

# CASE STUDY: YOUTUBE

---

AI PROFESSIONAL

AKAR GUPTA	S3808546
ANURAG DINESH KARMARKAR	S3829070
CHINMAY PRADEEP KARANGUTKAR	S3827733
SIDDHARTH KUMAR	S3816208

# Table of Contents

---

INTRODUCTION - .....2

BACKGROUND - .....3

CASE STUDY - .....3

AI/DATA SCIENCE TASK DESCRIPTION - .....3

FATE ANALYSIS -.....4

*F: Fairness* - .....4

*A: Accountability* - .....4

*T: Transparency* - .....4

*E: Ethics* - .....5

Legal Counter Provisions - .....5

Desirable Approach - .....6

Conclusion - .....6

Reference - .....7

# INTRODUCTION -

---

*"If You're Not Paying for It, You're the Product" – Anonymous, Internet*

YouTube an online video sharing platform, founded by Steve Chen, Chad Hurley, and Jawed Karim in 2005, is currently the world's biggest video sharing platform and world's second most visited website.

As a pioneer of video sharing and streaming, YouTube's services are available over their website, mobile application, tv application and even via embedding into other websites. The platform contains plethora of content like music videos, educational videos, live streams, vlogs, movie trailers, news, etc, primarily uploaded by various users, while the company provides availability of uploaded content in flexible video quality formats. The registered users can provide feedback about the content after watching by rating, commenting on, and liking it, and could also show support to the content creator by subscribing them on the platform.

This availability of versatile content, an adaptive technological infrastructure capable of serving any user around the globe irrespective of the device they are using, direct engagement with the content creator and while being a free service helps understand why it is the world biggest platform. It is estimated that collectively users watch around 1 billion hours of YouTube videos every day.

While being a free service the main source of revenue for YouTube is through target ads and user data collection. The platform generates, store and sell user behavioural data for user targeted marketing. YouTube being acquired by Google in 2006 for USD 1.65 billion has proven to be the most lucrative subsidiary of Alphabet, the parent company of Google. Its revenue in 2020 was USD 19.8 billion with a big chunk of it coming from Google's AdSense, the targeted ad service, and the rest via pay-to-play subscription-based model operated under the name of YouTube Premium and YouTube Music.

## BACKGROUND -

---

The conceptualization of YouTube came after one of its founders could not easily find clips of the 2004 Indian Ocean Tsunami and other such incidents on the internet. The founders thought it would be a good idea to develop an online platform where anyone with a working internet connection can share a video. This video can then be accessed by everyone on the platform irrespective of their location. During its initial rollout, the website had the capacity to upload 100 megabytes (30 seconds) of footage at a time. Overtime, it has evolved and now can have hours and hours of content uploaded to it.

We all have heard the saying “Picture is worth a thousand words”, well now we live in the world of video communication. Due to the advent in consumer technology, the most popular content consumption formats have defaulted to videos. It allows to for a direct two-way communication where people can have at-length discussion and explore a given topic. Most of the current search engines today tend to favour video over other formats. Hence, a good formatted and titled video is instrumental in improving one’s search engine ranking.

Video is the perfect medium for branding. Generally, people respond very well to visual information and videos conveys the message very effectively. The visual medium has the capability to tap into emotions of the viewers. YouTube is a very good place to publish such videos and extend your outreach. YouTube as a platform is sharable. The videos can spread far and wide very easily.

Educational content can be effectively delivered to students. Concepts which cannot be explained efficiently in the text form can be thoroughly explained by videos. News can be broadcasted live because of live streaming. YouTube is a mobile friendly platform and does not require the user to have a computer. Thus, it is accessible to users while they are travelling. A video sharing platform such as YouTube has many applications that can prove to be very useful if used in the right context.

## CASE STUDY -

---

Recently, YouTube was found to be collecting personal information such as persistent identifiers of children under the age of 13 without the consent of their parents. This data was then used by YouTube’s algorithm to advertise products to these children that fit their preferences. Children, being under-age will be susceptible to targeted and / or false advertising. The concerning aspect to this scenario is that the parents of these children are not aware of this data collection.

