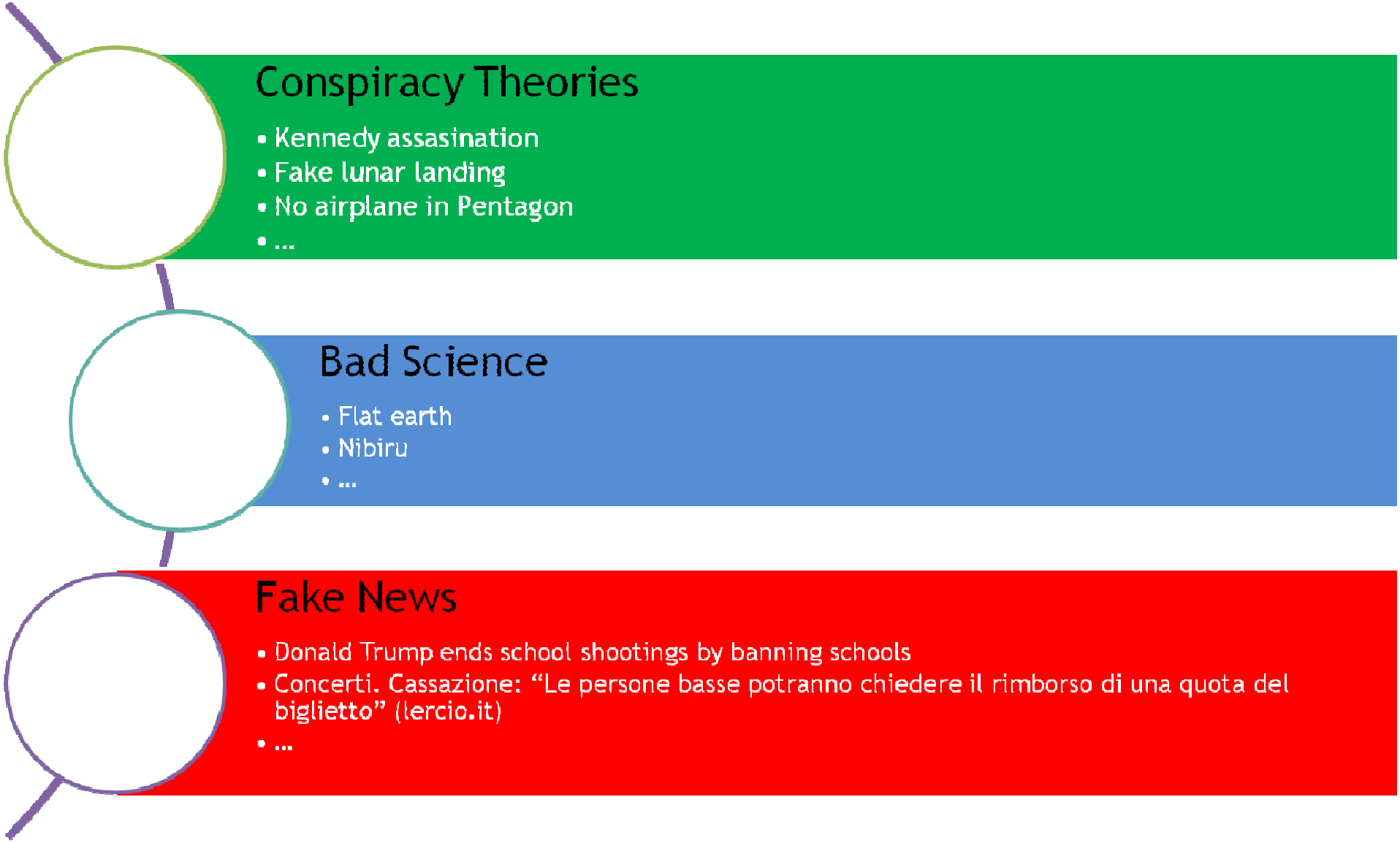
Ferrara 28 Novembre 2019

Deep Fake, cosa sono e come proteggersi.

