



REAL or FAKE: DEEPFAKES IN THE AGE OF AI





In the age when **seeing** is no longer **believing**







Zephania Reuben

Data Scientist & AI Trainer




Outline

1. State of Deepfakes: [Overview](#)
 2. Deepfakes
 3. Tech Behind
 4. Usage
 5. The Worry
- 



Since emerging in late 2017 the phenomenon of **deepfakes** has developed rapidly, both in terms of **technological sophistication** and **societal impact**.





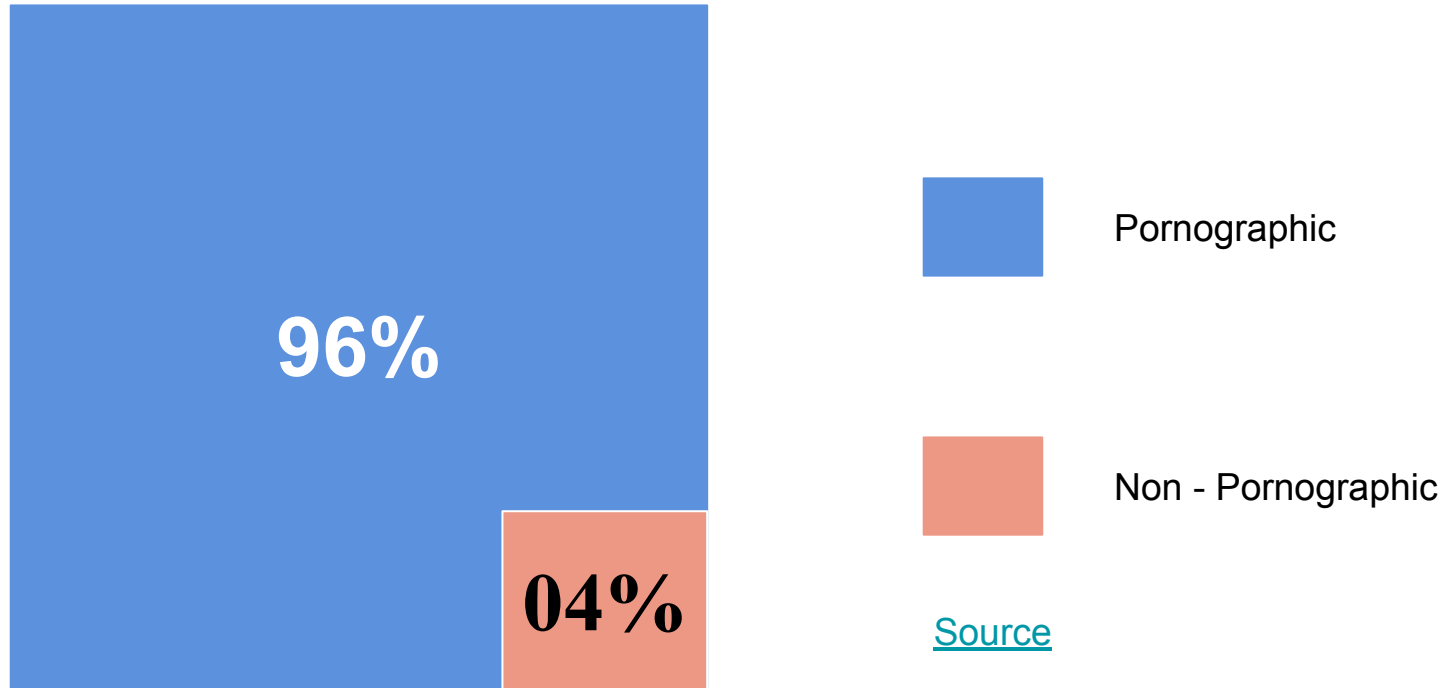
Total number of **deepfake** videos online

14,678

[Source](#)



Percentage of deepfake videos online by





What is Deepfakes

A **deepfake** refers to a specific kind of **synthetic media** where a person in an image or video is swapped with another person's likeness.

- Meredith Somers





Reflect

What if you appear **doing** or **saying** something you didn't say or do?





Deepfakes Forms



Reenactment

Using your facial or body movements to dictate the movements of another person. ML Models [MarioNETte](#) and [FLNet](#).



