Final Project Proposal

Concept

       YT is a data visualization video project about *YouTube* culture. *YouTube* was and still plays an essential part in my Internet culture. It is the digital platform that I've used the most since I can access a digital device. I grew up with the platform by watching various YouTubers at different stages of my life. What caught my curiosity about this platform is its popularity and stable use among digital users. *YouTube* is the second most used social media after *Facebook* and the most visited website after *Google*.[1] *YouTube* did not face a downfall compared to other platforms such as *Facebook* or *X* (i.e., previously known as *Twitter*). Since its creation in 2005, the platform has gained new users, channels, and videos after almost 20 years of activity. However, I have not found an artistic project exploring *YouTube* culture. I even asked the generative AI *ChatGPT* if they could provide names of artists or artworks that used *YouTube* as their subject, and none of them were related to the platform. They only referenced the Internet culture in general. Because it is a platform that was unexplored in an artistic sense, I want to dive deeper into the subject and try to define *YouTube*'s identity. I want to discover how and why this platform resists and adapts to the changes and trends of the digital culture. I want to find its recurrent themes that will answer why this platform is essential to our digital culture.

I want to display popular and niche *YouTube* videos of the past 18 years that are data-generated on a webpage. The videos will come from various categories: beauty, gaming, vlogs, news, music video clips, comedy sketches, tutorials, etc. The page will generate new videos after a delay. The delay will be randomized, so the video time length on the screen will vary. The usual thumbnail icons will display videos instead in these smaller windows. Generated comments will be displayed. If the video does not interest the user, they can generate new videos by clicking the "skip" button. But if they do, the comments will start floating around on the webpage instead of staying in the comment section. Each time the user clicks on this button, more floating comments will be generated, and even generated videos will start floating around on the webpage. The more the button is clicked, the closer the users will be to create an abstract representation of *YouTube*'s identity. The web page design will be like *YouTube*'s video layout. My targeted audience is young adults because they are the main demographic that grew up with *YouTube*. According to *Oberlo*, users aged between 25 and 34 are the age group that used the most

---

[1] "Biggest Social Media Platforms 2023," Statista, accessed September 30, 2023, https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/.; "Global Top Websites by Monthly Visits 2022," Statista, accessed September 30, 2023, https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/.

YouTube (i.e., 20.7%.).[2] The content generated will resonate more with this demographic because the references will be familiar to them; they were probably the users who viewed, shared, liked, or commented on those videos.

This project joins the readings well because, weirdly enough, I did not find any artist who specifically did a project about *YouTube*, even though it had greatly influenced the Internet culture. The artist community has not used and explored such a big video database that went through scandals, content restrictions updates, design changes, and trends. I want to use the concept of apophenia mentioned in Steyerl's text, which is the perception of patterns within random data; it draws connections unrelated to the subject and are displayed only due to perceptual simultaneity. Apophenia will happen if the user presses the "skip" button too many times. Comments and images will blend to create an abstract representation of the platform. It will blend to become a colour pattern, lines, and noise of videos. I want to find if there is a pattern in the type of content made over time, its themes, its trends, or its aesthetics. I also think that *YouTube* is somewhat recognized as dirty data; it was a platform that was not explored and studied much, and there is probably missing information about it. I agree with Onuoha about data reflecting our society and evolving. Collecting *YouTube* videos and comments will indicate the platform's identity and reflect people's values and interests. Moreover, collecting data over 18 years will also show the social-historical context of this period and its trends. It will also show what is missing on this platform.

Technicalities

I want to use video and web development for this project because it represents *YouTube*'s essence, a digital platform where you can upload videos. I do not think other ways of collecting and displaying this data would portray the subject better. Creating a 3D model or statistics graphics would make no sense when the study is about discovering the platform's identity, mostly from videos and the exchange between the content creators and their fanbase. I will use *Node.js* and explore *npm* packages to upload, download, or display YouTube videos and generate comments on the webpage. For example, the *npm* package *node-ytdl-core* allows downloading *YouTube* videos and returns a readable stream. The package will help me to display the *YouTube* videos on the webpage. Some packages require *React.js*, such as *react-youtube* (i.e., a customizable player UI for YouTube videos), so I will probably use it, depending on my needs. I will use *HTML* and *CSS* for the layout and design of my webpage. I want to use these programming methods because I am new to data collection and its way of collecting it through programming, so I want to use something I am more familiar with because I am still learning data collection. They are also the coding languages that will help me to create the webpage layout, design, and interaction properly.

My data will be *YouTube* videos and comments from various channels from 2005 to 2023; the total amount of videos displayed must be determined. I will use pre-existing

---

[2] "YouTube Age Demographics [Updated Aug 2023] | Oberlo," accessed September 30, 2023, https://www.oberlo.com/statistics/youtube-age-demographics.