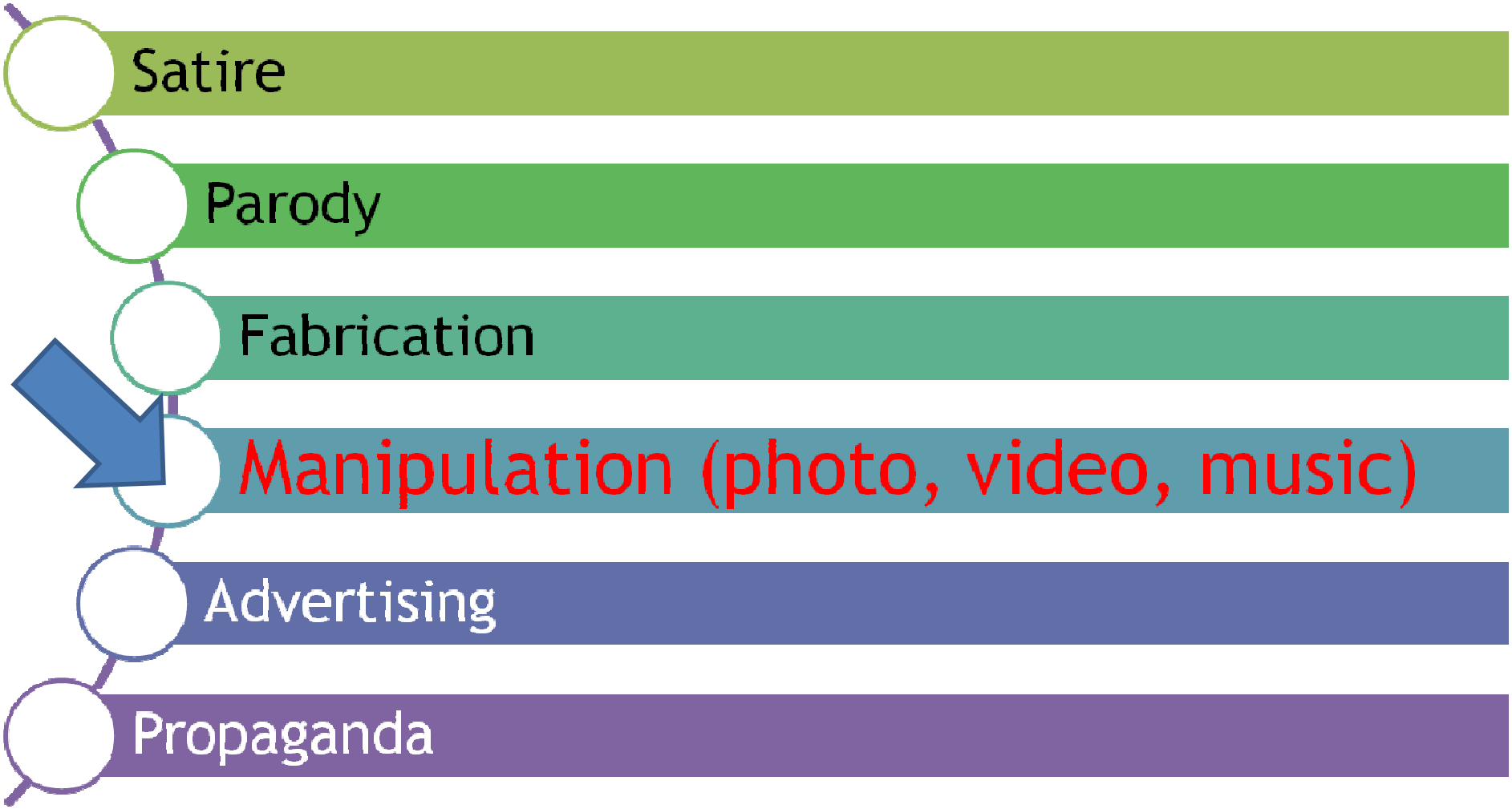
Giorgio Pedrazzi Cineca

Dalle teorie del complotto alle Fake News

Dalle teorie del complotto alle Fake News

