateful content linked to offline events
- Measure the impact of interventions taken in online spaces
- Help trusted flaggers to monitor if reported content is being removed

Benefits

- Automate detection of hate speech in all languages
- Filter and rank hateful, toxic and violent content
- Visualise Networks, Patterns and Relationships
- Set up a thorough and customised investigation
- Narrative evolution, from past to present

GDPR compliant

- METIS is developed by the principle of Privacy by Design
- Ethical and GDPR compliant investigation methodology and workflow
- Checked by an external Data Protection Officer





Our team

Our multidisciplinary team is passionate to help you make a difference. By combining social sciences with artificial intelligence, we provide solutions that work in practice with a minimal chance of bias. At Tilt, our team consists of individuals from diverse social, cultural, and academic backgrounds. Together we work on innovative ways that help our clients recognize (and deal with) online challenges. Based on science, and verified to be effective.

Work with us

Do you want to detect, analyse, report or combat online hate speech? We are happy to provide you with our solution to become fast, efficient and effective in doing so. It's also possible to work with our analysts for high quality analytics and customised reports upon online hate speech. With Tilt's Hate Speech Detection, you are one step closer to stopping hate speech in its tracks. Let's talk!



Interested in Hate Speech Detection? Please contact Esther at esther@tiltstudio.co

Innovators in
information resilience

www.tiltstudio.co