YouTube datasets that cover the years between 2005 (i.e., *YouTube*'s first year of existence) and 2023 to create a randomly generated list of videos. I have found some spreadsheet datasets on the *Kaggle* platform. One of them is named *YouTube Videos Dataset (3400 videos)*. This dataset contains around 3400 *YouTube* videos divided into four categories: food, travel vlogs, history, and art and music. If popular channels or videos do not appear on those lists, I will create my own dataset about this missing data covering the years 2005 to 2023. Not showing these videos would remove an influential aspect of *YouTube*'s identity. I will find articles that list the most popular channels on *YouTube* that cover each year of its existence and create a dataset based on that. The same goes for the comments. I will use pre-existing datasets from the platform *Kaggle*, such as *MOST LIKED COMMENTS ON YOUTUBE*, a dataset about the most liked comments on *YouTube* from 2005 to 2020. Suppose the pre-existing datasets do not have enough variation in the data collected. In that case, I will create a comments dataset by selecting comments from the videos listed in the existing video datasets. I want to use pre-existing datasets because there is so much content on the platform and so many channels that I do not know it would be nearly impossible to collect all the data myself. Using those datasets will help me to create a database that will cover my intention with this project and help me to learn more about this platform.

My knowledge of algorithms is too low to tell precisely what kind of algorithms will be needed or how they will be used, but I can give a framework. The goal is to loop sequences of generated *YouTube* videos on a webpage. I will find and build, if necessary, *YouTube* datasets. I will use the *YouTube* datasets to select and search several videos on the platform. Once I find them, I will collect their URLs and add them to the chosen *npm* package. Once the package is tested and works, I will try to find a way to generate a new video after a specific delay. If this works, I will try to randomize this delay. If this whole generative algorithm works, I will test if I can generate and display more than one new video on the webpage. If it works, I will start working on the webpage's design. Once the design is done, I will start working on the skip interactive button and generate comments on the webpage. If comments can be generated, I will work on adding more floating comments by clicking on the button. If this interaction works, I will try to add videos to the mix. If every step works successfully, then the project will be complete. I still need to find more resources to create this generative algorithm.

Storyboard
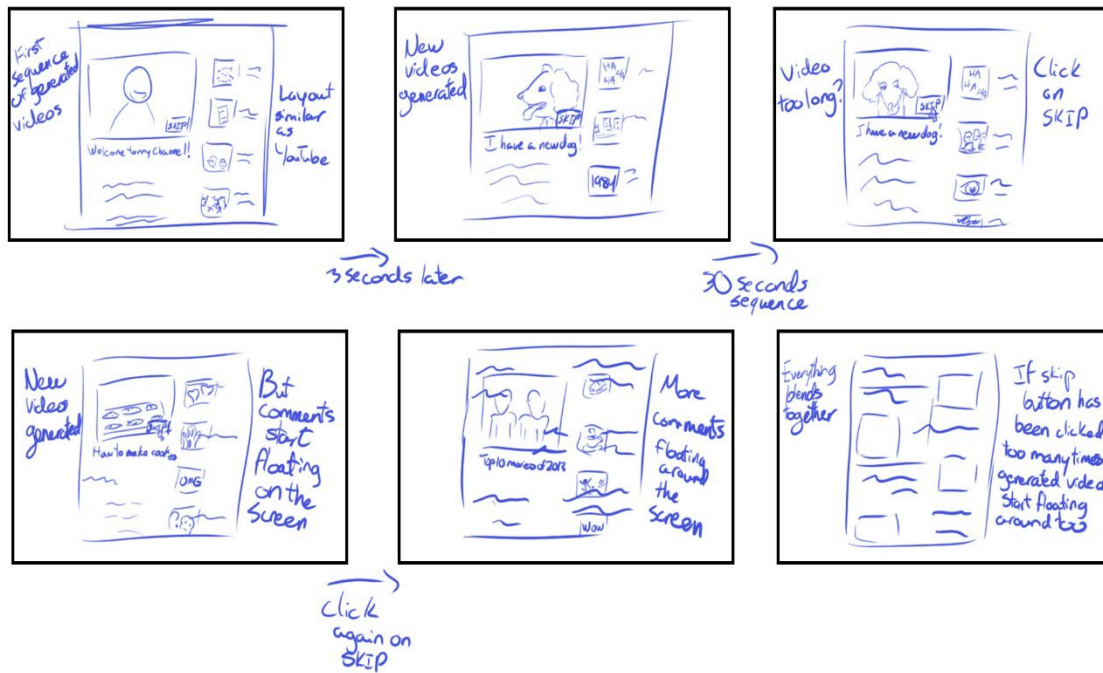
Artistic Inspiration

Three projects inspired me for this creation. The first one is *All the Ways (The Simpsons)* made by Jason Salavon. Jason Salavon is an artist who works around art, information technology, and daily life. He designs his software processes to generate and reconfigure masses of communal material to present new perspectives of the familiar. His projects manipulate the roles of individual elements arranged in diverse visual populations, unfolding unexpected patterns. His work reflects the natural attraction to popular culture by using common references and resource material. His practice uses a lot of amalgamation and unification. Salavon made the *All the Ways* series to explore all *The Simpsons* episodes as data, covering season 1 to season 26. I focused on the video of this series that gathers all the frames of every episode of the 26 seasons into a single episode. The project contains 17.7 million frames in total, including standard and high-definition scenes. The composition also includes algorithmically generated audio. Salavon and his team created software to change the generated videos' parameters. The blending result of this video is something that I would like to accomplish in my project; I would like to generate *YouTube* videos at a similar speed. Unfortunately, the code or details of Salavon's software are unavailable, so I cannot use it as a reference for my dataset. However, his aesthetic inspires me to have a similar result, especially with the amount of content provided on *YouTube*.

