

Press release

Bern, 27 February 2023

## Stop Hate Speech: alliance F and ETH Zurich set up Switzerland's first foundation for online public discourse

alliance F and ETH Zurich are launching the [Public Discourse Foundation \(PDF\)](#), which will take an in-depth look at hate speech and its impact on democracy. The aim is to research and strengthen public discourse on the internet. The foundation's first projects will support media companies with comment moderation and empower civil society to effectively deal with hate speech through an interactive knowledge centre.

Hate speech – that is, toxic language, insults and discriminatory language directed at people because of the way they look or how they identify – has become a widespread problem. alliance F launched the [Stop Hate Speech](#) project in 2020, supported by the Migros Pioneer Fund, to identify online hate speech and to develop effective counterstrategies. Now alliance F and ETH Zurich are establishing the non-profit Public Discourse Foundation (PDF), which aims to research and strengthen public discourse on the internet. Mercator Foundation Switzerland is supporting the foundation with a contribution of CHF 300,000 over two years.

*“Hate speech is a threat to our democracy. We cannot have people being intimidated by a loud minority. The Public Discourse Foundation helps ensure that everyone can take part in fair debates, thereby strengthening democracy.”* – Sophie Achermann, new Managing Director of the Public Discourse Foundation.

Scientific insights and scalable solutions can help to reduce toxic language and promote constructive dialogue, as Dominik Hangartner, Professor of Public Policy at ETH Zurich and Chair of the Foundation's Board of Trustees, explains: *“Today, we have the scientific and technical knowledge to be able to make a decisive contribution to this vexing issue. However, only by working in close cooperation with civil society and media companies are we able to successfully translate our findings into practice and generate positive impact. This is where the Public Discourse Foundation comes in and can serve as bridge builder.”*

## **Strengthening freedom of speech and variety of opinion in comment sections**

To establish this solution-centered approach right from the start, the foundation is striving for close cooperation with media companies in German-speaking Europe and in French-speaking Switzerland. The comment sections of media websites feature a lively exchange of opinions, but the moderation of these comment sections is a major challenge for media companies. This is where the foundation would like to make a first contribution by providing transparent algorithms to support the media companies in this difficult task.

The **Public Discourse Foundation** will begin its work on 1 April 2023 and will be based in Zurich. ETH Zurich is supporting the foundation through targeted research in this area and is contributing CHF 100,000 to its establishment. Stop Hate Speech's ongoing project work will be incorporated into the foundation. Dominik Hangartner, Professor of Public Policy at ETH Zurich, will become Chair of the Foundation's Board of Trustees, and Sophie Achermann will leave her position at alliance F at the end of March 2023 and take up the role of Managing Director of the Public Discourse Foundation.

### **For further information:**

#### **Sophie Achermann**

CEO of alliance F, designated Managing Director of the Public Discourse Foundation

Phone: +41 79 274 67 53, email: [sophie.achermann@public-discourse.org](mailto:sophie.achermann@public-discourse.org)

#### **Dominik Hangartner**

Professor of Public Policy, ETH Zurich, designated Chair of the Board of Trustees of the Public Discourse Foundation

Phone: +41 76 540 27 33, email: [dominik.hangartner@gess.ethz.ch](mailto:dominik.hangartner@gess.ethz.ch)

### The **Stop Hate Speech** project

The Stop Hate Speech project initiated by alliance F, in cooperation with ETH Zurich, the University of Zurich and a community of 1,200 volunteers, has developed an algorithm (Bot Dog) that automatically detects hate speech, using more than 420,000 annotated comments as training data. Strategies for successful counterspeech were also tested and refined in several field experiments. By combining civil society engagement with behavioral insights and technological advances, the project is pioneering solutions for dealing with harassment and abuse on the internet. The project was also based on a citizen science approach, meaning that volunteers could actively participate by checking comments for hate speech to help improve the algorithm.

The Stop Hate Speech project was made possible by the Migros Pioneer Fund, supported by the Raiffeisen Jubilee Foundation, the Ernst Göhner Foundation, the Gebert Rüt Foundation, Innosuisse and various cities and cantons. The establishment of the foundation is supported by the Mercator Foundation Switzerland.

Further information can be found at: [www.public-discourse.ch](http://www.public-discourse.ch).