**Assignment: Investigating and Analysing Ethical Risks of a Digital Engagement application**

I’m going to be talking about twitter as it’s a digital platform with huge reach and has had big changes made to its ethics in recent years after its sale to Elon musk.

**Introduction:**

Twitter is widely recognized as a leading social media platform worldwide, playing a pivotal role in facilitating instantaneous communication, sharing of information, and public discussions. Through its concise and influential format of brief messages, referred to as tweets, Twitter has transformed the dynamics of online interaction for individuals, groups, and governmental entities alike.

In this assignment, I choose Twitter as the focal point for my assignment, due to its enormity and significance in digital engagement. With over 330 million monthly active users worldwide, Twitter has become a vital tool for sharing news, expressing opinions, creating communities, and starting a public dialogue on a lot of topics ranging from politics and social issues to entertainment and sports.

Beyond its user base, Twitter's influence extends to shaping public discourse, a way for pushing your own agenda, and home to a lot of citizen journalism. Also, Twitter serves as a breeding ground for AI-driven technologies, using machine learning algorithms from content the user has watched to create personalized recommendations, sentiment analysis, and trend detection.

Understanding Twitter's role in digital engagement necessitates a comprehensive examination of its impact on society, its utilization of AI technologies, and the ethical considerations inherent in its operations. By delving into these aspects, in this assignment I aim to show the ethical Risks twitter has by being a social media which has in recent times made a point on the need for free speech and other risks which I shall explain later. Also, I will talk about Twitter as a real-world application and its implications for ethical AI-driven digital engagement.

**Description of Twitter**

Twitter is a prominent American communications company headquartered in San Francisco, California, renowned for its microblogging and social networking platform of the same name. The platform enables users to post and interact with messages, known as "tweets," which are limited to 280 characters. These tweets can include text, images, videos, and links, fostering real-time communication, information sharing, and social interaction.

Users can follow other accounts to receive their tweets on their timeline, engage with tweets through likes, retweets, and replies, and explore trending topics through hashtags. Twitter serves as a digital hub for a diverse range of content, including news updates, personal anecdotes, promotional material, and cultural commentary.

**Utilization of AI-driven Technologies**

Twitter leverages artificial intelligence (AI) across various facets of its platform to enhance user experience, optimize content delivery, and analyse user engagement. Some key areas where AI-driven technologies are utilized include:

**Content Recommendation:** Twitter employs AI algorithms to recommend relevant content to users based on their interests, engagement history, and trending topics. These recommendations help users discover new accounts to follow and engage with, thereby enriching their Twitter experience. For example, Twitter's recommendation system suggests tweets, accounts, and topics tailored to individual user preferences, fostering personalized engagement.

**Trend Analysis:** AI algorithms analyse vast amounts of tweet data to identify emerging trends and popular topics in real-time. Twitter's Trending Topics feature highlights the most-discussed subjects on the platform, allowing users to stay informed about current events and conversations. AI-powered trend analysis enables Twitter to adapt its content delivery and prioritize timely and relevant information for users.

**Sentiment Analysis:** Twitter utilizes AI-powered sentiment analysis to assess the emotional tone and sentiment expressed in tweets. By analysing language patterns and contextual cues, AI algorithms can categorize tweets as positive, negative, or neutral, providing insights into public opinion and sentiment trends. Sentiment analysis helps users gauge public perception, track brand sentiment, and identify emerging issues or crises.

**Role of AI in Enhancing Digital Engagement:**

AI plays a pivotal role in enhancing the digital engagement experience on Twitter by facilitating personalized content discovery, real-time trend monitoring, and sentiment analysis. Here are real-world examples showcasing how AI enhances digital engagement on Twitter:

**Personalized Content Recommendations:** When users sign up for Twitter or follow new accounts, the platform's AI algorithms analyse their interests, engagement history, and interactions to suggest relevant tweets, accounts, and topics tailored to their preferences. For example, if a user frequently engages with tweets about cooking and follows several food-related accounts, Twitter's recommendation system might suggest tweets from chefs, cooking enthusiasts, or trending recipes.

