

**School of Computer Science and Statistics**

**Individual Assessment Submission Form**

|  |  |
| --- | --- |
| **Course Title** | TEP |
| **Module Title** | TEP 00062 |
| **Lecturer(s)** | Prof Dave Lewis, Prof Vincent Wade |
| **Assessment Title** | Individual Assignment on Ethical Digital Engagement |
| **Date Submitted** | 31/03/34 |
| **Word Count** | 3300 |
| **Student Number** | 21365366 |
| **Student Name** | Luke O’Rourke |

**Assignment Submission Declaration:**

I have read and understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: <http://www.tcd.ie/calendar>

I have also completed the Online Tutorial on avoiding plagiarism ‘Ready, Steady, Write’, located at <http://tcd-ie.libguides.com/plagiarism/ready-steady-write>

I declare that this assignment, together with any supporting artefact is offered for assessment as our original and unaided work, except in so far as any advice and/or assistance from any other named person in preparing it and any reference material used are duly and appropriately acknowledged.

Any use of Chatbots/Generative AI tools in researching the materials research for this report is fully described at the end of the report. I confirm that no text produced by such tools has been directly used in the report.

Signed ………Luke O’Rourke……………………………… Date …………31/03/24…………………

Description of use of Chatbots/Generative AI tools:

1. Prompted ChatGPT with “Summarize the principles of social responsibility in the attached pdf”, in order to complete the “identification of ethical risk” section with a better understanding of the concepts.
2. Prompted ChatGPT with “Summarize the risks categorized by Latzer et al. (2016) in the attached pdf”, for the same reason as Prompt 1.
3. Prompted ChatGPT with “Summarize the findings of the attached study”, attached was ‘Mitigating the Privacy Paradox Through Higher Privacy Literacy? Insights from a Lab Experiment Based on Facebook Data ‘ given in the reference section, as I thought the findings were difficult to understand, aided inspiration for the “Discussion of Mitigations Measures for Misuse of personal data” section, but the text is still my own.
4. Microsoft word editor was used to improve sentence structure.

………………………………………………………………

**Table Of Contents**

Introduction…………………………………………………………………………………4

Description of application……………………………………………………………4

Identification of Stakeholder Roles involved in the application and its governance……………………………………………………………………………..7

Identification of Ethical Risks……………………………………………………….8

Discussion of Mitigations Measures for Risk………………………………12

References…………………………………………………………………………………13

***Facebook’s (META’s) Content Personalization and Recommendation Algorithms***

**Introduction**  
Social media platforms have gone past their role as ways of communication, evolving into complex environments that shape how information is consumed and shared by the users. At the heart of this change are artificial intelligence (AI)-driven content personalization and recommendation algorithms, with Facebook being a prime example of this evolution. These algorithms are designed to give and recommend content that users prefer to see, therefore enhancing user engagement and satisfaction on platforms like Facebook. Even though these technical developments have greatly benefited society, they have also brought up important ethical issues that I will discuss in this report. In particular, the ethical environment of Facebook's recommendation and personalisation algorithms will be examined, along with important stakeholders, ethical ricks, and mitigation techniques.

**Description of Application**

**The Functionality of Facebook's Personalization and Recommendation Algorithms**

Facebook links billions of users worldwide. Their recommendation and content personalisation algorithms are what lets the company reach such a high number of users. These algorithms are engineered to make tailored feeds for individuals, suggesting content such as posts, news articles, videos, and advertisements based on their historical behaviour and preferences. This focus on tailored content delivery is central to Facebook's strategy, aiming to maximize user engagement and satisfaction by giving the user the most appealing content.

In order to deliver relevant posts to users based on what they like and engage with, sourced from ‘‘Here’s How the Facebook Algorithm Works in 2024’ sourced below, the algorithms main components include :

* **Listing of Posts**: Quickly gathers posts from individuals, groups, Pages, and commercial accounts, with a preference for non-commercial sources.
* **Signals Identification**: Uses hashtags, keywords, and other signals to understand and categorize post content.
* **AI Predictions**: Predicts engagement likelihood based on historical user data, including likes, group memberships, and post topics.
* **Relevancy Scoring**: Assigns scores to posts based on predicted user engagement, favouring content likely to interest the user.