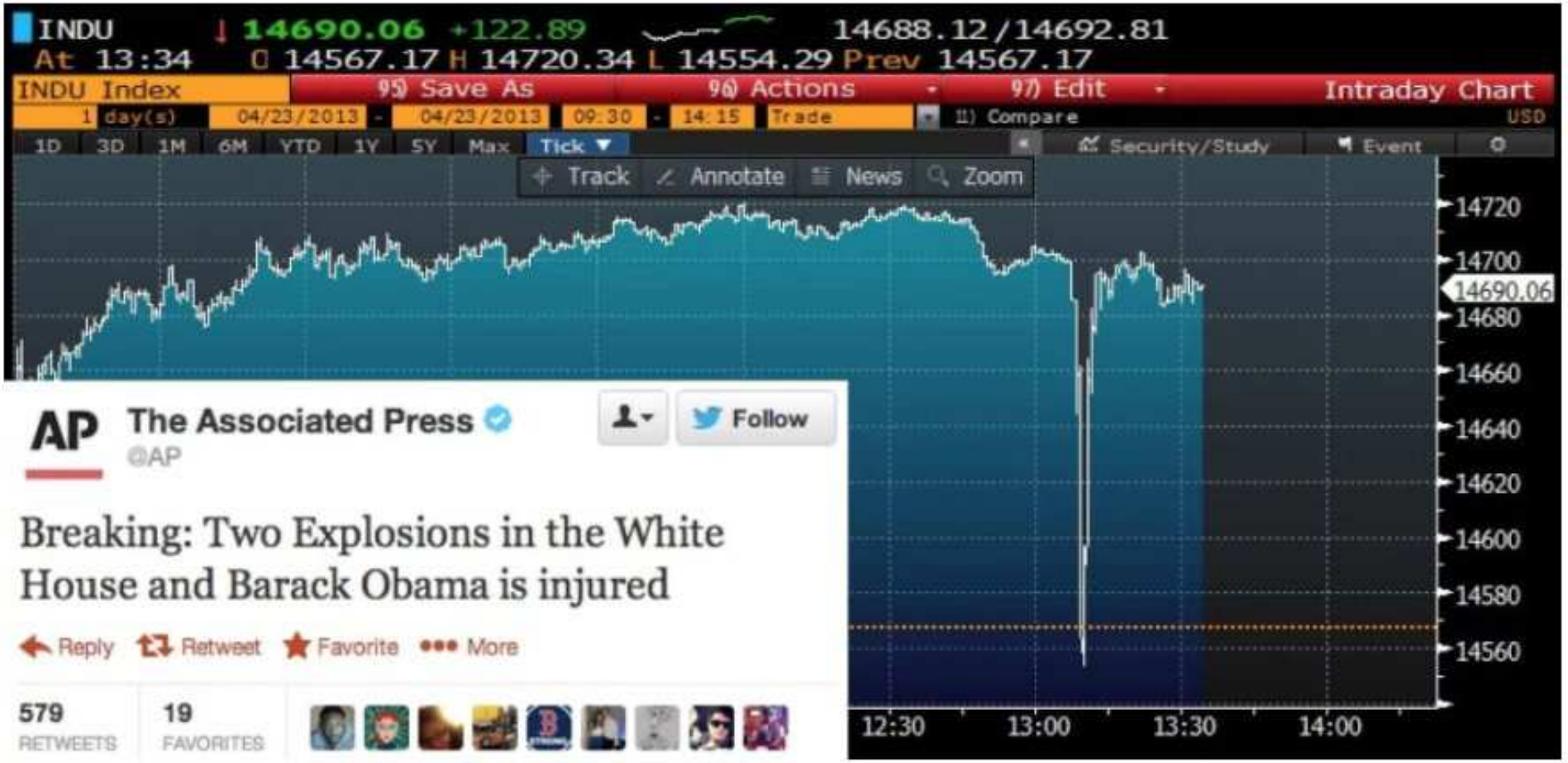


Classificazione delle Fake News



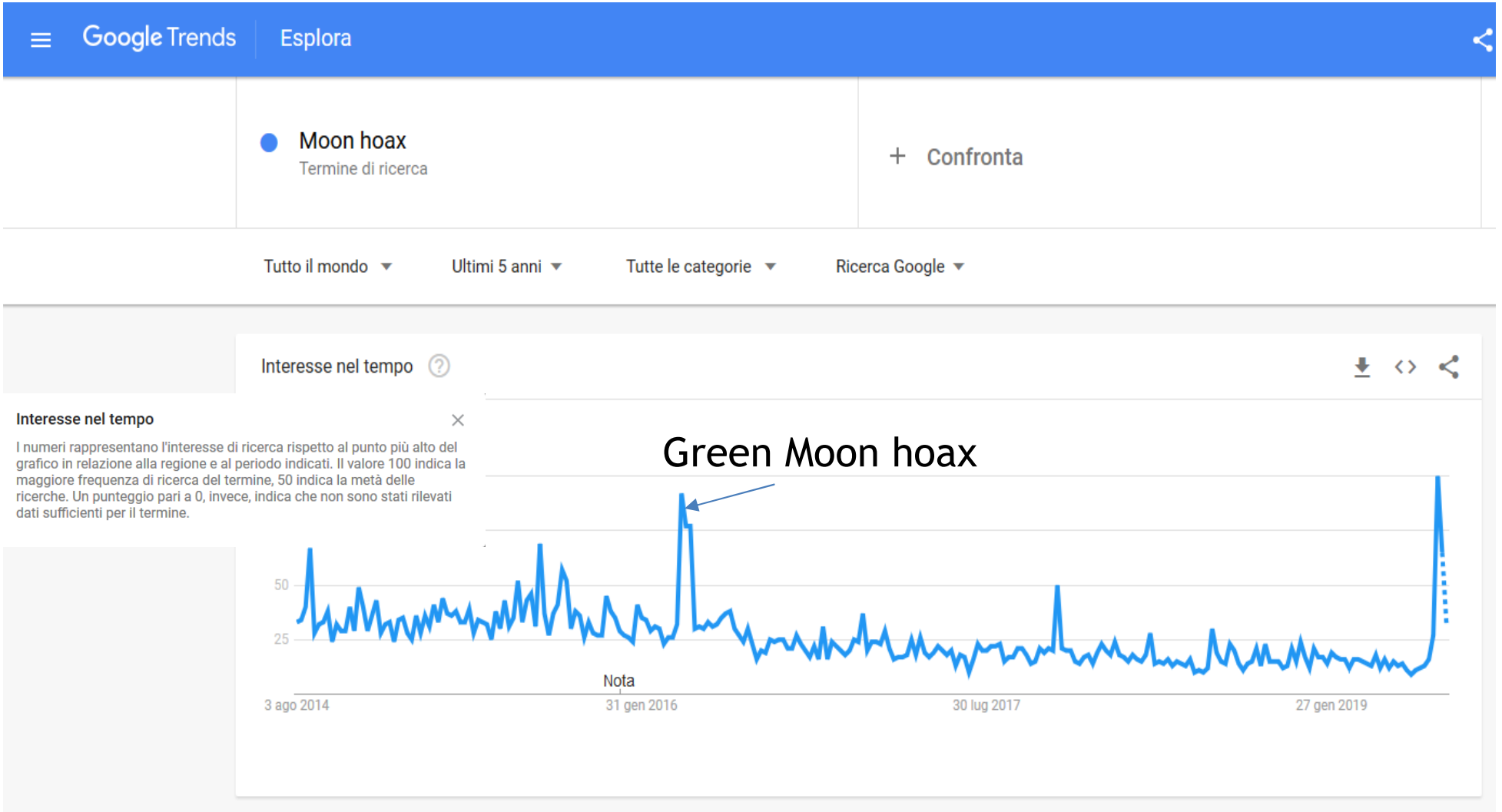
Tandoc, E. C., Lim, Z. W., & Ling, R. (2018, February 7). Defining “Fake News”: A typology of scholarly definitions. *Digital Journalism*. Routledge.

