Fiona Liu

PREFACE

In recent years, there have been an increase in sites, apps, and general platforms that aim to provide awareness and higher transparency surrounding products or processes that consumers of the digital world use daily. Examples of such platforms can include WebsiteCarbon, which calculates a website's carbon footprint, SurvivalOfTheBestFit, which immerses users in an educational game to reveal algorithmic bias, and many more. I envisioned my ethics primer to function in a similar manner, as an accessible means that does not shame technology or the power structures surrounding it, nor feeds a certain position to users, but instead educates and challenges users with the transparency that is often clouded from them.

Cautions

For now, I am setting this up in the form of an app and/or site, but want to caution that it is a mere framework of ideas, shared in this vessel because of its access. Problem solving in the face of digital technology can often be whittled down to the easiest solution being an app, but that is by no means true, and was actually one of the many hypocritical concerns and critiques I had, and would encourage you to continue to have as you read through this primer, about creating an ethics app that evaluates other apps. Some questions I considered included: How does evaluation of ethics truly work in a digestible manner? Inclinations often point towards a score, but the act of scoring itself is problematic in its devalue of the app; however, to solely give information and expect users to rely on their individual moral strength can be either pressure-inducing or ineffective. What is the app and its creator's accountability if the proposed ethics information is extracted into a harmful context? On a greater scale, how might ethics judgements be biased by ownership over the platform? Who is writing these ethics evaluations in the first place, especially when the number of apps to evaluate increases? How susceptible can the reliability of ethics judgements be to being bought out by advertisers, or acquisitions by wealthy investors - how might introducing "more ethical" app alternatives to apps being evaluated contribute to that jeopardization?

How It Will Work

Currently, the proposed platform acts as an awareness service, and attempts to avoid recommending or deterring the evaluated app, instead prompting users to weigh their costs and benefits to see how this app fits into lives in a personal, interpersonal and global context. Users should have the opportunity to weigh values that are important to them and receive information that will prioritize those values while continuing to be exposed to the scope of considerations; in an effort to mitigate the possibility of monopolized bias from the app conceptualizer, it will also try to solicit verified user feedback.

The most critical component of ethics evaluation is the criteria on which it is evaluated upon, and the following section will demonstrate the possible considerations that would be brought to a user's attention if Snapchat were used as a case study.

FRAMEWORK OF CRITERIA

Snapchat is described to be a multimedia instant messaging application, that's primary method of communication is through timed photos, or "snaps", and messages.

ATTENTION AND ADDICTION

One of the most pressing concerns with digital technology is the time spent on apps, and how that translates into attention and addiction. The ethical dilemmas that arise diverge in a personal manner - autonomy over individual attention - and in a corporate manner - monopolization of user proneness to addiction and influence. Those can be categorized as such:

Means	**Our ends**	**Other ends**
Reflective	Engagement/Justification	Persuasion
Non-reflective	Nudging	Manipulation

Snapchat's most relevant features to this concern involve:

- Streaks: A counter of the consecutive days that two users snap each other. This feature triggers an inclination to continuously return to the app on a daily basis, if not more, to uphold the streak.
- Scores: A running tally of activity on the app (Snaps sent, Snaps received, users added, stories sent, messages sent, etc.). Higher SnapScores indicate more frequent usages, and when placed in conjunction with a user's Snap Username, and consequential identity, can encourage comparisons in popularity and general definition of self-worth in a social context.
- Ad Presentation: Ads are generated in between Stories (snaps viewable to everyone, or at least all the user's friends), with time limits that don't permit skipping. This has less to do with personal addiction in attention as it does with influence in attention; it's not about how much of the attention being taken up, but what is taking up the attention in the first place, and whether users have the opportunity to even have a say in that.

Attention can often be difficult to conceptualize because it is both conscious, something from will, and subconscious, something that can be subject to change by outside influences; both facets are necessary for personal and social perception, but what may help users is considering the substance that is being consumed to impact attention. The addiction that can be associated with Snapchat may facilitate regular contact that would not otherwise be upheld, but the less developed quality of those bursts of contact may also lead to deficits in meaningful attention.

PRIVACY

Privacy in a data-driven digital sphere can be extremely difficult to parametrize; on one hand, users are entitled to an inherent right to privacy, which is essentially a right to not be infringed upon, but on the other hand, the supposedly democratic nature of the digital sphere can make it difficult to claim ownership over information. Privacy is most fairly defined, then, as the relationship of conscious intention to divulgence between all parties, contingent upon the situation, space, and time.