Replacement

One person's **identity** is mapped to another **person** (for example the face swap filter on Snapchat). **ML Models** **FaceShifter** and **FaceSwap-GAN**.

Face Replacement



Transfer

Swap

Editing

Altering the **attributes** of a **person** in some way (like changing their age or glasses).

Face Editing



Synthesis

Generating completely **new images** or videos with no **target person**.





The **Tech** Behind





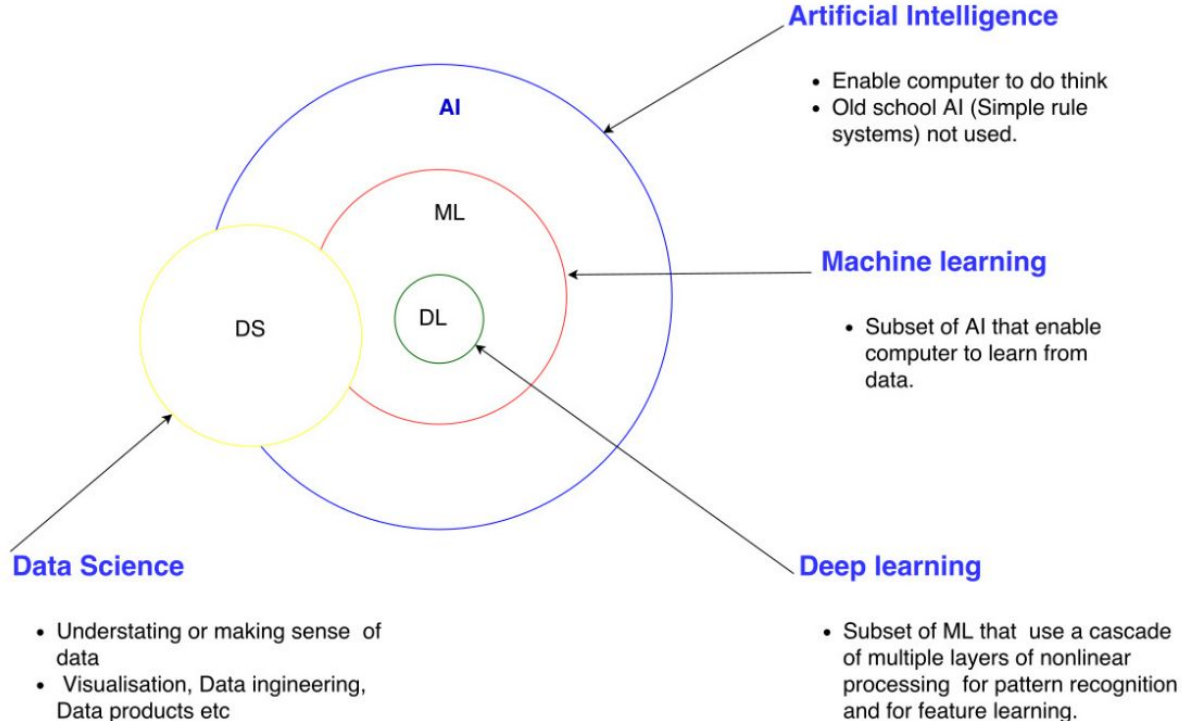
Artificial Intelligence (AI)

Is a way of making **machine** behave in the way that would be called **intelligent** if **human** were so behaving.

- John McCarthy, 1957.



Artificial Intelligence (AI) Approaches






Machine Learning (ML)

Is the field of study that gives computers the ability to learn without being explicitly programmed by

- Arthur Samuel, 1959.
- 




Machine Learning Algorithms

- Linear Regression
 - Logistic Regression
 - Decision Trees
 - Random Forest
 - Neural Networks ...
- 




Deep Learning (DL)

Is a subfield of machine learning (ML) in artificial intelligence (AI) that deals with algorithms inspired from the biological structure and functioning of a brain to aid machines with intelligence.