Fake News



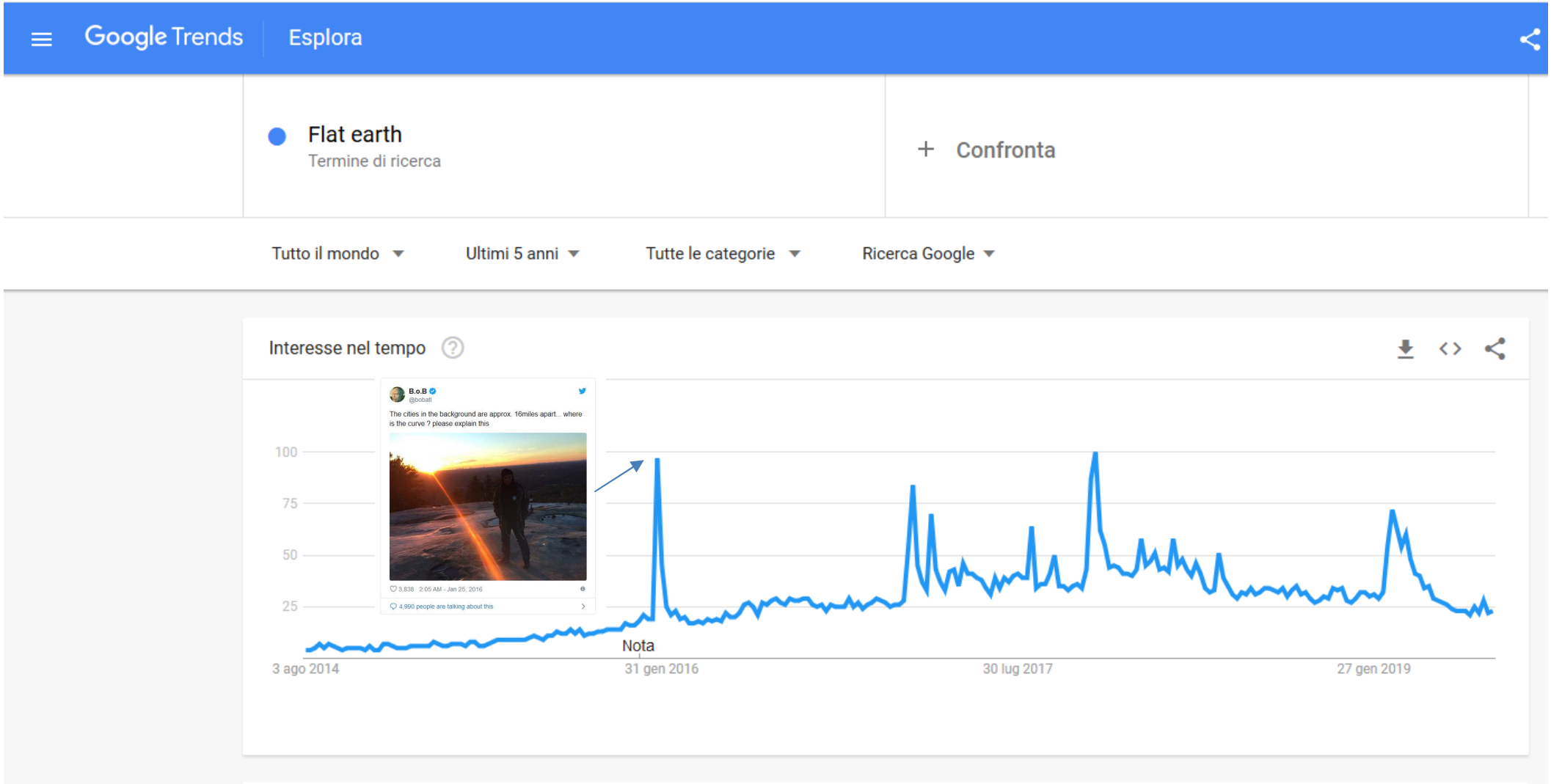
https://www.marketwatch.com/story/this-day-in-history-hacked-ap-tweet-about-white-house-explosions-triggers-panic-2018-04-23

