

Authentic and Justifiable Decision-Making: The Political Influence of Recommender Systems

Samuel French

Master Thesis

PSTS Final Project 20 EC (together with Internship): 201300089

University of Twente, Faculty of Behavioural, Management, and Social Sciences, Enschede, the
Netherlands

23/08/2022

Supervisor: Dr. Adam Henschke

Second Reader: Dr. Dina Babushkina

MSc Philosophy of Science, Technology, and Society – PSTS

Word count: 19955

Contents

Acknowledgements.....	4
Summary	5
1. Introduction	6
1.1. The Workings of Recommender Systems	7
1.2. Why are Recommender Systems Important?.....	8
1.3. Methodology.....	10
1.4. Overview	11
2. The Political Context	12
2.1. Reflective Endorsement.....	13
2.1.1. Freedom of Thought	14
2.1.2. Information	15
2.2. Deliberation, Democracy, and Justification of Decision-Making.....	16
2.3. Conclusion.....	18
3. The Influence of Recommender Systems on Political Decision-Making.....	19
3.1. Nudging.....	19
3.2. Recommender Systems and Nudges	20
3.3. Beyond Traditional Nudging Techniques	21
3.3.1. Hypernudges	21
3.3.2. Online Manipulation	22
3.4. Cambridge Analytica	22
3.5. Influence on Political Decision-Making.....	25
3.5.1. Influence of Recommender Systems on Reflective Endorsement.....	25
3.5.2. Echo Chambers	27
3.5.3. Influence of Recommender Systems on Democratic Deliberation.....	29
3.5.4. Deliberation vs Discourse	31
3.6. Conclusion.....	32
4. Solutions.....	33
4.1. Sociotechnical Solutions	33
4.1.1. The Shortcomings of Wholly Technical Solutions.....	35
4.2. Addressing Recommender Systems.....	35
4.3. Reflective Endorsement.....	36
4.3.1. Diversity of Information	36
4.4. Deliberation in Democracy	39
4.4.1. vTaiwan	39

4.4.2.	Facilitating More Open Discourse.....	41
4.4.2.1.	The Social Context of Political Decision-Making.....	43
4.4.2.2.	Targeted Political Advertising.....	43
4.4.3.	Recognise Lack of Deliberative Space.....	45
4.5.	Conclusion.....	47
5.	Final Thoughts.....	49
5.1.	Overview.....	49
5.2.	Recommendations.....	51
5.3.	Limitations.....	51
5.4.	Further research.....	53
5.5.	Final Statement.....	53
	Bibliography.....	55

Acknowledgements

Both for the internship and throughout this thesis, I would like to thank Dr Adam Henschke for his expert supervision, continued guidance, and stimulating feedback.

I would also like to thank Dr Dina Babushkina for her provoking critiques and welcoming support.

Finally, I would like to thank my parents whose support meant any of this was possible in the first place.

Summary

Recommender systems have transformed how people discover and interact with multimedia. Extensively found on major platforms such as Facebook, YouTube, TikTok, and Google, recommender systems employ big data analytics to disseminate information to millions across the globe. Among these recommendations are politically relevant content, such as news articles, political advertisements, blogs, opinion pieces, and podcasts. This thesis explores the relationship between recommender systems and political decision-making. Particularly, it analyses the effects that daily exposure to recommender systems has on the authenticity and justifiability of political decision-making.

In this thesis, I outline *reflective endorsement* and *deliberation* as necessary criteria for authentic and justifiable political decision-making. I then explore the relationship between recommender systems and these criteria, analysing whether the economic motivations and technological workings of such systems are conducive with authentic and justifiable political decision-making. Given that recommender systems have a tendency to prioritise misinformation and promote ideologically entrenched groups, I argue that they negatively affect citizens' ability for reflective endorsement and deliberation.

Identifying recommender systems as sociotechnical systems, I provide sociotechnical solutions that address the phenomena undermining the possibility of reflective endorsement and deliberation. These solutions do not attempt to impose changes on individuals but create less divisive and more reflective conditions in which citizens can engage in deliberative communication that can practically inform policy and decision-making.

1. Introduction

Every action and inaction online generates data. Whether it is the biographical information provided when creating an account, your search history, the people you interact with, your likes, reactions, favourites, and bookmarks, the locations and times of use, the notifications and pages you decide to click on, and even those you ignore, every action generates data. In isolation, some argue that this data is neutral and without meaning (Turilli & Floridi, 2009), or is mostly morally innocuous (Henschke, 2017, p. 202). Simply knowing phone *X* was last used at set of coordinates *Y* are singular pieces of data and meaningless. When ordered and combined with other data, however, it becomes *information* in a thick sense by virtue of acquiring a context. Compared to data, information possesses meaning and is thus more powerful (Henschke, 2017, p. 202). For instance, a single pixel on a screen is largely meaningless and morally irrelevant. However, if one zooms out, allowing the other pixels and data points to come into focus, the image starts to acquire meaning. The individuals pictured in combination with what they appear to be doing, the background of the photograph, and previous knowledge of cultural norms create meaning for the viewer. This is morally relevant as such an image could be intimate, contain sensitive information, or be evidence of wrongdoing. When applied to one's phone usage, given the user's past location data, their coordinates suggest whether they are at home or their workplace, and, similarly, their past phone activity will disclose whether they are likely to only check the time or if they will start browsing social media and are thus susceptible to engaging with recommended content.

Personal and behavioural data is collected precisely to make these predictions and inferences. It allows for apps and search engines to curate content to the perceived wishes of the user. These often come in the form of targeted advertisements and recommended multimedia, such as news articles, blogs, films and television, music, videos, clothes, restaurant suggestions and any variety of other content.¹ This content is curated and promoted by *recommender systems*, and, in ways similar to advertisements, influences peoples' behaviour (Susser et al., 2019). Be it through engaging users' attention, resulting in sales, generating comments, shares and clicks on links, recommender systems infer an individual's preferences and provide them with curated content to personalise their browsing experience and maximise engagement on their platform. Such systems have become ubiquitous and highly profitable throughout the internet,² arguably becoming the most visible and successful application of artificial intelligence (AI) in practice (Jannach & Jugovac, 2019, p. 1).

¹ Although advertisements and multimedia are seldom distinct.

