Deepfakes & Algorithms

Threat or Opportunity?

November 2020

(club) prAxis

12 recommendations

to reposition Europe as an expert on content generation algorithms and a leader in the fight against deepfakes

Abstract

Nowadays, deepfakes appear to be a manipulation tool whose **impact on society is still poorly understood**. Their existence and use raise many **legal** and **ethical** questions. Despite the laws and governance rules that may be implemented in response, the **inability to detect a deepfake** remains a fundamental concern. As technology continues evolving, **it becomes more and more complicated to identify a fake**. Developing a European knowledge on tools for detecting fakes and authenticating originals appears urgent.

To answer this challenge, the latest Praxis report, *Deepfakes & Algorithms*, makes twelve recommendations around **four major strategic pillars**:

- Making Europe **a leader in the fight against deepfakes**

- **Strengthening the responsibility** of platforms at the European level

- Building a **regulatory environment** adapted to an efficient fight against deepfakes

- **Protecting citizens** from the impact of deepfakes

Download the full report here:
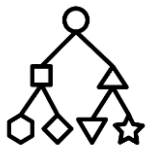https://www.clubpraxis.com/publications

About Club Praxis

Club Praxis is a French think tank based in New York that brings together expatriate businesses and public policy experts whose experience abroad helps shedding light on public debate in France. Founded in 2007 by several French leaders, including Henri de Castries, Club Praxis has embraced its role as an outsider in French discourse. Its mission is to promote new and innovative ideas that will reinvigorate both democratic institutions and economies.

(club) prAxis

1st Pillar
Making Europe a leader in the fight against deepfakes

1. Support the development and improvement of techniques for classifying content, whether true, false, or suspect (fact-checkers, watermarks, algorithms, blockchain), while **defining European standards in the meantime**, to support the population in identifying reliable content.

2. Invest in scientific research on content generation algorithms, to enable a better capacity to fight against them, but also use them for positive applications for society.

3. Foster international partnerships, mutualize resources to provide the means for catching up with the leap of China, Russia, and the United States in this area.

4. Strengthen the responsibility incurred by platforms (for instance within the Digital Services Act's framework). Compel each platform to set up an internal governance structure for the regulation of its content. Place each governance body under European supervision, drawing inspiration from the regulations put in place in the banking sector following the 2008 crisis.

5. Encourage platforms to implement tools to reduce the impact of deepfakes.

Sanctions such as:

- A penalty for authors posting false or malicious content in a non-explicit manner.

- A clear statement warning the user that the post could not be verified, or even that it is suspicious.

- A limit on the number of recipients of a publication.

- A limitation of transfers to one post per user.

6. Compel platforms to share information they hold on alleged deepfakes perpetrators with judicial authorities as part of criminal investigations, with no possibility for the platforms to oppose any lawful reason for refusal.

(club) prAxis

7. Facilitate legal access for victims of deepfakes, identity theft or harassment (to legal remedies and information on their rights) through the development and visibility of **support structures** (both physical receptions and virtual state platforms).

8. Strengthen the existing criminal sanctions against the authors of deepfake content and **reassess** them, taking into account **the new technological and digital environment** in which the creation of deepfakes is made possible.

9. Strengthen at the state, and above all, supra-state levels (the transmission of information knowing no borders) the platforms' responsibility. To do so:

- **Develop a supranational organization for police and justice cyber-cooperation focused on the topic in order to facilitate the exchange of information between States**

- Allow in some exceptional cases, **such as the broadcast of clearly illegal content, to order the permanent closure of a site**, a profile, an account, a page, a group, with no possibility of recourse for the incriminated party

- **Limit and legally quantify the withdrawal deadline for illegal content** reported to platforms under the threat of heavy financial penalties

- Require platforms **to verify** and retain identification data when creating an account and, **in the absence of actual identification of the author, close the account without delay**

- Require hosts located outside of France to designate a representative in France to take responsibility should they fail to withdraw content reported as illegal within the legally fixed deadline

- Systematically combine the financial sanctions imposed on platforms with an advertising measure

(club) prAxis

4th Pillar
Protecting citizens from the impact of deepfakes

10. Make the fight against deepfakes a European objective for 2021. Set up **awareness campaigns** to educate about deepfakes, their uses, and the associated risks in terms of the misappropriation of images, voices and videos, while emphasizing the possible applications in opinion manipulation.