Ricerca dei termini



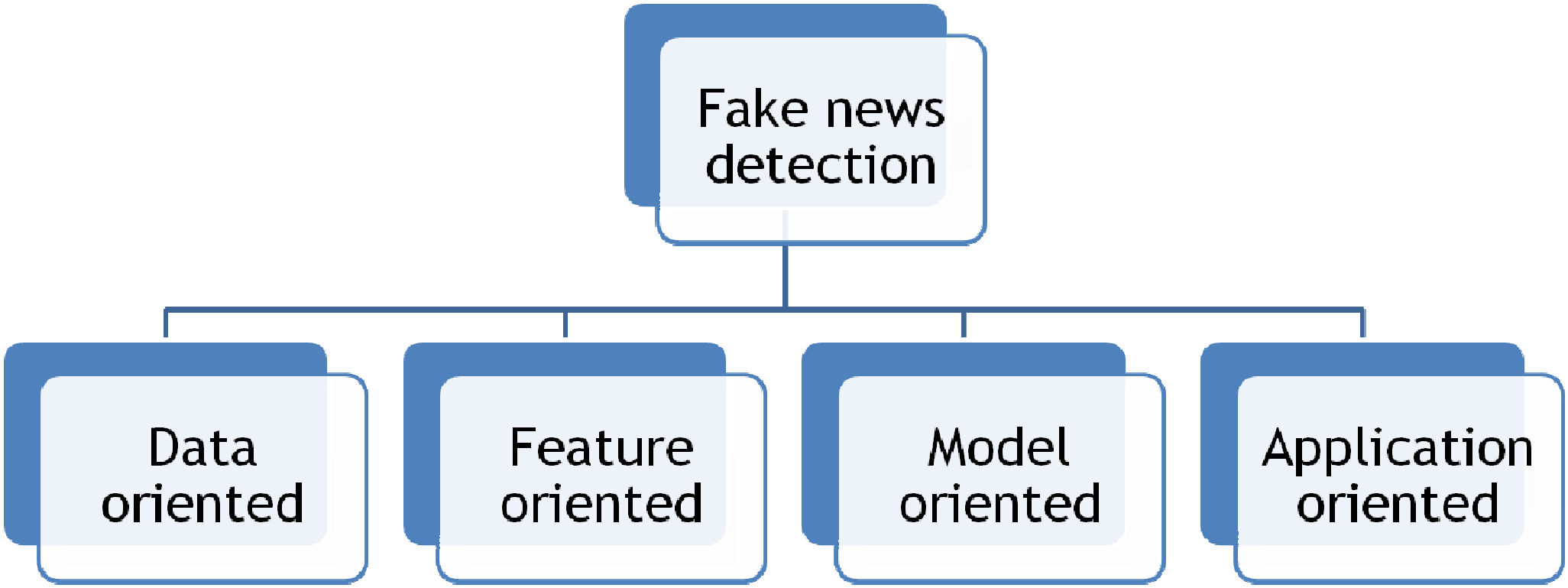
Fonte dei dati: Google Trends (https://www.google.com/trends)

Ricerca dei termini



Fonte dei dati: Google Trends (https://www.google.com/trends)

Fake News detection



Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36.