² Describing recommender systems as ubiquitous within something as broad and diverse as the internet is perhaps contentious. As this work is focused on recommender systems in the context of political decision-making, I am primarily focusing on social media platforms and video-sharing sites that are commonplace in Europe. These include Facebook, YouTube, TikTok, Twitter, and Instagram but this is by no means an

1.1. The Workings of Recommender Systems

Recommender systems are automated forms of content personalisation algorithms. They are commonly found on major internet platforms, such as Google, Facebook, Amazon, YouTube, Instagram, TikTok etcetera. Recommender systems analyse an individual's data in conjunction with the data of other platform users to generate personalised recommendations (Elahi et al., 2021; Bozdag, 2013) and infer population-wide trends (Yeung, 2016). Moreover, as recommender systems are *adaptive algorithms*, meaning they enhance their workings based on the effectiveness of recommendations, the very process of recommending content improves the effectiveness of the systems (Bhargava & Velasquez, 2020, pp. 13-14; Chesson, 2018; Schou & Farkas, 2016). Typically, the desired end states for recommender systems are maximised user engagement and click-through rates (Seaver, 2018, pp. 429-431; Susser & Grimaldi, 2021, p. 5). Approaches to recommender systems include *collaborative filtering* (content that is recommended according to how it is valued by other users with similar interests) (Bozdag, 2013, p. 214), *content-based filtering* (associations that are based on a user's past choices and descriptions of new items with similar criteria) (Bozdag, 2013, p. 214), *reinforcement learning-based recommendations* (machine learning-based³ recommendations that self-improve, learning from the success or failure of AI-generated suggestions) (Mehdi Asfar et al., in press), and *hybrid methods* that incorporate a mixture of different techniques simultaneously.

These approaches are underpinned by *big data analytics*, meaning the techniques used to process the vast quantities of data and metadata generated and created by (internet) technologies (Yeung, 2016, p. 119). These analytical methods allow for the discovery of new trends which are inferred from large numbers of different variables and represent a fundamental shift in how knowledge is created, how we understand the world, and how businesses operate (Macnish & Galliot, 2020, p. 1). Through big data analytics, patterns and trends can be observed from quantities of data too vast for any human to consider, let alone gather new information from. These observed trends facilitate the improvement of algorithms and form the basis for personalised recommendations (Yeung, 2016). With collaborative filtering, for instance, big data analytics identifies new groups of people a user shares interests with, thus allowing for more personalised recommendations. So, instead of simply recommending to purchase the next book in the series, recommender systems

exhaustive list. Recommender systems are fundamental constituents of each of these platforms. By extension, less well-known sites and future platforms that also employ recommender systems could be considered as relevant to this work. For more extensive discussions regarding internet trends, the general characteristics of different stages of internet development, and predictions about which characteristics will define future stages, see Fuchs et al. (2010), Alabdulwahhab (2018), and Rudman and Bruwer (2016).

³ Machine learning is a form of AI that learns from data and experience to improve its performance towards a desired end state. To learn more about machine learning, its promises, shortcomings, and ethical challenges see Broussard (2018), O'Neil (2017), and Coeckelbergh (2020).

analyse the behaviour of thousands of shoppers with similar interests and subsequently provide suggestions across a variety of different domains.

For individuals, too, recommender systems represent a significant shift in how one discovers new music, films, shops, products, online content, news, and other multimedia. Extending far beyond word-of-mouth or traditional forms of advertising, recommender systems are designed to personalise a user's internet experiences, introducing them to content they may enjoy or be interested in and limiting or excluding irrelevant suggestions. In this sense, recommender systems should be thought of as a broad classification of algorithmic techniques that provide users with personalised content which includes but is not limited to targeted advertising. For example, in addition to targeted advertisements, recommender systems also disseminate news articles, videos, blogs, and other multimedia. Consequently, when referring to recommender systems throughout this thesis, targeted advertising should be understood as one of multiple forms of personalised recommendations that users interact with on a daily basis.

Ultimately, recommender systems are highly profitable and boost engagement. Their profitability is predominately derived from targeted advertising (Zuboff, 2019). The platform's advertising revenue is based on *click-through rates*, meaning their fee is dependent on the number of clicks an advertisement receives, as opposed to if anyone makes a purchase. As recommender systems use a variety of techniques to infer a user's preferences, they offer more sophisticated targeted advertising, reportedly resulting in a 30% increase in click-through rates (Kirshenbaum et al., 2012; Garcin et al., 2014). When applied to the scale of platforms such as Facebook, recommender systems generate significant profits. For instance, Facebook's advertising revenue was nearly \$115 billion USD in 2021 and is predicted to reach over \$200 billion by 2026 (Statista, 2022).

1.2. Why are Recommender Systems Important?

The techniques, ethics, and consequences of recommender systems are not unexplored territory. The frontpeople of Silicon Valley and Big Tech corporations publicly laud recommender systems. For example, CEO of Meta (formerly known as Facebook, Inc.), Mark Zuckerberg (2019), explained that "People consistently tell us that if they're going to see ads, they want them to be relevant." In this sense, understanding users' interests and behaviour through their likes, interactions and online activity is framed as to the benefit of the platform, the recommended party, and the end user. As the host company's revenue is mostly dependent on *click-through rates* (Zuboff, 2019, p. 82), meaning their fee corresponds to the number of clicks an advertisement generates, recommender systems benefit host companies as they allow them to target individuals with more personalised recommendations thus prolonging their engagement, exposing them to more advertisements and

therefore more chance of successful clicks and greater revenue (Zuboff, 2019, pp. 130-133). Those advertising their products, multimedia, and messages benefit as recommender systems provide a more sophisticated method of reaching new users by going directly to those likely to be interested and thus only paying for advertisements that receive clicks. Finally, end users also benefit as they experience recommendations that reflect their interests, discovering content that has been enjoyed by people similar to them thus making their online experience more personalised (Jannach & Jugovac, 2019).

On the other hand, recommender systems require philosophical consideration because they structure and define interactions that we deem to be valuable. For example, when recommending films, music, places to eat, and sights to visit they directly influence how people live their lives, spend their leisure time and money, and the extent of their cultural experiences. A significant socially valuable area of consideration is the extent of recommender systems' influence on political decision-making. For instance, news, political advertisements, petitions, opinion pieces, and blogs are routinely disseminated by recommender systems (Mittelstadt, 2016; Saunders, 2020; Bozdog, 2013). Through this process, recommender systems distribute information, providing access to what there is to know, how to know it, and what is not worth knowing (Gillespie, 2014). In other words, recommender systems determine what is politically salient for users to interact with and what is not (Gillespie, 2014). For example, imagine both news source *A* and news source *B* are covering topic *X* from different ends of the political spectrum. Based on the perceived interests of a user, a recommender system suggests news source *A*'s coverage. Firstly, irrespective of the news provider, the recommendation suggests that topic *X* is important enough to be worth knowing about. This necessarily involves dynamics of inclusion and exclusion as topic *X* has been deemed more salient than an alternative issue (Gillespie, 2014, pp. 170-172). Secondly, given the impartial framing of recommender systems (Gillespie, 2014, p. 179), recommending news source *A*'s coverage implies that *A*'s perspective is the better way of learning what there is to know about topic *X* and how one should know it, such as whether topic *X* is positive, negative, controversial etcetera. In this sense, recommender systems structure our relationships with information and mediate our understanding of the wider world, thus directly informing our political beliefs and actions. As liberal democracies are premised on individuals being capable of authentic decision-making (see chapter 2) (Susser, 2019), the nature and consequences of recommender systems require philosophical and ethical consideration to ensure their influence is not undermining democratic processes. The focus of this thesis, then, is much narrower than the general consumption of multimedia, such as news and advertisements. Rather, it centres on the effects that the daily exposure to politically-relevant recommendations has on the authenticity and justifiability of political decision-making, questioning the extent of the influence on this socially valuable activity.