**Real-time Trend Identification:** Twitter’s Trending Topics feature showcases the most-discussed subjects on the platform in real-time. AI algorithms continuously analyse tweet data to identify emerging trends and popular conversations. For instance, during major events like sports games or award shows, Twitter's AI can quickly identify trending topics related to those events, allowing users to stay informed and engage with relevant content as it unfolds.

**Brand Sentiment Analysis:** Many brands use Twitter's AI-powered sentiment analysis tools to monitor public perception and track sentiment trends regarding their products or services. For example, a company might track mentions of their brand on Twitter and analyse the sentiment expressed in those tweets (positive, negative, or neutral). By doing so, the brand can gain insights into customer satisfaction, identify areas for improvement, and respond promptly to customer feedback or concerns. Additionally, sentiment analysis can help brands gauge the effectiveness of their marketing campaigns and adjust their strategies accordingly.

In summary, Twitter harnesses the power of AI-driven technologies to deliver personalized content recommendations, monitor real-time trends, and analyse sentiment, thereby enhancing the digital engagement experience for users and brands alike. Through AI-enabled features and functionalities, Twitter continues to innovate and evolve its platform to meet the dynamic needs of its global user base.

**Stakeholders**

**Users:**

**Creators:** Users are not only consumers but also creators of content on Twitter. They contribute by sharing tweets, posting updates, sharing multimedia content like photos and videos, and engaging in conversations.

**Engagers:** Users engage with content through actions such as likes, retweets, replies, and mentions. These interactions shape the visibility and spread of tweets, influencing trends and discussions.

**Community Builders:** Users also play a role in building communities and fostering connections on Twitter. They participate in discussions, form networks based on shared interests, and collaborate on various initiatives.

**Feedback Providers:** Users provide feedback to Twitter through various channels, including direct messages, tweets, and surveys. This feedback helps Twitter understand user preferences, identify issues, and improve the platform.

**Developers:**

**Tool and App Creators:** Developers create third-party tools, applications, and integrations that enhance the Twitter experience. These can include analytics tools, scheduling apps, social media management platforms, and more.

**API Users:** Developers utilize Twitter's API (Application Programming Interface) to access platform data and functionality for their apps and services. They build upon Twitter's infrastructure to create innovative solutions.

**Contributors to Platform Innovation:** Developers contribute to platform innovation by experimenting with new features, creating plugins, and proposing enhancements. Their contributions can influence Twitter's product roadmap and user experience.

**Advertisers:**

**Promoters:** Advertisers use Twitter's advertising platform to promote their products, services, events, or brand messaging. They create targeted ad campaigns to reach specific audience segments and achieve marketing objectives.

**Revenue Contributors:** Advertisers contribute to Twitter's revenue stream through advertising spend. Their investments sustain the platform financially and support ongoing development and operations.

**Content Sponsors:** Advertisers may also sponsor content or collaborate with influencers and content creators to reach wider audiences. Sponsored content adds diversity to the content ecosystem and can drive engagement.

**Regulators:**

**Legal Compliance Enforcers:** Regulators ensure that Twitter complies with relevant laws, regulations, and industry standards governing online platforms. They monitor content moderation practices, privacy policies, data protection measures, and consumer rights.

**Policy Makers:** Regulators develop policies and guidelines that govern social media platforms' operations, including Twitter. They may introduce new regulations or updates to address emerging challenges such as misinformation, hate speech, and user safety although with free speech being a huge point made by Elon Musk when he bought twitter some policies have aren’t enforced as much as other.

**Oversight Bodies:** Regulators establish oversight mechanisms and enforcement agencies to monitor Twitter's activities and address complaints or violations. They investigate issues related to user rights, data privacy, and platform integrity.

By understanding the roles of these stakeholders, Twitter can better engage with them, address their needs and concerns, and foster a collaborative and sustainable ecosystem. Each stakeholder group contributes to Twitter's success and influences its development, making their engagement and collaboration crucial for the platform's evolution.

**Identification of Ethical Risks**

The rise of social media platforms like Twitter has revolutionized global communication, providing a forum for individuals to express themselves and engage with others. However, this newfound freedom of expression comes with ethical responsibilities and challenges. The ethical risks associated with Twitter's censorship policies, labour practices, fair operating procedures, consumer issues, and community involvement and development.