Different Types of NNs

- Regular Neural Networks
 - Convolutional Neural Networks
 - Recurrent Neural Networks
 - Graph Neural Networks
 - Generative Adversarial Neural Networks
- 



Generative Adversarial Networks (GANs)

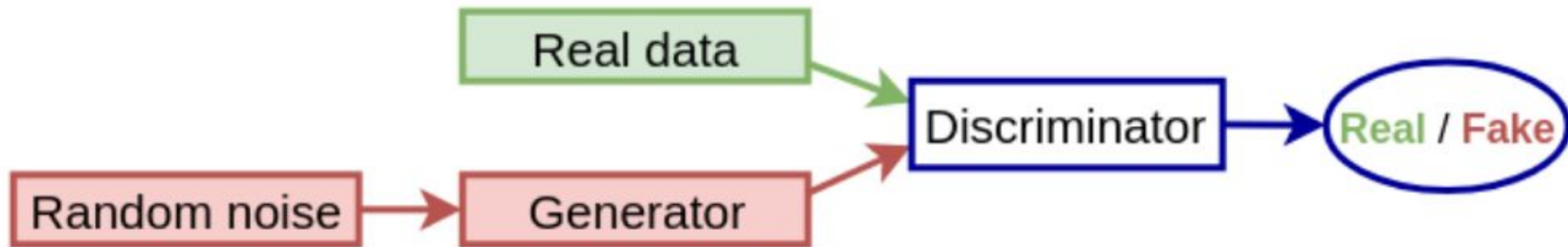
A **GAN** is a system of two components (**neural networks**)

Generator

Discriminator



Generative Adversarial Networks (GANs)






Generative Adversarial Networks (GANs)

Generator:

This is the **generative model** itself. It takes a probability distribution (**random noise**) as input and tries to generate a realistic output image.





Generative Adversarial Networks (GANs)

Discriminator:

This takes two alternating inputs: the real images of the training dataset or the generated fake samples from the generator. It tries to determine whether the input image comes from the **real images** or the **generated ones**.





How **Deepfakes** Are Being Used?





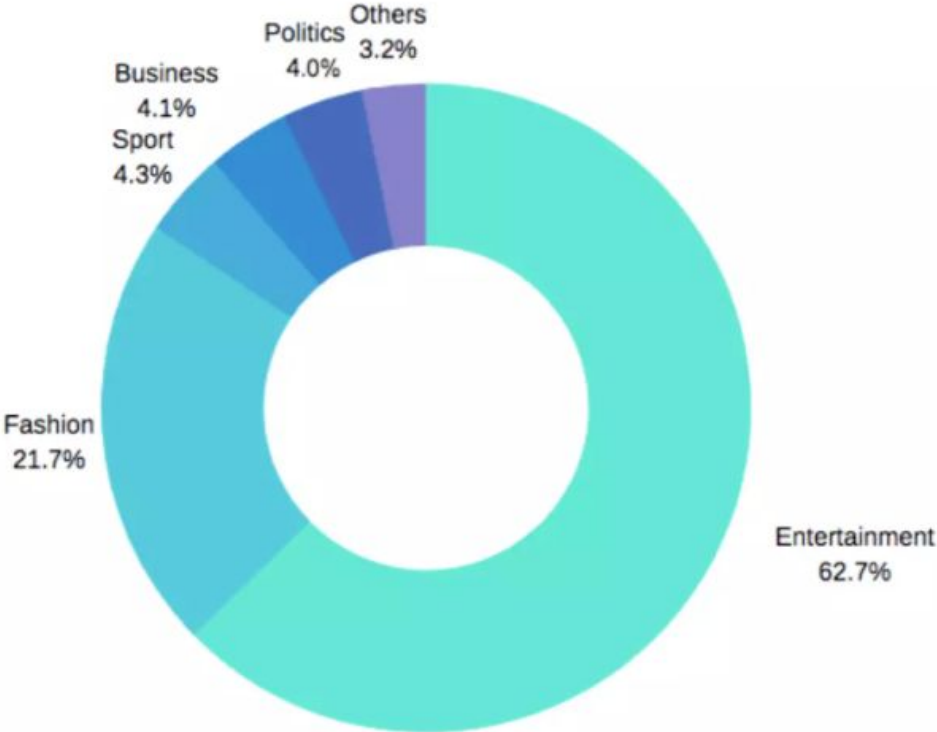
Politics

Cybersecurity

Pornography



Distribution





Should We be Worried?





Deepfakes *Detection*





Microsoft Video Authenticator

Sensity

Deepware





Open Discussion