In addition to mediating peoples' relationships with information, recommender systems are a form of behavioural influence. Through personalised recommendations, platforms attempt to change a person's beliefs, desires, and behaviours by influencing what they want, feel, and think (Susser & Grimaldi, 2021, p. 3). In chapter 3, I will outline *nudge* theory (Thaler & Sunstein, 2003, 2008; Sunstein & Thaler, 2003) and explain how recommender systems use these techniques to influence behaviour. This effect is particularly relevant in the context of political decision-making, as this thesis questions whether political advertising and curated news coverage have undue influence on a person's beliefs and actions. As recommender systems have been identified as having a tendency to promote homogenous content (Mittelstadt, 2016; Elahi et al., 2021; Rajtmajer & Susser, 2020), this paper explores the effects that repetitive information may have on (a) a person's ability to reflect on their beliefs and (b) propensity to interact with alternative views. This is fundamentally important for political decision-making in liberal democracies as democratic processes are (a) dependent on authentic voting and (b) pluralist beliefs are a prerequisite. Consequently, a further point of consideration is whether recommender systems facilitate healthy interactions between irreconcilable political opinions or if they create division and hostility.

1.3. Methodology

In this thesis, I use conceptual analysis to study political decision-making. As a philosophical method for studying the nature of concepts (Daly, 2010, p.10), adopting conceptual analysis allows me to define the subject matter of this thesis (Jackson, 2000, p.30) – namely, authenticity and justifiability. In chapter 2, I analyse the nature of authenticity and justifiability in the context of political decision-making. I undertake this to clarify what is meant by these concepts and demonstrate their importance for political decision-making, identifying that authentic decision-making is achieved through reflective endorsement and more justified political decision-making is dependent on deliberation.

The conceptual analysis of authenticity and justifiability serve as premises for my ethical analysis of political decision-making. Having outlined that in the context of political decision-making, authenticity and justifiability consist in reflective endorsement and deliberation respectively, I identify the ethical importance of these criteria. As democratic processes are a mechanism for assuring the equality of citizens, protecting and promoting values that ensure the integrity of political decision-making is of fundamental importance. Inauthentic and unjustified decision-making threatens democratic processes as the decision will not authentically reflect the wishes of the populace and will appear to be imposed upon them without consideration of the plurality of beliefs within a society. Through this ethical analysis, I argue that circumstances which threaten authentic and justifiable

political decision-making require moral consideration as they could undermine democratic processes, the mechanism by which a degree of equality amongst citizens is ensured.

In chapter 3, I analyse the relationship between political decision-making and recommender systems. I outline how the homogenous and ideologically entrenched recommendations endemic to recommender systems threaten reflective endorsement as they overwhelm users with (mis)information and limit the diversity of sources, thus reinforcing their beliefs and restricting reflection. I then argue that recommender systems offer a means of attempting to bypass deliberation through targeted advertising and give a false sense of having deliberated through the ideologically homogeneous groups they routinely disseminate. Given that recommender systems are obstructing reflective endorsement and deliberation, two processes by which the authenticity and justifiability of political decision-making are ensured, I am justified in proposing sociotechnical solutions in chapter 4.

1.4. Overview

This chapter has outlined recommender systems, introducing a variety of techniques for generating recommendations and explaining their basis in big data analytics. It provided an overview of the relationship between recommender systems, maximised engagement, and advertising revenue, demonstrating the profitability of these systems. This chapter demonstrated the philosophical importance of analysing recommender systems insofar as they structure our relationship with politically-relevant information and are a form of behavioural influence. Consequently, this thesis explores the extent of the relationship between recommender systems and political decision-making, questioning whether the authenticity and justifiability of political action are affected. Chapter 2 will outline the necessary criteria for the functioning of liberal democracies, focusing on the requirements for authentic and justifiable political decision-making. This will give further context to the types of reflective behaviour and deliberative activities that are conducive to more authentic and justifiable decision-making in liberal democracies. Chapter 3 analyses the relationship between recommender systems and these necessary criteria, identifying that the economic imperatives of the attention economy result in less authentic and justifiable political decision-making. Finally, in chapter 4, I provide sociotechnical solutions to respond to the undesirable effects of recommender systems identified in chapter 3.

2. The Political Context

This thesis analyses the relationship between recommender systems and political decision-making. When referring to *political decision-making*, its focus is on decision-making processes within liberal democracies. A prerequisite for such a democracy is the freedom of its population, with established inalienable rights that protect this. A liberal democracy is also defined by a degree of equality amongst citizens and established collective decision-making mechanisms for attaining the will of the populace, such as votes, elections, and referendums. Political decision-making, however, is not limited to organised votes and instead happens on a near-daily basis. Whether it is proposing or rejecting new policies, (re)distributing funding, responding to global or national events, providing aid, building new facilities, or choosing where to host the party conference, political decision-making incorporates far more than the build-up to and aftermath of an election.

Although policies and decisions are mostly determined by elected officials and other political decision-makers, the basis of their power, political legitimacy, and moral authority is gained through the will of the populace. Where campaign pledges and manifestoes establish the candidates' and party's intentions once elected, they do not provide a free pass for action or shield them from scrutiny. The freedom of the population ensures that the policies and their practical application remain open to criticism or debate from fellow party members, opposition parties and groups, and the public. Moreover, events beyond the decision-makers' control may lead to the creation of new policies and decisions to combat emerging issues. Particularly as they were not elected on these policies and promises, there will be greater scrutiny regarding whether their response is appropriate, sufficient, legal, or ethical.

This chapter explores the criteria necessary to foster and maintain authentic and justified political decision-making. Given that citizens play a fundamental role in selecting political decision-makers and what they have mandate to do, these criteria focus on how they arrive at their beliefs, but also incorporate what is necessary for relationships and interactions between (potential) decision-makers and the population. The first category I outline is authenticity criteria. This is *reflective endorsement* (see 2.1) which is necessary for individuals to determine their beliefs and conceptions of the good. The following requirement is *deliberation* (see 2.2) which acts as a justificatory mechanism by which political decisions become more acceptable.