Meta still needs to narrow down the options before recommending content to users. Facebook’s algorithm considers the following signals in **Figure 1** before pushing posts to the user’s news feed.



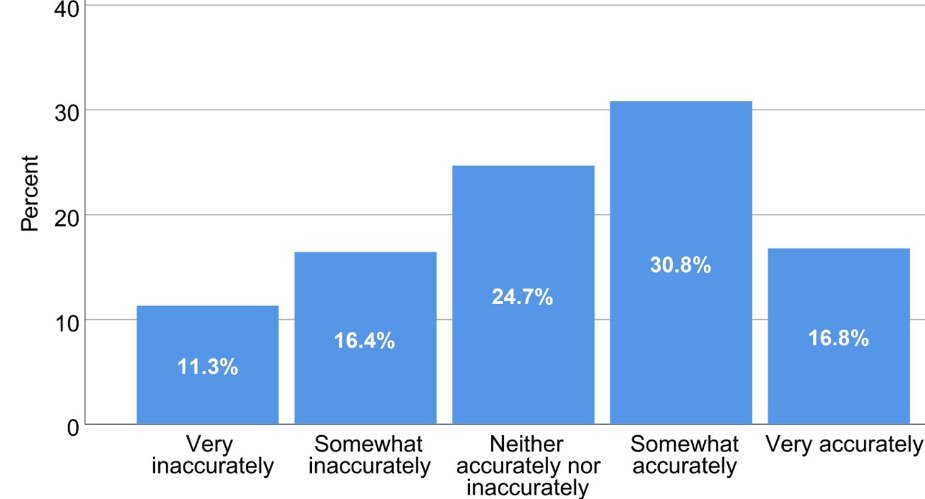
**Figure 1** : Content Recommendation Decision making. Image from ‘Here’s How the Facebook Algorithm Works in 2024’ referenced.

* **Type of content**: This is about whether the post is a video, picture, or link, and how long people spend looking at it compared to other similar posts. It basically asks what kind of posts are catching people's eyes the most.
* **Recency**: This refers to when the post was shared. Usually, newer posts are preferred because old news isn't as interesting to most people. There might be a few exceptions, but they're rare.
* **Relationships**: This looks at how much you interact with others and their pages on Facebook. For example, if you often like or comment on your college roommate's posts, you'll see their posts more on your feed. The same goes if you're active on a particular page.
* **Popularity or engagement**: If a post gets a lot of likes, comments, or shares, it means it's popular. Popular posts are shown to more people, even if they don't usually engage with that kind of content.
* **Creator's performance over time**: This considers how well the person or page who made the post has done in the past. If their posts usually get a lot of attention, their new posts are likely to be shown to more people.

**The Role of AI and Data Collection**

Facebook's AI algorithms are not fixed, they continuously learn from new data, which is fundamental to the effectiveness of the platform's content personalization. As users interact with content, the algorithms fine-tune the recommendations in real-time, creating a highly personalized and engaging user experience. These AI models progressively identify nuanced user preferences, enhancing the content personalization over time.

However, a study based on an online survey (taken from ‘Making sense of algorithmic profiling: user perceptions on Facebook’, referenced in the ‘References’ section) indicates that this process may not always accurately reflect individual identities. In the study, participants assessed the accuracy of Facebook's 'Your Categories' section in the Ad Preferences, revealing a spread of perceptions: 27.7% felt the categories reflected them inaccurately, while 47.6% believed they were accurate, as illustrated in **Figure 2**. This suggests a potential difference in the platform's ability to understand and reflect users' true interests and identities. Such insights emphasize the complexities and limitations of algorithmic profiling.



**Figure 2: the accuracy of Facebook's ad categories**

Another crucial aspect of Facebook's digital engagement strategy is its recommendation algorithms. These algorithms are intended to present users with fresh information, relevant pages, and groups based on their interests By analysing the collective data of users with similar interests and habits, the AI may suggest new content that individuals are likely to find fascinating even if they haven't previously interacted with similar content. Maintaining up-to-date and pertinent content not only boosts user engagement but also facilitates communication between content creators and advertisers and their target audience.

However, the dependence on AI and data collection for content personalisation and suggestion has brought to light concerns about data security, privacy, and the potential for echo chambers. By providing users with more control and more transparent data procedures, Meta has been attempting to allay these worries. over their data and the content they see.