Twitter has ushered in an era of unprecedented connectivity and communication, facilitating global interactions and the exchange of ideas on an unprecedented scale. However, the proliferation of free speech on these platforms has brought to light a myriad of ethical challenges that demand attention and consideration. As individuals exercise their right to express themselves freely, the consequences of unchecked speech become increasingly apparent, raising concerns about the spread of misinformation, hate speech, and social discord. Elon Musk, as a prominent figure in the tech industry and owner of the platform Twitter, plays a significant role in shaping the discourse and policies surrounding free speech online. His hands off approach to free speech has enabled the expansion of divisive rhetoric and harmful content, highlighting the need for responsible platform governance and ethical oversight. The ethical risks associated with Musk's approach to free speech on Twitter and the implications of his decision to fire the ethical AI team are huge while he is allowing a basic human right of free speech he is also opening opportunities for hurtful/hateful speech and complete misinformation to be passed around freely and as I said in the beginning of my essay twitter has over 330 million monthly users so it has a global impact on society this all sheds light on the complex interplay between freedom of expression, ethical responsibility, and platform governance in the digital age.

**Human Rights Implications:** Twitter's censorship policies raise concerns about potential infringements on basic human rights, particularly the freedom of speech and expression. In regions where censorship is enforced, users may face limitations on their ability to participate in public discourse, impacting individual users and activists striving for social change. This poses a high-severity risk, with a moderate likelihood of occurrence due to Twitter's global reach and potential government intervention.

**Labour Practices and Ethical Oversight:** The recent termination of Twitter's ML Ethics, Transparency, and Accountability team raises concerns about compromised ethical oversight and labour practices within the company. The elimination of key teams focused on ethics and transparency heightens the risk of biased algorithm development and unethical behaviour. This poses a risk, with a high likelihood given the complexity of content moderation and the pressure to balance competing interests.

**Fair Operating Procedures:** Twitter's censorship policies may result in inconsistencies in content moderation and unfair treatment of users based on geographical location or political influence, undermining trust in the platform. This poses a risk, with a moderate likelihood of occurrence due to varying legal requirements across borders and evolving content moderation practices.

**Consumer Issues and User Satisfaction**: Before Twitters sale restrictions on content to Twitter's censorship policies may have led to dissatisfaction among users who value freedom of expression and unrestricted access to information with a high likelihood given the complex differing landscapes and competing interests that social media platforms navigate. This is where freedom of speech has equalled the field for a lot of users in countries that allow it. Earlier I talked about the risks of freedom of speech, but this is one of the positives that we have gained when using twitter.

**Community Involvement and Development:** Twitter's new censorship policies have the potential to help community involvement and development by giving diverse voices and unlimiting the exchange of ideas. This poses a high risk, with a high likelihood of spreading harmful content and misinformation. Given the challenge of maintaining an open platform while addressing concerns about this sort of content Twitter must learn to balance the good and the bad

In conclusion, Twitter's censorship policies present significant ethical concerns across various social responsibility categories. From potential violations of basic human rights to compromised ethical oversight and consumer dissatisfaction, the ethical risks associated with Twitter's operations are complex. Addressing these risks requires a balanced approach that respects freedom of expression while promoting transparency, accountability, and inclusivity in the social media landscape. By addressing these challenges, stakeholders can work towards creating a more ethical and responsible online environment.

**Discussion of Mitigations Measures for Risk**

The termination of Twitter's ML Ethics, Transparency, and Accountability team has raised significant concerns about compromised labour practices and ethical oversight within the company. To address this risk, Twitter can implement several mitigation measures aimed at creating better ethical oversight and transparency.

One mitigation measure is the establishment of an independent ethics board composed of external experts, industry professionals, and representatives from civil society organizations. This board would be responsible for reviewing and evaluating Twitter's algorithmic decision-making processes, ensuring that the ethical guidelines and principles are upheld. By involving external stakeholders, such as academics, ethicists, and human rights advocates, Twitter can benefit from diverse perspectives and expertise in addressing ethical challenges. This measure would reduce the likelihood of biased algorithm behaviour by providing an independent body to oversee and review Twitter's practices.