This list of necessary criteria for political decision-making does not claim to be exhaustive. Rather, it focuses on two aspects that recommender systems play a significant role in – the dissemination of information and facilitating communication. Further research could incorporate the interrelationship between recommender systems and other politically important criteria, such as

freedom of association or freedom of expression. However, for the sake of this work, these considerations are beyond its scope.

2.1. Reflective Endorsement

For political decision-making processes in liberal democracies to accurately reflect the wishes of the population, it is dependent on citizens voting authentically. If voters are bribed, threatened, manipulated, or tricked into voting for or against candidates, a decision will not reflect the populace's will as the electorate was prevented from acting as they may have otherwise intended. In absence of foul play, authentic decision-making also requires reflecting on the content of one's beliefs and convictions to determine their value and suitability (Frankfurt, 1971; Korsgaard, 1996). This is called *reflective endorsement*, a process whereby one's behaviour becomes authentic if their actions and will are aligned (Frankfurt, 1971; Korsgaard, 1996; Rosner, 2000). To illustrate this, Frankfurt distinguishes between first-order desires, second-order desires, and second-order volitions (Frankfurt, 1971). First-order desires are simply wanting to do something or not (Frankfurt, 1971, p. 7), such as smoking. Second-order desires are the desire to have or not have a first-order desire (Frankfurt, 1971, p. 7). For instance, one may want a cigarette but also resent this desire as they know it is detrimental to their health. A second-order volition is when one wants to have a certain desire or wishes for this desire to be their will (Frankfurt, 1971, pp. 10-11). Here, their volition could be their will to quit smoking as they want to be a healthier version of themselves. Through reflective endorsement, people reflect on their desires and determine the suitability of their convictions and actions. It is through this process of reflecting on one's volitions and endorsing behaviour that directs them towards their desired goals that their behaviour becomes more authentic as their actions are not impulsive or unconsidered, but aligned to who they are or would like to be.

When applied to political decision-making, impulsively voting for a candidate because their name is first on the ballot would be inauthentic as the person has not reflected on their volitions and whether it is aligned with the candidate's policies. However, if they were to reflect on what they believe to be politically or socially valuable, interrogate their convictions and ask 'do I want to have this desire to vote for the candidate?' or 'is the world I would like to live in aligned with the intentions of the party I am voting for?', their decision-making would become more authentic when their actions and volitions are aligned. This form of endorsement is not limited to elections and is equally applicable when considering the authenticity of one's beliefs, party allegiances, and political expression. Consequently, this does not impose restrictions on what one is allowed to believe or whom they need to support, it merely creates conditions for ascertaining whether their actions and convictions are authentic insofar as they have been reflectively endorsed.

2.1.1. Freedom of Thought

To reflect on the content of one's actions and volitions, individuals require freedom of thought. If one is not able to think freely, then reflective endorsement will not be possible as their volitions are not theirs to endorse. Freedom of thought is the space for individuals to think freely without undue influence from external parties and it is widely protected as an inalienable human right, such as in the Universal Declaration of Human Rights (United Nations, 1948, Article 18) and the International Covenant on Civil and Political Rights (United Nations, 1966, Article 18). In political philosophy, too, freedom of thought and conscience are fundamental preconditions for participating in the political sphere (Mill, 1985, p.71; Locke, 1983) and is understood as the foundation for existing as equals in a free society (Rawls, 1999, p. 53).

For Rawls (1993, p.19), freedom of thought consists of each person's ability to form, revise, and pursue their *conception of the good*. A requirement for this conception is that it incorporates what the individual believes is valuable for human life, going beyond narrow conceptions of what is solely in their self-interest (Rawls, 1993, p. 19). For this, individuals need an autonomous space in which they can form opinions, consider merits of political options, and have authorship over their own beliefs. Applied to reflective endorsement, this process of forming and revising one's conception of the good can inform their second-order volitions. Conceptions of the good and second-order volitions, however, are not interchangeable. Second-order volitions are higher-level states where one chooses to endorse their will. Conceptions of the good are beliefs about what an individual thinks are valuable for human life. This is a special class of political belief that extends beyond one's self-interest. If one reflects on their conception of the good and has the volition to endorse it, one's behaviour will be authentic. For instance, if one's conception of the good is universal free access to healthcare, supporting a candidate who denies this would be inauthentic as the citizen's convictions and actions are not aligned. However, were they to reflect on their conception of the good and revise their belief in universal healthcare, then their endorsement of the candidate would become more authentic. Similarly, were they to reflect on their actions and withdraw their support for the candidate their decision-making would become more authentic as their conception of the good and second-order volition are aligned. As political decision-making that accurately reflects the wishes of the populace is dependent on citizens considering their beliefs and supporting candidates that are commensurable with their views, reflective endorsement is central to liberal democracies as the means by which the populace authentically establish their respective wills and conceptions of the good.

2.1.2. Information

To develop and reflect upon one's beliefs, life plan, and conception of the good, they require access to reliable and accurate information (van den Hoven and Rooksby, 2008, p. 382).⁴ Reflective endorsement, then, does not happen in a vacuum but requires additional criteria for citizens to effectively endorse their beliefs. For informed and authentic political decision-making, the populace requires access to accurate information about their society, the rest of the world, and the intentions of political decision-makers. Without this, citizens are unable to make informed choices during elections as they lack the information on which to authentically base, reflect, and act upon their volitions. Moreover, given the need of freedom of thought for reflective endorsement, information is required to help inform individuals' conceptions of the good and volitions as opposed to constitute it. If information sources begin to overwhelm and close off their space to reflect, their behaviour will become less authentic as it will no longer be aligned to their will.

To illustrate the necessity of accurate information for effectively endorsing one's conception of the good, consider a pledge from the 2019 United Kingdom general election. One conception of the good that may be shared by many of the electorate is that the health and well-being of citizens is important. A means of pursuing this is providing the health service with greater infrastructure and funding. Consequently, the Conservative and Unionist Party manifesto (2019, p.2) promise of 'building forty new hospitals' reflected this conception, gaining widespread support. However, when it later emerged that a 'new hospital' was defined not only as 'an entirely new hospital', but 'a major new clinical building or wing of an existing hospital' or 'a major refurbishment and alteration of an existing building' (Department of Health and Social Care, 2021, p. 12), there were accusations that the electorate had been misled (Barrett, N. & Palumbo, D., 2021; Walker, P. & Campbell, D, 2020). In this situation, where the individuals were able to freely establish their conception of the good, defining this in part as improved health services, the quality of information they received limited their ability to endorse their beliefs. Henschke (2017, pp. 230-232) would describe this representation of their policy as an *incomplete information harm*. Where in a narrow sense, according to the Conservative and Unionist Party definition, it is strictly true that forty new hospitals will be built, this definition is decontextualised from the common understanding that a 'new hospital' is built where previously one did not exist. The policy failed to give enough contextual information for citizens to understand the coalition's genuine intentions. This becomes a harm when the decontextualised information is recontextualised by others, using their interpretation to guide their action. Incomplete information limits authentic political decision-making as effectively endorsing one's actions becomes more

⁴ This is not to say that each person is required to have a *rational* life plan, but that one requires access to information in order to *rationally* pursue their goals, be these rational or otherwise.

challenging. To pursue behaviour that authentically corresponds to one's volitions, access to reliable and accurate information is required. Political decision-making, then, can be described as authentic if their volitions are their own and they have reflected on and endorsed their conception of the good. To effectively pursue this conception of the good, however, one requires accurate and reliable information sources.