Snapchat's most relevant features to this concern involve:

- Locations: Users can see where their friends are using SnapMap, and the last location shown is the last time a user has had Snapchat open. GhostMode is available to hide your location.
- Notifications: One of Snapchat's most notable features is the openness in user activity. Not only are snaps timed, but chat and photo screenshotting, saving, and deleting are all documented for all participating users to be notified.
- Astrology: Bitmoji is a customizable avatar service that joined Snapchat in 2018, and is accessed through third-party services. They introduced an astrological feature in 2020, which provides users with a detailed horoscopic profile spanning ten planets, as long as they enter their date of birth and exact time and location. Because it was meant to engage users, and test compatibility between them, this is also an Attention and Addiction concern as well.
- My Eyes Only: Snapchat has Memories, which are photos and videos taken on Snapchat and stored on Snapchat, as well as access to users' Camera Rolls, and My Eyes Only, which is essentially a locked version of Memories.
- Filters and lenses: Snapchat allows creation and usage of augmented reality, often used to distort faces or your environment.

Snapchat's privacy is both user- and corporately-dependent. Snapchat attempts to be more intimate by limiting who and the duration of what can be seen; though this may speak to an addiction pattern, it also, in some ways, protects its users. Users aren't consistently publicly sharing their location, and the ephemeral nature of the messages themselves discourages receipts - this has, however, encouraged sexting amongst young teenagers, though, which is another question relating to the privacy of minors, specifically; if their confidence is breached, whether by someone screenshotting their message, Snap, or Story, they are notified, which can be seen as an effort of mitigation. These are instances of user-dependent privacy in the regard that Snapchat has these implementations, and they can be optimized for privacy from other users as long as used responsibly; protecting one's own location, being alert to screenshots from other users, etc. However, corporately-dependent privacy is what the user can't see and can't control. Users may turn on GhostMode to ensure that their privacy can't be compromised by other users, but there is little known whether Snapchat holds records of user location, if messages between users themselves are actually ephemeral, whether Snapchat has access to

one's My Eyes Only, or what Snapchat, or the third-party service that mitigates Bitmoji-Snapchat astrology, does with users' birth data; for unopened Snaps, they're not deleted from Snapchat servers for 30 days, leaving them susceptible to being hacked, in the case a user loses their phone. It was also reported that Snapchat's filters and lenses spent years scanning user faces without notifying users that it was storing that data. In this sense, users may be acting responsibly to ensure their own privacy amongst the user-base, but have little to no control over what is documented otherwise, and should be vigilant of that data and surveillance.

ALGORITHMIC BIAS

Algorithmic systems are predefined systems purposed to obtain output from a set of inputs, and how bias can seep into this is through the inputs of both users and developers. Algorithmic bias rears its head in multiple forms, such as the types of advertisements that are being presented to a specific demographic, the ease at which certain features may incite facial recognition, etc. This bias can pose a threat to justice both in the ways we value each other as human beings, and what assets we are willing to distribute to each other.

Snapchat's most relevant features to this concern involve:

- Best Friends List: The people a user snaps most will appear in a Best Friends List of 8 users; users have no control over this, and can only mitigate by sending more snaps to someone to try and get them higher on the list.
- Spotlight: Similar to Tik Tok, Facebook and Instagram Reels, Youtube Shorts, Snapchat has Spotlight, which is their section dedicated to short-form video content that is user-generated and then algorithmically user-curated.
- Discover: The Discover page is an amalgamation of ads, Stories from public users, article tidbits from sources such as DailyMail, People, Vogue, Now This, CNN, The Washington Post, as well as snap series like Satisfying Beauty and Snap Originals.
- Filters and lenses
- Scores

Snapchat is different from a lot of other social media platforms such as Instagram, Facebook, and Reddit, because there is no centralized feed. The algorithm still works based on users' Discover and Spotlight Subscriptions, and recommends short stories relevant to the ones already viewed, which Snapchat reportedly tries to diversify with more inclusive content about the LGBTQ+ community and underrepresented minorities.

The main contender for algorithmic bias in Snapchat is the visual component, specifically in camera-inclusivity. Light calibration in cameras is by default built on white skin, and apertures are still not wide enough today to capture all skin tones; Snapchat is reportedly improving their photo processing software to reflect the physical identities of their users. The machine learning and augmented reality used for filters is also in question, as face-tracking

datasets are often trained mainly on white users, leaving all other demographics at a disadvantage, and Snapchat has been trying to optimize their variance in their algorithm audits.