## AI/DATA SCIENCE TASK DESCRIPTION -

---

The AI task being considered is a targeted ad system implemented by YouTube that displays targeted ads to user based upon the user’s behavioural data collected by its video recommendation system. The recommendation system suggests videos to the users on the platform after considering the users’ watch history and based on the user’s interaction with the genre of video builds a behavioural profile.

# FATE ANALYSIS -

---

## *F: Fairness -*

When talking about fairness in context of Data Science it refers to an AI/ML system that has the least number of biases so that it does not discriminate based on socio-economic factor such as race, creed, gender, or sexual orientation of a person. Thus, Fairness in Data Science is concerned with actions, processes and consequences which results in the model treating all individuals with equality.

In reference to the ML model under scrutiny as part of the study, even though the ML model did not actively discriminate among individual while providing targeted ads, in the interest of high children on its platform YouTube should have altered the model so that behavioural data of kids were not collected and later used to serve then targeted ads.

## *A: Accountability -*

It refers to the responsibility in AI and Data Science decision making. Presently, there are many AI systems which make decisions to boost the productivity of the system. The decisions taken by the system often are the result of complex mathematical or statistical operations based on the guidelines laid down by the developers. However, the developers cannot always predict all the decisions taken by the AI system. So, the question arises that the if the decision taken by the AI system proves to be a wrong one, who is to blame? This is the accountability that is being referred to in the FATE analysis.

As for accountability of the ML model in question, we firmly believe that YouTube/ Google are shying away from taking full responsibility of the model. To avoid any kind of discovery and investigation into their practices the company opted for out of court settlements in the previous lawsuits.

## *T: Transparency -*

Most of the systems driven by AI are often the products of collective efforts of various underlying algorithms that take appropriate decisions based on the external features. There is no fixed way to determine how the model learns the correct responsive action to any given scenario. This is the result of various small chip-level decisions happening in the system. It is very difficult to log these actions and trace the computations that lead to the output. This is the issue of transparency toward the nature of the algorithms where nobody knows exactly how the algorithms take the decision.

YouTube created a specialized platform for children under the age of 13, called “**YouTube Kids**”. The highlighting features were that children will not be able to log in without the consent of parents and YouTube promised that it will not collect behavioural data of children because it violates the Children Online Privacy Protection Act (COPPA). Allegedly, even after declaring the above YouTube went ahead and collected the behavioural data of children and used it for targeted advertising. This was juxtaposed to what YouTube promised its customers and can be considered as breach of contract.

The way YouTube operates and implemented its ML model the users, and parents about the lawsuit, did not knew what kind of data was being gathered. This highlights a bigger transparency and ethical concern. As a parent, not knowing that a big corporation is storing your kids' data is unscrupulous. Furthermore, since the collected data was used to target children for advertisement, there is no way to know how much information and what kind of information was shared with the third-party advertisers.

## *E: Ethics -*

Ethics are defined as shared values which basically assists the humanity to distinguish right practices from wrong practices. Ethics have its standards which follow up the entire process of data science practices and then according to the standards it has been decided that the following practice is fair or unfair.

As mentioned before the idea of collecting children's data is itself unethical but the bigger concern is that the data was collected without the knowledge of the parents / legal guardians. The parents have a right to decide whether they want to share the data of their kids with the company. The unethical part of this issue is that the parents were not given a choice to opt out of the data collection.

## **LEGAL COUNTER PROVISIONS -**

---

In April of 2018, a coalition of 23 groups (including the CCFC, CDD, as well as Common Sense Media) filed a complaint with the Federal Trade Commission USA(FTC), stating that YouTube has collected information from users under the age of 13 without parental consent, in violation of the COPPA laws. By September of next year i.e., 2019, YouTube was fined \$170 million by the FTC for collecting personal information from minors under the age of 13 without parental consents. It is to be noted that there were legal provisions made by Government after this lawsuit so that they could fight for their people on these issues.

Even after rules added by the FTC another lawsuit has been filed by the State of Massachusetts against YouTube regarding the same issue. The State filed a class action lawsuit against YouTube in violation of the Children's Online Privacy Protection Act. This act stated that any online platform cannot collect data of children under the age of 13. It restricts the corporations from collecting and exploiting the sensitive data of children. This case suggests us that an amount even as big as \$170million is not a big thing for such big corporation and there should be other strict laws in place which we have discussed in next section.