2.2. Deliberation, Democracy, and Justification of Decision-Making

Reflective endorsement alone is insufficient for ensuring the integrity of the democratic process. When a populace is comprised of people with highly diverse backgrounds, religions, ages, ethnicities, genders etc., there is an equally eclectic range of priorities, beliefs, ideologies, cultures, and conceptions of the good – especially when individuals have been given the freedom to authentically form their own opinions. Naturally among citizens, there will be some beliefs held by the majority, some by relative minorities, and other concurrently-held antithetical positions which seem to have effectively equal support. To overcome these disagreements, particularly in cases where an outcome will be intolerable to one group, one can take inspiration from *deliberative democracy*, a position that understands the justification of political decision-making to be dependent on *deliberation*, as opposed to the mere aggregation of votes (Boham & Rehg, 1997, pp. ix-x; Cohen, 1997a, pp. 67-69). This is largely because democratic processes often favour the majority group, meaning minority opinions can easily be ignored, leading to them becoming alienated or disenfranchised. Instead, deliberation is a process for discussing alternative conceptions of the good, outlining their details, and determining how to apply them to policy issues (Rawls, 1999, p. 362; Cohen, 1997a, p. 68). Precisely because deliberation over a decision requires listening to alternative perspectives, outlining merits and pitfalls, explaining reasonings, and making concessions, the final policy or consequence becomes more justifiable to citizens as they can recognise how it benefits society and aligns with the beliefs of those within it. Furthermore, as deliberation encourages and requires input from different members of the population, the decision becomes more justifiable and legitimate as the public has participated in the decision-making process, as opposed to being subject to it (Fishkin, 2009, pp. 76-78). Through participating in deliberative activities, seeing how one's conception of the good contributes to the decision, and understanding the reasoning behind the decision, political action becomes more justifiable. By extension, subjecting citizens to decisions without consultation, consideration, or explanation makes the action less justifiable.

Under deliberative democracy, prerequisites for more justified decision-making are public deliberation towards the common good, some degree of equality amongst the citizens, and that the deliberation shapes the identity and interests of the citizens in a way that contributes to the common good (Cohen, 1997a, p. 69). Although there are differing accounts of precisely what processes

deliberative democracy should or must consist in (Habermas, 1991, 1996; Cohen, 1997b; Fishkin, 2009, 2016), it fundamentally presupposes a state of *reasonable pluralism*, meaning the populace holds multiple reasonable but incommensurable political, religious, and philosophical values or beliefs (Rawls, 1991, p. 36). Inasmuch that these irreconcilable beliefs are *reasonable*, they are not founded solely on the interest of themselves or their group, but arrived at through rational and practical reasoning. Deliberation acts to justify political decision-making, basing the justification for exercising political power on the process and products of free public reasoning amongst people (Cohen, 1997b, pp. 412-414). This is not to entirely subscribe to deliberative democracy or argue for the superiority of this approach but to outline the importance of deliberation and pluralism when justifying political decision-making in liberal democracies. In other words, some degree of interaction and the possibility of concession is a necessary constituent for justifying reasonable yet irreconcilable worldviews.

In practice, deliberation can happen through a broad range of activities. This includes (one-to-one) interactions between individuals, constituents and their elected officials, political decision-makers, and within groups. Additionally, deliberation can occur between a community of individuals and political decision-makers, both nationally and internationally. This includes but is not limited to hosting consultations with the local community, panel discussions, debates, question and answer sessions, and radio phone-ins. These mechanisms provide the ability for citizens and political decision-makers to discuss varying conceptions of the good and how they may be achieved through policy and action. Given the requirements that deliberation ought to shape the interests and identity of citizens, a necessary condition for deliberative activities is that they engage with citizens' beliefs as opposed to dogmatically defending one's intentions. Moreover, a decision does not require concessions for it to be justified. Rather, the process of deliberating and establishing how a decision or policy is conducive to the common good is sufficient if it includes participation and an openness to engaging with competing ideals.

As with reflective endorsement, information plays an important role in deliberation. As a process through which conceptions of the good are openly shared and listened to, participants will routinely provide each other with information about their convictions. Furthermore, political decision-making presupposes a particular context about which deliberation is happening. For instance, to borrow an earlier example, the Conservative and Unionist party pledge regarding forty new hospitals was in response to a social context where citizens believed health services were underfunded and under-resourced (Barrett, N. & Palumbo, D., 2021; Walker, P. & Campbell, D, 2020). Were the Conservatives to engage in deliberation about this policy, the proposal would likely be unjustifiable to someone who has been exposed to a misrepresentation of the context – instead being told that there are currently too many hospitals, with many sitting empty. No degree of explanation would satisfy this

hypothetical individual as the policy and discussions regarding its merits would be failing to address the context their conception of the good corresponds to. Deliberation, then, requires that participants have access to accurate information about the social context in order to be receptive to others' beliefs and for their conception of the good to be reasonable. If participants and their conceptions of the good are all based on vastly incommensurable information sources, more justified political decision-making will not be achievable.

Although I have argued that some degree of deliberation is required for more justifiable political decision-making, the process of deliberation is often not preferable on the individual level. Typically, people prefer to rely on their routine scripts for approaching the political world and its social issues (Ryfe, 2005, p. 56). Deliberation requires going beyond one's everyday reasoning habits as their beliefs are challenged, often making them feel anxious or frustrated (Ryfe, 2005, p. 56). Accepting one's convictions is easier than interrogating them, whilst having them challenged by others is harder still. As such, having deliberation as a necessary requirement for more justified political decision-making entails elected officials and the population striving to go beyond typical political discourse, accepting the fallibility of their conceptions of the good, and a willingness to discuss their reasoning. Incorporating deliberation, then, requires collective commitments to deliberating rational conceptions of the good among the populace and political decision-makers.