**The benefits provided**

Facebook's AI-powered algorithms have revolutionised social media interaction by enhancing user experience, increasing engagement, and personalising content for individual users. By selecting material according to user preferences and behaviour, these algorithms increase user happiness and lengthen engagement times. They also aid in discovering new content, creators, and communities, improving the user experience, and community building. For advertisers and content creators, these algorithms offer targeted reach, improving ad value and resource use. Despite potential concerns about echo chambers, the algorithms promote content diversity and encourage exploration beyond usual interests, demonstrating the substantial benefits of AI in social media.

**Identification of Stakeholder Roles involved in the application and its governance.**

In analysing the complex landscape of Facebook's content personalization and recommendation algorithms, a variety of stakeholders play pivotal roles in its application.

**Direct Stakeholders**

* **Meta (formerly Facebook Inc.)** Meta's primary benefit is economic, deriving significant revenue from targeted advertising aided by personalized content. Leadership figures such as Nick Clegg, Meta’s president of global affairs, underscores the importance of direct stakeholders in steering the ethical direction of Meta's AI tools. Clegg plays a crucial role in the decision making in relation to META’s AI tools, as he “anticipates and addresses legal, regulatory, moral, ethical and cultural issues that may arise with the introduction of new AI products”, source – “Meta Addresses Legal and Ethical Issues in AI Development” referenced below.
* **Users** of the platform are both consumers and producers of content. Their interaction with the application feeds the data algorithms need for personalization, creating a loop that enhances their experience. The benefit for users is a more relevant and engaging social media experience, leading to increased satisfaction and time spent on the platform.
* **Advertisers** are pivotal in this ecosystem, leveraging the personalized recommendation system to target ads more effectively. The algorithmic targeting allows for a higher return on investment, as ads are more likely to reach interested users.

**Indirect Stakeholders**

* **Regulatory Bodies and Policymakers** globally impact the control of Facebook's algorithms. They ensure the platform adheres to privacy laws, content regulation, and ethical standards. Their involvement benefits the public by protecting privacy rights and ensuring content does not harm societal norms. The ongoing developments in AI control, as demonstrated by Meta's efforts to address ethical, legal, and cultural challenges, highlight the dynamic interplay between technology companies and regulatory frameworks aimed at ensuring responsible AI use.
* **Data Providers**, such as third-party websites and apps, indirectly participate by supplying additional user data to improve personalization algorithms. These entities benefit through partnerships, data sharing agreements, or financially.
* **Academic and Research Institutions** contribute through research and development, often working with Meta to improve algorithm models or investigate the social impacts of algorithm-driven content curation. Their involvement helps in developing more ethical and effective algorithms.

Each stakeholder in this big system gets specific benefits, from economic gains to societal value. The interaction between these roles is what I think to be creating this on going evolution of Facebook’s content personalization and ethical risks.

**Identification of Ethical Risks**

In this section I will Identify ethical risks, assess the severity and likelihood of occurrence, identify affected stakeholders and I will justify my assessments. Using the principles of social responsibility outlined ethical digital engagement lecture 2 slides, and the risk classifications by Latzer et al. (2016) from lecture 1 slides, I will identify and analyse the ethical risks associated with META's algorithms.

**Human Rights**

**Risks to Civil and Political Rights:** By interfering with democratic processes like elections and limiting freedom of speech, false news, social media bots, deepfake videos, filter bubbles, and censorship pose serious threats to civil and political rights. Referenced as ‘An Ethics Perspective On Facebook’ Michael Thate, who teaches Ethics at Princeton University expresses how Facebook's algorithms may prioritize engagement at the expense of these rights. Stakeholders affected include the general public, who may suffer from misinformation and manipulation, and political entities, whose reputations and operational integrity may be compromised. The severity of impact on these stakeholders is high due to the potential to influence election outcomes and public opinion. The likelihood of these risks is also high, due to the amount of content generated and shared on social media platforms. Justification for these assessments is based on documented instances of election interference and large dissemination of fake news.