Another mitigation measure is the implementation of regular audits and assessments of Twitter's algorithmic systems by independent third-party organizations and internal teams dedicated to ethical oversight. These audits would identify and mitigate potential biases, discriminatory practices, or ethical lapses in Twitter's algorithmic decision-making processes. The internal audits would fix parts of the algorithmic systems that have biases and discriminatory practices, but internal audits would have their own biases as these employees would work for twitter. The external audits could look at the working of twitter as a company fixing the unseen biases which have built up in the company while also looking at the algorithm to ensure the internal fixes don’t have new unseen biases. By proactively monitoring and evaluating these processes, Twitter can identify and address ethical concerns before they escalate into significant issues.

Twitter can prioritize employee training and education on ethical considerations and best practices in algorithmic decision-making. Twitter can cultivate a culture of ethical conduct and responsibility within the organization to develop comprehensive training programs tailored to the unique challenges of content moderation and algorithmic decision-making. By ensuring employees have the necessary knowledge and skills to identify and address ethical challenges. This measure would reduce the likelihood of unethical behaviour and biased algorithm.

Overall, these mitigation measures aim to reduce the likelihood and impact of the compromised ethics which gave been left compromised by allowing free speech, figuring out which tweets or posts go too far, is a job twitter must do and fast as leaving such problematic tweets up even for a few moments can have substantial affect on a person or group.

**References**

Manku, C. (2021, June 21). How Twitter Is Embracing Artificial Intelligence. LinkedIn. https://www.linkedin.com/pulse/how-twitter-embracing-artificial-intelligence-chetna-manku

IFTTT. (n.d.). AI-Generated Tweets. IFTTT. <https://ifttt.com/explore/ai-generated-tweets#:~:text=AI%2Dgenerated%20tweets%20are%20tweets,patterns%20found%20in%20the%20data>.

Emerj. (n.d.). Artificial Intelligence at Twitter. Emerj. https://emerj.com/ai-sector-overviews/artificial-intelligence-at-twitter/

Wikipedia. (n.d.). Twitter. Wikipedia. https://en.wikipedia.org/wiki/Twitter

All Tech is Human. (n.d.). Elon Musk's Impact on Twitter's Stakeholders: A Candidate's Perspective. All Tech is Human. https://alltechishuman.org/community-voices/blog-post-elon-musks-impact-on-twitters-stakeholders-a-candidates-perspective

Solis, B. (2016, June 22). The Twitter Conflict: Twitter's New Algorithm and the Battle Between Shareholders and Stakeholders. Medium. https://medium.com/@briansolis/the-twitter-conflict-twitter-s-new-algorithm-and-the-battle-between-shareholders-and-stakeholders-b9b400fbd0

Hall, B., & Tait, J. (2022). Twitter's Ethical AI Team: Response to Ethical and Governance Risks. MDPI. https://www.mdpi.com/2071-1050/14/24/16769

LaFrance, A. (2019, May 2). Inside Twitter's Attempt to Automate Away Abuse. Wired. https://www.wired.com/story/twitter-ethical-ai-team/

Park, M., McDonald, D. W., Cha, M., & Waber, B. N. (2014). Activity Life Cycles of Mobile Applications. Journal of Medical Internet Research, 16(12), e290. https://www.jmir.org/2014/12/e290/

(2020, January 16). The Business Ethics of Elon Musk, Tesla, Twitter, and the Tech Industry. Harvard Law School. https://hls.harvard.edu/today/the-business-ethics-of-elon-musk-tesla-twitter-and-the-tech-industry/

Sherry, B. (2022, January 6). Elon Musk's Twitter Firings: What the Ethical AI Team Told Us About Artificial Intelligence. Inc. https://www.inc.com/ben-sherry/elon-musk-twitter-firings-ethical-ai-artificial-intelligence.html

Edubirdie. (n.d.). Social and Ethical Implications of Twitter Censorship. Edubirdie. https://edubirdie.com/examples/social-and-ethical-implications-of-twitter-censorship/

Chatgpt was used to help format the better organise the essay.