2.3. Conclusion

In this chapter, I have outlined reflective endorsement and deliberation as necessary criteria for more authentic and justified political decision-making in liberal democracies. I argued that reflective endorsement is necessary for authentically ascertaining the will of the populace and that deliberation achieves more justified decision-making by virtue of citizens' participation and incorporating the plurality of people's beliefs. Moreover, for both criteria, I identified the importance of accurate information. In chapter 3, I outline recommender systems as a form of *nudging* that attempts to influence user behaviour, before exploring how this affects the possibility of reflective endorsement and deliberation. This identifies how recommender systems can be argued to make political decision-making more authentic and justifiable, but also how, at times, these criteria are inadvertently undermined by the economic demands of the attention economy.

3. The Influence of Recommender Systems on Political Decision-Making

This chapter provides an overview of recommender systems, describing their behavioural influence as a form of *nudging*. First, I outline nudges, providing examples, before demonstrating how recommender systems exhibit these techniques. I then explain how recommender systems go beyond traditional forms of nudging using the work of Yeung (2016) and Susser (2019; Susser et al., 2019). Based on this difference, I explore whether the behavioural influence manifest in recommender systems matters in the context of political decision-making. Building on chapter 2, I outline reflective endorsement and deliberation as necessary requirements for political decision-making to be more authentic and justifiable for citizens.

3.1. Nudging

A *nudge* is “any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives” (Thaler & Sunstein, 2008, p. 6). In other words, by making seemingly insignificant changes to the decision-making context, it is possible to affect people’s behaviour, making particular options more likely to be chosen. This is based on findings in cognitive psychology which demonstrate that much of human decision-making occurs subconsciously, without active reflection (Kahneman, 2012; Tversky & Kahneman, 1981). Instead, these decisions are guided by heuristics and cognitive shortcuts. Consequently, it is possible to construct decision-making contexts that subvert reflection on one’s actions and thus limit reflective endorsement.

This form of behavioural influence is not necessarily insidious or a threat to free will as an entirely neutral choice architecture is not possible (Thaler & Sunstein, 2008, p. 3). Rather, all designs have some form of choice architecture and thus influence an agent’s decision. Moreover, by definition, a nudge does not completely remove freedom of choice as it is a technique for making certain outcomes more likely (Schmidt & Engelen, 2020, p. 4). Thaler and Sunstein (2003, 2008; Sunstein & Thaler, 2003) describe this form of influence as *libertarian paternalism*, meaning individuals are free to do as they please but the choice architecture directs them towards behaviour that – as an example – is claimed to make their lives better, longer, and healthier. This can be through increasing the proximity and reducing distance between desired choices, such as placing healthier food alternatives at the supermarket checkouts (Kroese et al., 2015); through determining default settings thus opting out becomes an active choice, such as in Austria where all citizens are placed in the organ donor pool but can opt-out at any time, with 99.89% remaining as donors (Goldstein et al., 2008); and through changing the salience of certain options or making them more difficult to choose, such as making more restrictive privacy permissions comparatively harder to find and select in cookie banners (Bauer et al.,

2021). Where this overview of techniques is non-exhaustive, it demonstrates that nudging is a diverse form of behavioural influence that, when applied, can be used towards arguably positive or negative ends. When reflecting on choice architectures, then, it is important to consider whether its design tries to elicit desirable behaviour (Susser et al., 2019), whether its goal is transparent (Mittelstadt, 2016), does it impose on peoples' autonomy (Susser et al., 2019; Susser, 2019), and if there are ethical or normative grounds on which to base a criticism.

Mitchell (2004), however, argues that libertarian paternalism mistakenly prioritises choice architectures that promote welfare, neglecting its libertarian foundations. For instance, by placing healthier food alternatives near supermarket checkouts, choice architects impose welfare values on individuals when a *libertarian* paternalist account should seek to disincentivise irrational choices and promote liberty, perhaps providing a diverse range of food choices instead. By taking directive choice architectures as a given and promoting welfare over liberty, libertarian paternalism is simultaneously overly paternalistic and insufficiently liberal. In order to more completely address commitments to libertarianism and paternalism respectively, it is important to consider whether a choice architecture encourages rational reflection about one's desires and volitions or if it continues to exploit cognitive shortcuts, merely replacing the intended outcome with actions targeted at promoting welfarist conceptions of better, healthier lives. Although not outright damning to libertarian paternalism, Mitchell's criticism is a valid perspective to consider, particularly when reflecting on the solutions I will propose in chapter 4. Namely, it will ask whether the proposed nudging techniques encourage reflective endorsement of one's actions – and thus more authentic behaviour – or if they simply redirect unconscious actions towards alternative ends.

3.2. Recommender Systems and Nudges

Recommender systems are diverse in appearance but demonstrate a variety of nudging techniques. For instance, the position of content on your news feed requires careful consideration as placing promising recommendations beyond the 300 posts typically seen by a user each day is unlikely to be fruitful (Luckerson, 2015). Additionally, whilst scrolling, images are often incomplete, revealing to the user what is ahead and prompting them to continue scrolling to relieve their curiosity (Eyal, 2014, p. 110). Moreover, headlines, thumbnails, reactions, comments, trending topics, and push notifications all serve as mechanisms for subconsciously redirecting users' attention towards particular content. For example, 'BREAKING NEWS' is visually striking and suggests the importance of an article, whereas a provocative image may redirect the user's focus, and posts that have mostly received 'angry' or negative reactions are more engaging than those which users 'like' or 'love' (Merill & Oremus, 2021).

As such, any variety of nudging techniques are employed to influence user behaviour to prolong engagement, generate clicks, and interact with promoted content.

Recommender systems are a ubiquitous technique within the *attention economy* (Bhargava and Velasquez, 2020).⁵ As outlined in chapter 1, recommender systems are tasked with analysing users' behaviour and interests to generate personalised recommendations which in turn increase engagement and click-through rates. This business model features a cyclical approach whereby the longer a user spends online, the more recommended content they are exposed to, thus increasing the likelihood of successful clicks. Simultaneously, their browsing generates behavioural data about their preferences, interests, and things they did not like, allowing for more sophisticated recommendations (Schou & Farkas, 2016; Chesson, 2019; Zuboff, 2019, pp. 130-133). Maintaining users' attention is imperative within this business model as it increases the likelihood of successful clicks both in terms of proximity to content and how accurately it is recommended.

3.3. Beyond Traditional Nudging Techniques

In one sense, recommender systems have choice architectures similar to other forms of attempted behavioural influence, such as advertisements. By *influence*, this means purposefully intervening so as to change beliefs, desires, and behaviour (Susser & Grimaldi, 2021, p. 2). That is, they attempt to change what people think, feel, and want. This is analogous to Thaler and Sunstein's definition of a 'nudge' as it attempts to alter their behaviour without forbidding options or changing the economic incentives. However, recommender systems go beyond traditional forms of nudging inasmuch that they have highly adaptive, diverse, and personalisable choice architectures (Yeung, 2016, p. 122; Susser, 2019, p. 2).