Particle is another project with a similar aesthetic to Salavon made by *D-Fuse*. *D-Fuse* is a visual arts company working on multiple media such as installations, film, experimental documentaries, photography, live cinema performances, VR, AR, and architectural projects. Their work addresses social and environmental themes and explores collaborative processes. *Particle* is a processed urban imagery that fluctuates between recognizable landscapes and abstract data-like patterns combined with dense sound textures. It deconstructs visual and audible urban landscapes. The purpose was to fragment the reality of everyday life. It explores processes of abstraction. The data collected is videos and sound from across the globe used for the *D-Fuse* documentary *Endless Cities*. The images and sounds have been broken into fragments and then reconfigured. I like the result of this project; the blurriness and transparency of the sequences give a dreamy effect, and the fragmentation of the audio and video in Particle is something that will also be found out in my project *YT*. I like the distinct layering of the sequences and the balance between longer and shorter sequences. A change of rhythm is also something that I consider implementing in my project. I think it would be an effective way for the audience to discern glimpses of the video platform. Layering text like it was done in *Particle* is what I intend to achieve a similar effect with the comments floating on the webpage once the "skip" button is pressed.

The last project is Mary Flanagan's *pile of secrets*. Mary Flanagan is an artist with a research-based practice investigating and exploiting the seams between technology, play, and human experience. She explores how data, computing practices, errors, glitches, and games reflect human psychology and the limitation of knowledge. Her approach involves both onscreen spaces and physical spaces. *pile of secrets* is a series that displays video game footage on LCD screens and is exhibited in festivals, art centres and museums. Flanagan collected video game footage from 1980 to 2011 to answer the question: "What makes a game a game?" Flanagan gathered common elements in video games into her database. The recurring themes that came across the footage are Jump, Ascend, Dying, Key, Treasure, Corridor, Shooting, and Doors. The project manifests a structural analysis of games in a creative form, exploring elements essential to the experience of digital games. The videos are displayed in an infinite loop, alluding to repetitive elements and repeat gameplays that games offer as a typical part of their structure. My research about *YouTube* is similar to Flanagan's project. I want to discover what makes the platform *YouTube*. I also want to create a loop with my database; even if the user has viewed the whole dataset, it will repeat itself. The looping effect also shows the rabbit hole of *YouTube*. The platform keeps suggesting new videos to its users after each video ends and has so many videos uploaded that it would take a while before watching everything on the platform.

Bibliography

"D-Fuse - About." Accessed September 30, 2023. http://www.dfuse.com/about.html.

"Jason Salavon." Accessed September 30, 2023. http://salavon.com/page/biography/.

"Jason Salavon | All the Ways (The Simpsons)." Accessed September 30, 2023.
    http://salavon.com/work/all-the-ways-video/.

JASON SALAVON - ALL THE WAYS Feb 25 - Apr 9, 2016 @ Mark Moore Gallery,
    2016. https://www.youtube.com/watch?v=ja1c1HbQdKQ.

Mary Flanagan. "About - Mary Flanagan -." Accessed September 30, 2023.
    https://maryflanagan.com/about/.

"MOST LIKED COMMENTS ON YOUTUBE." Accessed September 30, 2023.
    https://www.kaggle.com/datasets/nipunarora8/most-liked-comments-on-youtube.

npm. "React-Youtube," November 22, 2022. https://www.npmjs.com/package/react-
    youtube.

npm. "Ytdl-Core," July 14, 2023. https://www.npmjs.com/package/ytdl-core.

"Particle." Accessed September 30, 2023. http://www.dfuse.com/particle.html.

Statista. "Biggest Social Media Platforms 2023." Accessed September 30, 2023.
    https://www.statista.com/statistics/272014/global-social-networks-ranked-by-
    number-of-users/.

Statista. "Global Top Websites by Monthly Visits 2022." Accessed September 30, 2023.
    https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/.

user, mf default. "[Pile of Secrets]." *Mary Flanagan* (blog), July 3, 2011.
    https://maryflanagan.com/pile-of-secrets/.

"YouTube Age Demographics [Updated Aug 2023] | Oberlo." Accessed September 30,
    2023. https://www.oberlo.com/statistics/youtube-age-demographics.

"Youtube Videos Dataset (~3400 Videos)." Accessed September 30, 2023.
    https://www.kaggle.com/datasets/rajatrc1705/youtube-videos-dataset.