Deep Fake

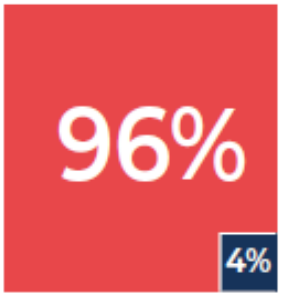
- **Le nuove tecnologie informatiche, specialmente il Deep Learning hanno aperto un nuovo fronte permettendo la generazione di volti di persone non esistenti realistici o la sostituzione di volti in filmati (es. il volto di Jim Carrey sostituisce quello di Jack Nicholson in Shining).**
- https://www.youtube.com/channel/UCKpH0CKItc73e4wh0_pgL3g
- http://www.whichfaceisreal.com/index.php
- https://www.creativebloq.com/features/deepfake-examples

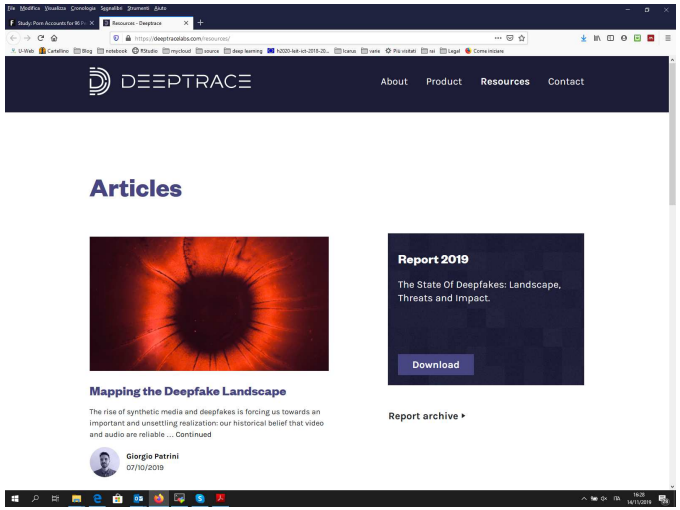
Deep Fake: qualche numero

Total number of deepfake videos online

14,678

percentage of deepfake videos online by pornographic and non-pornographic content