**Complicity in Partner Violations of Rights:** META's collaboration with value chain partners who may commit rights violations presents a complicity risk. This includes partners involved in data collection and processing who may misuse personal data, violating privacy rights. Stakeholders affected are users of META's platforms, whose personal data may be at risk, and bodies reliant on META's algorithms for business operations. The impact severity can be high for individuals whose data is misused, leading to privacy breaches and related harms. The likelihood of incurring this risk is medium in my opinion, relying upon META's due diligence processes and partnership policies.

**Labour Practices**

**AI Automation Leading to Labour Displacement:** The advancement of AI within META's operations posits a significant risk of labour displacement, particularly affecting employees in operational and content moderation roles. I believe the impact severity as high, given the economic and social results for displaced individuals. I also consider the likelihood to be high due to the technological path towards increasing automation.

**Worker Physical and Mental Health:** The physical and mental well-being of content moderators and operational staff develops as a critical concern, with both the impact severity and likelihood high in my opinion. Employees in these roles are exposed to potentially upsetting content. This exposure shows the ethical requirement for META to establish robust support systems. My justification for the high impact severity is from the direct correlation between job function and exposure to harmful content, which can cause significant health issues.

**Legal Risks Arising from Discrimination in AI-Assisted Recruiting:** The deployment of AI in recruiting processes introduces potential legal risks related to discrimination, affecting potential employees. The impact severity is medium however, the likelihood is assessed as low in my opinion, due META's measures to refine AI algorithms and maintain ethical AI practices.

**The Environment**

**Increased Carbon Emission Due to AI Training and Service Operation**

The process of training complex AI models and maintaining META's extensive service operations has a significant environmental impact, chiefly due to the high energy consumption that leads to increased carbon emissions. This is worrying because it speeds up climate change, which has an impact on ecosystems all around the world and puts people's health at risk. The public at large, and actually the next generation are the stakeholders that are most impacted. The severity of the impact is considered high outlined as “AI has a fast-growing carbon footprint” by Sundberg, G, in Tackling AI’s Climate Change Problem, and the likelihood of occurrence is also high, given the ever-increasing demand for more sophisticated and engaging content.

**Resource Usage and Pollution from AI-Driven Product Creation and Disposal**

The production and disposal of AI-driven products pose medium environmental risks, affecting local communities, regulatory bodies, and environmental NGOs with medium likelihood and severity. Despite recycling efforts, challenges in managing resource depletion and pollution persist. META's role in this cycle demands enhanced sustainability, including renewable energy investments and better recycling, to mitigate impacts on key stakeholders.

**Fair Operating Procedures**

**Use of AI in Corrupt or Anti-competitive Practices:** This risk involves the manipulation of financial, investment, or procurement processes. Stakeholders affected include competitors, investors, and the general public. The impact on competitors and the market at large is High, due to the potential for market bias and unfair advantage. For investors, the impact is Medium, influenced by potential financial gains or losses. The likelihood of this risk is Medium, given the increasing scrutiny and monitoring of tech giants.

**Undermining the Public Political Process:** The use of AI for excessive lobbying, creating deep fakes, or spreading targeted misinformation impacts voters, political entities, and public trust in democratic processes. The severity of the impact on the public and democratic institutions is High, due to the potential to significantly influence public opinion and election outcomes. The likelihood of occurrence is also High, considering past instances of misinformation spread via social platforms.

**Violation of Intellectual Property Rights:** This involves the unauthorized use of copyrighted material through AI algorithms. Content creators and copyright holders are the primary stakeholders affected, facing a High impact due to potential loss of revenue. The likelihood of this risk is Medium, reflecting the ongoing challenges in effectively policing content on large platforms.

**Consumer Issues**

**Unsustainable Consumption**

The risk of fostering unsustainable consumption patterns is high, particularly as personalized content can encourage overconsumption by promoting products and services tailored to users' interests. This issue is outlined by Michael Thate again, by the "engagement game," a strategy where platforms like Facebook design algorithms to keep users engaged for as long as possible. This approach not only aims to increase the time users spend on the platform but also aggressively pushes them towards consuming more content and products, often without regard to the consequences on their financial well-being or the environment. Such relentless pursuit of user engagement and spending magnifies the risk of overconsumption. The likelihood of this risk is high, considering the algorithm's underlying goal to maximize engagement and spending. The impact on consumers is significant, leading to potential financial distress and contributing to environmental degradation.