3.3.1. Hypernudges

As recommender systems are underpinned by big data analytics, their choice architecture can be refined in real-time. Yeung (2016) observes these dynamic processes happening in three directions: (a) the refinement of the individual's choice environment in response to their behaviour and constantly expanding data profile; (b) data feedback to the choice architect that is collected and repurposed for other big data applications; and (c) refinement of individual's choice environment based on population-wide trends. Consequently, recommender systems and their choice architectures are constantly adapting to the users' data, generating more sophisticated nudging techniques and exerting more behavioural influence. Such dynamic refinement goes beyond structuring the supermarket layout to promote healthier choices as now it is as if each individual

⁵ Other techniques include eroding natural stopping queues, exploiting the desire for social validation (Bhargava & Velasquez, 2020), and intermittent variable rewards (Griffiths, 2018, pp. 66-67; Bosker, 2016).

simultaneously interacts with a different layout (Susser, 2019, pp. 2-4). To reflect this, Yeung (2016) labels big data-driven recommendations as *hypernudges*.

3.3.2. Online Manipulation

Susser (Susser et al., 2019) argues that some nudges go beyond the techniques outlined by Thaler and Sunstein (2003, 2008; Sunstein & Thaler, 2003), instead becoming a form of online manipulation. This is not to say that all nudges in the online environment are manipulative but that this distinction is dependent on whether the nudge is hidden and exploits vulnerabilities (Susser et al., 2019, p. 6). For instance, Susser is in favour of nutritional labels as they encourage individuals to hesitate and make more conscious choices about their food selection. However, consider online nudges that try to exploit cognitive biases without the user's awareness. This undermines their autonomy, leading them to act for ends they have not chosen and for reasons that are not their own (Susser et al., 2019, p. 9). Moreover, online manipulation limits or subverts an individual's capacity for reflective endorsement by concealing that a decision was ever made. Targeted advertising informed by emotion analytics could fall into the manipulative category, given that it would employ the sentiments and behavioural shortcuts associated with each emotion (Zuboff, 2019, pp. 282-284). For example, waiting for a user to demonstrate feelings of sadness or inadequacy before exposing them to adverts would fall into Susser's definition of online manipulation as it is covertly exploiting their cognitive vulnerabilities to sell products. When considering recommender systems as a form of behavioural influence it is important to consider whether the nudging techniques stray into the manipulative territory that precludes reflection on the authenticity of one's actions.

Online manipulation goes beyond nudging as it does not encourage the user towards actions that will make their life healthier or better, but waits to exploit their vulnerabilities for ulterior motives. Under libertarian paternalism, the user remains free to act as they please whilst being faced with a choice architecture that encourages actions that are supposedly in their long-term interests. Online manipulation, on the other hand, exploits cognitive vulnerabilities to the benefit of the choice architect and does so in part by concealing that a decision was ever made. This classification is also different to hypernudging. Where the focus of hypernudging is curating personalised choice architectures to exert greater influence over a user, online manipulation focuses on exploiting specific vulnerabilities. Hypernudging, then, should be thought of as hyper-personalised choice architectures, whereas online manipulation centres on hidden, exploitative nudging.

3.4. Cambridge Analytica

Yeung's label of *hypernudging* and Susser's concerns about online manipulation are typified by the now-infamous practices of Cambridge Analytica. Building on research that established a strong

empirical correlation between a person's Facebook activity and their 'OCEAN' psychological profile (openness, conscientiousness, extraversion, agreeableness, and neuroticism) (Kosinski et al., 2013), Cambridge Analytica used online targeted advertising to influence voter behaviour during elections (Isaak & Hanna, 2018). They developed a personality quiz to determine people's OCEAN score which included a third-party blanket privacy permission, meaning those who took the quiz unwittingly gave Cambridge Analytica access to the data of every person on their friends list, resulting in the data of 87 million Facebook users being harvested, many without explicit consent (Hern, 2018). By integrating this data with a range of other sources purchased from data brokers, they established 5000+ data points on 230 million US adults (Anderson & Hovath, 2017). This allowed Cambridge Analytica to *behaviourally micro-target* internet users, claiming they could identify whom they would vote for, their values and reasons for doing so, and whether their vote could be changed (Anderson & Hovath, 2017).

The advertisements were microtargeted according to a user's perceived personality. To illustrate this, then-CEO, Alexander Nix (2016), writes that their analytics identified a small pocket of voters in Iowa, United States who strongly believed that citizens should show an ID card whilst voting. Subsequently, 'temperamental' individuals, as those who were categorised to typically dislike commitment, received adverts suggesting "it's as easy as buying a beer", whereas those labelled as 'stoic traditionalists' saw messaging stating that showing an ID is simply part of the privilege of living in a democracy. This level of microtargeting allowed Cambridge Analytica to target adverts at millions of internet users during the 2016 United States Presidential Election, 2016 United Kingdom European Union membership referendum, Nigeria's 2015 Presidential Election, and more than two hundred elections around the globe (BBC News, 2018). In each of these democratic events, Cambridge Analytica attempted to identify those who could be persuaded to vote for their client or discouraged from voting for a rival candidate. Where it is not possible to calculate the influence that Cambridge Analytica had over election results, the global scale of their operations means that even if they were partially successful, we should be significantly concerned with the effects of political recommendations. For instance, in the United States alone, microtargeting delivered 1.4 billion exposures across 4000 different campaigns supporting Trump's presidential bid (Ward, 2018, p. 140). Millions, then, were repeatedly exposed to personalised advertisements that attempted to influence their voting behaviour which is politically significant in itself.

For this paper, the most salient aspect of Cambridge Analytica's practices is the use of targeted advertising and recommender systems in the political context. For instance, in Donald Trump's 2016 United States Presidential campaign alone, it is estimated that up to 50000 variants of an advertisement were run each day, with the content being continuously adapted and improved based

on the audience's response (Cadwalladr, 2016). This typifies Yeung's (2016) classification of hypernudging, as the advertisements were consistently improved to reflect user behaviour, data feedback, population-wide trends and responses to the advertisements. Moreover, the practices allude to Susser's concerns of online manipulation as the microtargeting exploited vulnerabilities in the users' psychological profiles and tried to utilise hidden cognitive shortcuts as opposed to promoting conscious reflection about their volitions, conceptions of the good, and whether the political candidates would pursue this (Susser et al., 2019).