YouTube Kids on paper was a good initiative by the company for segregating the content availability based on age. As all the YouTube content are not children friendly so it was a great move on YouTube's part. In YouTube Kids, the feature where the children need parental consent before viewing any content. This feature was brought back due to federal law which supplied data access control to the parents.

## DESIRABLE APPROACH -

---

The responsibility lies with the tech companies itself, for being transparent and communicate with their users when implementing such targeted machine learning models. These AI/ML technologies store, manipulate and share user's behavioural data to "better" serve the user, but there is still lack of declaring notice for consent of the user, whether to opt in or out of these services.

One of the ways companies can operated is to implement and expand the GDPR principles in most of the ML models. In addition to that companies should clearly outline what kind of data is being collected and for what purpose. This will allow user to make an educated decision whether to opt in or out of a service.

Government bodies should enact new data protection laws which is in accordance with our current internet infrastructure, one of the examples of this is GDPR. Also, provisions should be enacted that can scrutinize these tech giant's operation practices and allows to impose heavy fines. Government can use the method of acquisition where they can hold a certain % of shares or ownership of such companies and can continue doing so until they make their records clear. We want to clarify that this can be implemented in an incremental way for e.g., if an organisation does not comply more than three times then the Government can take ownership of 5% of the corporation's shares.

Federal Government should bring some more methods where it will include enquiry of corporations on the account of how they collect data and where it has been used. This way of enquiries will keep big tech giants under pressure, and it may reduce the privacy issues which are currently present.

## Conclusion -

---

We understand that YouTube has violated data collection and protection laws pertaining to the data of its users. This is a serious breach of the users' trust on behalf of the company. This case highlights the growing concerns pertaining to data collection ethics and transparency to the consumer.

We firmly believe that the "Big Tech" should strive to implement F.A.T.E principle in most of their existing and upcoming machine leaning model. Adoption of these principles will reduce the likelihood of questionable data collection and handling practices being employed further.

However, these actions by the company have not gone unnoticed, as there have been legal actions against YouTube on multiple instances in the past. This report aims to shed some light on these issues and the out of settlements by the company to silence and pacify the accusers.

## REFERENCE -

---

- Variety.com. 2021. *YouTube's Child-Directed COPPA Content Rules for 2020: What to Know - Variety*. [online] Available at: <[https://variety.com/2020/digital/news/ftc-rules-child-directed-content-youtube-1203454167/?sub\\_action=logged\\_in](https://variety.com/2020/digital/news/ftc-rules-child-directed-content-youtube-1203454167/?sub_action=logged_in)> [Accessed 25 April 2021].
- Stempel, J., 2020. *Google faces \$5 billion lawsuit in U.S. for tracking 'private' internet use*. [online] U.S. Available at: <<https://www.reuters.com/article/us-alphabet-google-privacy-lawsuit-idUSKBN23933H>> [Accessed 25 April 2021].
- Boston.cbslocal.com. 2020. *Massachusetts Among Dozens Of States Suing Google Claiming It's A Monopoly*. [online] Available at: <<https://boston.cbslocal.com/2020/12/18/google-lawsuit-monopoly-massachusetts-attorney-general/>> [Accessed 25 April 2021].
- Federal Trade Commission. 2019. *Google and YouTube Will Pay Record \$170 Million for Alleged Violations of Children's Privacy Law*. [online] Available at: <<https://www.ftc.gov/news-events/press-releases/2019/09/google-youtube-will-pay-record-170-million-alleged-violations>> [Accessed 25 April 2021].
- International Comparative Legal Guides International Business Reports. 2020. *Data Protection 2020 / Laws and Regulations / Australia / ICLG*. [online] Available at: <<https://iclg.com/practice-areas/data-protection-laws-and-regulations/australia>> [Accessed 25 April 2021].
- Federal Trade Commission. n.d. *Children's Online Privacy Protection Rule ("COPPA")*. [online] Available at: <<https://www.ftc.gov/enforcement/rules/rulemaking-regulatory-reform-proceedings/childrens-online-privacy-protection-rule>> [Accessed 25 April 2021].
- En.wikipedia.org. n.d. *YouTube - Wikipedia*. [online] Available at: <<https://en.wikipedia.org/wiki/YouTube>> [Accessed 25 April 2021].