**Misuse of Personal Data** The misuse of personal data represents a critical ethical risk, impacting users' privacy and autonomy. Stakeholders affected include the platform's users and external entities that might exploit this data without consent, as Facebook is a popular target for hackers, outlined by Jain, G, in ‘Ethical issues faced by Facebook’, referenced below. The severity of the impact is high, as misuse of personal data can lead to identity theft, privacy invasion, and manipulation. The probability of this risk is also high, given the vast amount of data collected and the complexities of data management and protection. My justification for this assessment comes from numerous instances of data breaches and misuse.

**Denial of Access to Essential Services** Denial of access to essential services, due to algorithmic biases or operational policies, affects marginalized communities the most. This risk has a medium likelihood, as efforts are made to improve inclusivity and access, yet gaps remain due to inherent biases in data and algorithm design. The impact of this risk is high for affected stakeholders, as it can worsen social inequalities by hindering access to vital information and services. My justification for this assessment comes from documented cases where algorithms have mistakeably marginalized certain groups.

**Community Involvement and Development**

**Difficulty Identifying Communities Suffering Negative Impact of AI Use**

Stakeholders affected include social networks, road users, medical patients, and even non-users. The risk severity for community cohesion and individual wellbeing is High, given the widespread influence of META's platform. The likelihood of this risk is also High, as algorithms often prioritize engagement over societal wellbeing, potentially leaving out the needs and harms to specific groups. My assessment is supported by the categorization of risks to societal wellbeing outlined by Latzer et al. (2016).

**Negative Impact on Local Health, Employment, and Wellbeing**

The pervasive reach of META's algorithms can worsen issues like deskilling, impact child development negatively, and fuel culture wars, thereby harming local health, employment, and general wellbeing. Stakeholders here include local businesses, educational institutions, and healthcare providers. The impact severity ranges from Medium to High depending on the community's reliance on digital platforms for information and social interaction. The likelihood of this risk is Medium, as it depends on local extent of digital integration in daily life. However, when these risks materialize, they can have profound long-term effects on community development.

**Concentration of AI’s Wealth and Income Creation**

This risk pertains to the economic disparities intensified by the centralization of wealth creation within AI-driven companies like META, at the expense of local communities. Affected parties include employees, local business owners, and the local economy as a whole. The likelihood of this happening is High, mirroring global trends of wealth accumulation in the tech sector and its effects on economic diversity and opportunity distribution.

**Discussion of Mitigations Measures for Misuse of personal data**

To address the misuse of personal data on Facebook, the mitigation technique that I have chosen is discussed in the study ‘Mitigating the Privacy Paradox Through Higher Privacy Literacy? Insights from a Lab Experiment Based on Facebook Data’ referenced below.

A lab experiment looking at whether more privacy literacy could help with the "privacy paradox," which refers to the difference between people's actual online sharing activities and their expressed worries about privacy, among Facebook users. The study found no relationship between Facebook users' levels of privacy concerns and how infrequently they use the social network. A greater tendency towards privacy-protective behaviour on Facebook is driven by a mix of high privacy literacy and high privacy concerns.

Improving user privacy literacy may be essential to reducing the risk of personal data exploitation. Creating and distributing thorough teaching materials about online privacy, the possible dangers of oversharing personal information, and useful techniques for controlling social media privacy settings are all part of this proposed strategy. To guarantee broad reach and effectiveness, workshops, online courses, and awareness campaigns could be planned with academic institutions, privacy advocacy groups, and regulatory agencies.

**Impact on Risk Reduction:** By closing this knowledge gap, users are more likely to take actions that protect their privacy, like changing their privacy settings to restrict data sharing and posting only the information they want to share publicly. By reducing the quantity of information that is easily accessible to third parties, this lessens the possibility that personal data will be misused and mitigates potential dangers to user privacy.

**Stakeholders Involved**:

* **Users**: Direct recipients of privacy literacy projects, responsible for applying knowledge to manage their privacy settings and data sharing practices more effectively.
* **Facebook** : Must develop and provide accessible privacy education resources and tools. They play a crucial role in designing user interfaces that promote understanding and ease of use of privacy settings.
* **Educational Institutions and Non-profits**: Can collaborate to design and implement privacy literacy programs, leveraging their expertise in education and outreach to reach a broader audience.
* **Regulators and Policy-makers**: Advocating for policies that require social media platforms to adopt transparent practices and provide comprehensive privacy education to users.