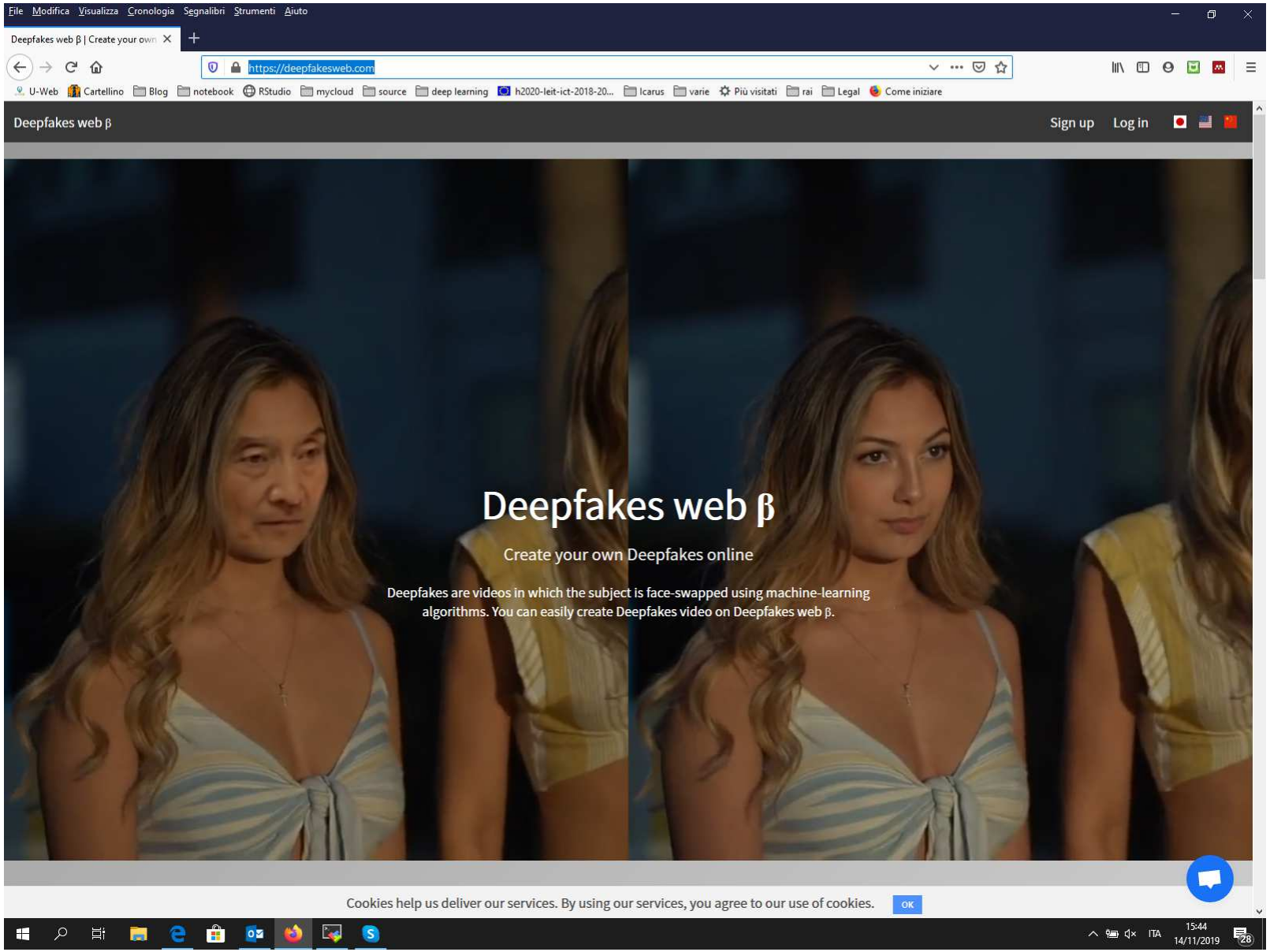


Total number of video views across top four dedicated deepfake pornography websites

134,364,438

Fonte: https://deeptracelabs.com/resources/ Deeptracelabs report 2019

Creare Deep fake online



Come riconoscere un Deep Fake

- **Presenza di artefatti**

Fonte http://www.whichfaceisreal.com/index.php

Come riconoscere un Deep Fake

- **Background e occhiali**

Fonte http://www.whichfaceisreal.com/index.php

Come riconoscere un Deep Fake

- **Per i video esistono algoritmi per l'individuazione basati sulla frequenza con cui una persona sbatte le palpebre normalmente. L'algoritmo identifica il Deep Fake poiché nel video rielaborato il soggetto non sbatte le palpebre come nella realtà.**
- **Per difendersi attivamente si possono inserire pixel di disturbo per ingannare gli algoritmi che riconoscono le facce.**

Fonte : https://www.wired.it/internet/web/2019/07/15/deepfake-video-difesa/?refresh_ce=

Deep Learning GAN

GAN è l'acronimo di **Generative Adversarial Network (rete antagonista generativa). Le GAN implementano due reti che si sfidano l'una con l'altra.**

- **La rete "Generativa" produce immagini che siano indistinguibili da quelle reali**
- **La rete "Antagonista" sottopone a verifica le immagine generate dalla rete "Generativa" .**

Alla fine della fase di apprendimento il modello generato è in grado di proporre immagini verosimili con un certo grado di accuratezza.

Applicazioni GAN

- **Generate Examples for Image Datasets**
- **Generate Photographs of Human Faces**
- **Generate Realistic Photographs**
- **Generate Cartoon Characters**
- **Image-to-Image Translation**
- **…**
- **Super Resolution**
- **…**

https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/