The technical constraints are being considered, but the content of the facial filters and lenses themselves are still in need of evaluation. Oftentimes, Snapchat recommends “facial-enhancing” filters; not only has this increased plastic surgery rates, but also perpetuates a parallel in impressionable audiences that enhancement is equivalent to larger eyes, thinner noses, smoother and lighter skin - eurocentric features. There was a Bob Marley filter that essentially utilized blackface, and another released filter for Juneteenth that prompted users to “smile to break the chains”. The recommendation system of filters is still unclear, but it is a consideration in algorithmic bias when evaluating how these factors influence user’s self worth, but also what values they’re promoting.

SPREAD OF INFORMATION

Being in the Information Age means we are naturally prone to a spread of information whether it is factual, false, or a mere opinion. There is a new pressure to keep facts afloat because of the sheer volume of information that can drown them out, but spread of information is not just about what the information is and whether it’s accurate or not - it’s also largely about the communities and vessels in which that information is spread, and how that may transform our social habits and abilities to empathize with differing opinions. Common concerns related to this could include echo chambers and filter bubbles.

Snapchat’s most relevant features to this concern involve:

- 👻 Discover
- 👻 Filters

As mentioned earlier, Snapchat is different from a lot of other social media platforms because it was intended as a communication method for close friends, where you can only really talk to people who have been mutually added, and not for broadcasting information. That has generally held true; misinformation is most directly spread to the limits of one user’s social circle, and it’s much more difficult to congregate with like-minded people, or even to mass follow popular figures; this can work to maintain existing relationships without a need for reaffirmation from a feed, but can also decrease opportunity for exposure.

However, that doesn’t mean Snapchat doesn’t facilitate app-wide interactions at all. Users can still post to the public, and with that crowd-sourced content, it can be difficult to monitor validity in claims, especially if Snapchat isn’t vetting public user content or the content of the sources they are partnered with.

Moreover, filters are real estate for propaganda as well. Not only has Snapchat been linked to deep fakes, but instances like 2016 presidential candidate Donald Trump purchasing a “Crooked Hillary” geo-filter viewed 80 million times also have influence on Snapchat’s role within ideological spreading.

POWER AND CONTROL

Apps are just one part of a larger corporation, and within the business of corporations, there are naturally stakeholders and investors. While users are a critical part of the stakeholder base, there are also wealthy investors who contribute to that stakeholder population as well. That wealth or power can then have large implications in content moderation.

Snapchat's most relevant features to this concern involve:

👻 Discover

Snapchat selectively chooses the programs that it partners with for its "news sources" within Discover. This does somewhat weed out notoriously gimmicky fake news partners, but it also skews the range of entertainment and information offered in favor of political and personal values of the app's investors.

WORK DISTRIBUTION

Work distribution is concerned with how a product or a service can affect the workforce. For many physical and laborious jobs, introduction of technology can either overtake jobs or enhance jobs, for a consequence of stripping a livelihood or preventing an injury and speeding up production. The integration of digital technologies can be more difficult to measure, as artificial intelligence has yet to reliably mimic human thinking patterns and morals, and society has yet to decide whether it even wants to implement that.

Aside from potentially increasing visibility of music or filter creators whose services get featured on the app or partner companies or influences who have a spot on the Discover page, Snapchat isn't as integrated with the workforce. Unlike Instagram's Shop section or Facebook Marketplace that give users estate to grow their monetizable platform, Snapchat's structure doesn't really allow for that.

ACCOUNTABILITY

Responsibility is often related to the task at hand, and accountability is contingent upon the outcome, but accountability is also a greater reflection on impact and future justifiability as well, and that starts before the outcome is even released.

As mentioned, Snapchat has come out with supposed ways to diverge fake news and implement more ethical engineering, but there is yet to be transparency about how those algorithms may be implemented or what councils will overlook that.

PROTOTYPE

In this particular screen, I've represented the way in which ethics might be presented to a user. The grid is based on values in personal (top; such as how it pertains to an individual user's safety and physical and mental health), interpersonal (lower left; such as how the app may impact close social relationships), and global (lower right; such as how the app feeds into a greater systemic issue).

The light teal in the background indicates user prioritization; this user, in particular, is concerned primarily with effects on personal and global ethics. The more saturated teal is representative of the moral center. The further a spot is from the center, the more it deviates from the user's moral values.

Spots are as follows:

Orange is attention and addiction,

purple is privacy,

fuschia is algorithmic bias,

yellow is spread of information

red is power and control,

brown is work distribution,

and green is accountability.

Once clicked, the corresponding consideration from the previous Framework section will be conveyed.