By empowering users and lowering the danger of data exploitation, addressing Facebook's misuse of personal data and educating consumers about it calls for cooperation from all parties involved in creating a safer online space where privacy issues are adequately handled.

**References**

[Meta, How Facebook distributes content

Available at: <https://www.facebook.com/business/help/718033381901819?id=208060977200861>

Meta, 2023, The AI behind unconnected content recommendations on Facebook and Instagram

Available at: <https://ai.meta.com/blog/ai-unconnected-content-recommendations-facebook-instagram/>

Ouchchy, L, Coin, A, and Dubljević, V. 2020, AI in the headlines: the portrayal of the ethical issues of artificial intelligence in the media

Available at: <https://link.springer.com/article/10.1007/s00146-020-00965-5>

Fosch-Villaronga, E, 2021, Making sense of algorithmic profiling: user perceptions on Facebook,

Available at : <https://www.tandfonline.com/doi/full/10.1080/1369118X.2021.1989011?scroll=top&needAccess=true>

Hitlin, P and Rainie, L, 2019 : Facebook Algorithms and Personal Data

Available at: <https://d1wqtxts1xzle7.cloudfront.net/59250358/Facebook_algorithms_report20190514-34280-jrzed3-libre.pdf?1557866216=&response-content-disposition=inline%3B+filename%3DFacebook_Algorithms_and_Personal_Data.pdf&Expires=1711802514&Signature=PDApGaKN-s1r92F~3LxS5Vdc2ZDR5~QeZrXDhoWhoG8jAOUBMjllNsMaTt~bveAPguTFtAcqX7URZlrG-D7yjTr9i7Eg0ALsM8RUsNKBMA0onTEDOEg89SlAXbUg5~b0p7MUxxlaT8n2ZpqyWW6ACSL3B~rtRgmEqOirylz9hOCv4WULEYs5mAamlV5E2rZXftpzhkgq9F5ODbqsgpLizE8HK7sZuj31~7fUJjTJ5BbFUNLhGl~Qws6mSSXB-aru1IZAEQvNEuDuPrWYRmOM~fY2dVZRU4Vn759j4v~iuFPKSPZLgMuptZREug8n7ROekzJ7WV5ZJRcKs3kK6hQtpg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA>

Miller, G, 2022, Stakeholder roles in artificial intelligence projects

Available at : <https://www.sciencedirect.com/science/article/pii/S266672152200028X>

PYMNTS.com, 2023, Meta Addresses Legal and Ethical Issues in AI Development

Available at : <https://www.pymnts.com/artificial-intelligence-2/2023/meta-addresses-legal-and-ethical-issues-in-ai-development/>

Tiku, N, 2023, AI researchers uncover ethical, legal risks to using popular data sets

Available at : <https://www.washingtonpost.com/technology/2023/10/25/data-provenance/>

Jain, G, 2021, Ethical Issues Faced by Facebook

Available at : <https://mvgt-gautam.medium.com/ethical-issues-faced-by-facebook-113a7296895f>

Steinhorst, C, 2021 : An ethics perspective on Facebook

Available at : <https://www.forbes.com/sites/curtsteinhorst/2021/10/22/an-ethics-perspective-on-facebook/?sh=6620a5251a3e>

Schaffer, N, 2024 : Here’s how the Facebook algorithm works in 2024

Available at : <https://nealschaffer.com/facebook-algorithm/>

Sundberg, N, 2023, Tackling AI’s Climate Change Problem.

Available at : <https://sloanreview.mit.edu/article/tackling-ais-climate-change-problem/>

Schubert, R, Marinica, I, Mosetti, L, and Bajka, S, 2022 : Mitigating the Privacy Paradox Through Higher Privacy Literacy? Insights from a Lab Experiment Based on Facebook Data

Available at : <https://deliverypdf.ssrn.com/delivery.php?ID=014117064098121090070101018064003069057020066018053053081103097094127113075127014093100057062011027030018064003002108013095003051035011087029126096106097121095096122061001031089117003116076068086030026022085120113025087028027122073067027023064111076002&EXT=pdf&INDEX=TRUE>