11. Reinforce the responsibility incurred by the authors. Require each author to clearly signal any deepfake content as such.

12. Involve employers in the fight against deepfakes. Encourage them to **adopt ethical charters**, to inform their employees, customers, suppliers, and partners of their good ethical practices on the matter. **Encourage the implementation of trainings around deepfakes.**

(club) prAxis

About the authors

Supervision

Emilie Rannou holds a double Master degree in Applied Mathematics and Statistics from ENSAE school and University Paris Dauphine. She manages a team of data scientists. Emilie, a former Executive Director at JPMorgan Chase and Senior Product Manager at Criteo, has worked 13 years in the tech and financial industries. As a senior manager at Ekimetrics, she defines and sets up new models and advises companies on their data science strategy.

Work Group

Alexis Benichoux has a PhD in signal processing from the University of Rennes. He has worked on the learning methods applied to sound signal processing. Alexis is also a graduate of the Master of Apprenticeship at ENS Cachan. Currently a research engineer at Yubo, he develops image recognition algorithms and content recommendation.

Rémi Forgeas oversees the France Business Services Center for an international audit and consulting firm. He's been based in the United States for over twenty years, where he assists French groups in their local development. He is a graduate of Sciences Po Paris and holds a master's degree in econometrics.

Simon Gaillard is a consultant in data protection. Based in New York City, he specializes in online privacy and helps companies complying with the GDPR and the CCAC. Being a graduate of Sciences Po Paris, he also advises international organizations on public policy, new technologies, and data protection issues.

Jérémie Mary has a PhD in computer science from the University of Paris-Sud. He is a lecturer, authorized to direct research at the University of Lille, and is currently on a provisional assignment as a researcher at Criteo. He is an associate member of the Inria SequeL team which is specialized in sequential decision making. He has won and organized 3 challenges for recommendation systems in major machine learning conferences using Yahoo! data, Adobe, and Twitter (ICML'11, ICML'12, RecSys'14). His current research work evolves around recommendation sequences and the use of generative models for recommendation systems.

(club) prAxis

About the authors

Minh Trinh is a graduate from the Ecole Polytechnique and the University of Paris-Dauphine and holds a PhD in economics from Harvard University. He is a consultant based in New York and specializes in quantitative methods and machine learning applied to economics and finance. He is the author of a book on AI project management. He has worked previously for Lehman Brothers, Neuberger Berman, Bloomberg and Kroll Bond Rating Agency. He is a member of the American Finance Association and the CFA Institute.

Gabriel Turinici is a former student of the ENS Ulm-Paris (G/S 95) and former member of the Institut Universitaire de France. He was a senior researcher at INRIA until 2005 and currently is a professor at Université Paris Dauphine - PSL. He is specialized in numerical simulation algorithms with applications in quantum physics, medicine, artificial intelligence and finance. He is the author of more than a hundred scientific publications (5400+ Google scholar citations) and a member of the Editorial Board of several peer-reviewed journals.

Emilie Waxin has been a business lawyer since when she was admitted to the Paris Bar in 2010. After 8 years being an associate in several law firms in France and in Luxembourg, she created her own law firm, WE Avocats, in 2017. She advises and represents her clients in the fields of commercial law, labour law, and IP law. She also teaches litigation within the Paris Bar School (EFB) and is a member of several societies such as AFFJ (Association Française des Femmes Juristes) and IDE (Institut de Droit Equin).

Club Praxis

Yann Coatanlem is the President of Club Praxis. He graduated from the National School of Computer Science and Applied Mathematics of Grenoble (ENSIMAG) and holds a Master of Finance from HEC. He is director of research for a financial institution, president of Maison Française de New York University, President of the American Foundation of the Paris School of Economics, Foreign Trade Advisor and Director of the European-American Chamber of Commerce.

In collaboration with:

Ekimetrics.

(club) prÁxis