Where the practices of Cambridge Analytica were heavily criticised for, among other things, the mass privacy violation in the quiz's terms and conditions, the level of psychological manipulation employed in its advertising, and the hidden nature of all these processes, recommended political content is still ubiquitous in online environments (see Anderson & Hovath, 2017; Hern, 2018). Within Europe, it would be challenging to operate to the same extent as Cambridge Analytica as the opaque and misleading privacy policies they employed have been outlawed by the General Data Protection Regulation (2016, Article 12 (1)). Despite this, recommender systems still diffuse politically-relevant content to users on a daily basis. Such recommendations come in the form of official party advertisements or messaging; through news articles that promote or have an allegiance to particular political beliefs; documentaries, podcasts, blogs, opinion pieces, thought leadership, and any content that is committed to particular ideologies, beliefs, and/or political parties with or without formal support from the party; and trending topics and discussions regarding politically-relevant topics. The reach of this information is also significant. For instance, in the UK, 70% of people engage with news online, with roughly 50% using social media as a news source (Statista, 2021). Moreover, 35% of people aged between 11 and 16 said their main news source was social media sites (Parnaby, 2022). Recommender systems, as a constitutive part of most social media sites, play a significant mediatory role between the populace and their access to politically-relevant multimedia, affecting political decision-making both today and among the voters of the future. Consequently, an important area of philosophical consideration is whether current recommendation techniques that bear similarities to Cambridge Analytica promote or undermine the context of political decision-making. This assessment focuses on political decision-making as a valuable social activity through which citizens reflect upon and authentically endorse their conceptions of the good, making the process and outcome itself more justifiable through deliberation. The analysis, then, goes beyond a purely consequential assessment of democracy, understanding the process of political decision-making to be valuable in itself rather than a mere mechanism for determining power.

3.5. Influence on Political Decision-Making

Having outlined the necessary criteria for more authentic and justifiable political decision-making in chapter 2, the next sections explore the relationship between these criteria and recommender systems. Particularly, it asks whether recommender systems help the populace freely form and reflect on their conceptions of the good and facilitate deliberation.

3.5.1. Influence of Recommender Systems on Reflective Endorsement

As outlined in 2.1, reflective endorsement is the process by which individuals arrive at authentic behaviour and beliefs. Through reflecting on one's desires, determining their suitability, and endorsing behaviour that is aligned with one's volitions, individuals act authentically as they have considered who they are or would like to be and directed their activity towards this end state (Frankfurt, 1971). I argued that freedom of thought and access to reliable information are necessary requirements for reflective endorsement. If one does not have space to think freely, reflecting on the content of their actions and volitions will not be possible. Furthermore, one needs access to reliable, truthful information to form and effectively endorse conceptions of the good that correspond to the current state of the world. Despite one's best intentions, if their information about the world is inaccurate, reflectively endorsing behaviour that corresponds to their volitions will be challenging. For instance, if one wishes to be a healthier version of themselves yet has been told smoking is an effective way of strengthening their respiratory system, despite authentically endorsing cigarettes, their behaviour will not be directed towards their intended volition. In this sense, to effectively pursue one's volitions or conceptions of the good, one requires access to reliable sources of information.

One could argue – together with Purcell & Rainie (2014) – that recommender systems aid reflective endorsement inasmuch that through suggesting politically-relevant information to users, they become more informed, providing greater knowledge with which to develop and interrogate their conceptions of the good. This is because individuals are dependent on information to develop and rationally pursue their conceptions (van den Hoven & Rooksby, 2009). Particularly as this now happens throughout their daily internet activities as opposed to actively seeking out news, meaning the politically-relevant information, in a sense, comes to them. In fact, it can be argued that precisely because traditional forms of politically-relevant multimedia, such as legacy news, have previously and still act as *gatekeepers* to newsworthy content (Smith et al., 2001, pp. 1400-1401), recommender systems provide more diverse information. For instance, recommender systems allow the public to set the news agenda by collectively determining what is *trending*, as opposed to a small group of powerful individuals in media organisations. Consequently, individuals' propensity for authentic

political decision-making is increased due to more readily available news sources and improved access to information, particularly as the access becomes less contingent on gatekeepers.

This argument, however, appears to be mistaken as the personalisation manifest in current recommender systems limits the diversity of sources a user interacts with. As the big data analytics underpinning recommender systems infers a user's preferences based on their behaviour, the behaviour of other statistically similar people, and population-wide trends (Yeung, 2016, p. 122; Zuboff, 2015, pp. 80-81), there is a risk that, in the political context, recommendations will not be diverse or offer alternative points of view (Mittelstadt, 2016, p. 4992; Susser et al., 2019, p. 12). Instead, recommendations are often homogenous, repeatedly suggesting content the user has already demonstrated an interest in, such as consistently recommending the same political commentator after a user interacts with one of their videos. This is perpetuated by the underlying economic imperative of increasing *click-through rates* as the platform's revenue is dependent on amassing clicks, making similar content a safer bet. Consequently, news recommendations do not challenge users with diverse or novel ideas (Elahi et al., 2021, pp. 105-106). This results in a phenomenon called a *filter bubble*, referring to recommender systems filtering out suggestions that are contrary to the user's beliefs, leading to them becoming isolated in a homogenous ideological bubble (Pariser, 2012). This limits individuals' ability to reflect about politically-relevant topics as their sources of information continuously confirm and reinforce their beliefs, resulting in unconsidered conceptions of the good. This is not to say that authenticity is dependent on one continuously changing their beliefs but that a variety of information sources and space to reflect are required to rationally consider alternatives. Within a filter bubble, the process of authentic political decision-making is obstructed by homogenous information sources reaffirming and even constituting a person's beliefs.

Moreover, recommender systems amplify the bias and agenda-setting issues associated with legacy media as the former's suggestions are framed in an air of neutrality and objectivity (Bozdag, 2013), and remain largely hidden in daily practices. Traditional media may have ideological allegiances or political agendas, meaning their coverage of topics can be biased or omitted altogether. However, there is a significant awareness of this phenomenon because it is associated with the wishes of a select few individuals or a widely accepted fact, consequently making it easier to govern (Ratjmajer & Susser, 2020, p. 1). For example, media moguls, such as Rupert Murdoch, have held significant influence over the news agenda, determining what is newsworthy and the tone of its coverage (McKnight, 2013). Furthermore, coverage by *Fox News* or *The Guardian* is well-known for its association with right- and left-wing beliefs respectively. Comparatively, recommender systems are not framed as having similar degrees of bias or ideological commitment. As a form of big data analytics, they are portrayed as neutral and objective, merely reflecting a user's perceived interests (Gillespie, 2014, p. 179). This not

only hides that recommendations can be sold to the highest bidder thus maintaining a degree of agenda-setting, but also lends implicit credibility and legitimacy to the suggested source (Gillespie, 2014). The design of recommendations also tends to obscure that this content is being promoted to users, with up to 62% of British citizens not realising their social networks affect the news they see (Dot Everyone, 2018). When compared with the branding and logos proudly emblazoned on legacy media, the signposting that indicates whether content is being promoted by recommender systems is largely hidden. In this sense, users remain unaware that they are subject to recommender systems, further obscuring the processes by which alternative information sources are excluded from their feeds (Gillespie, 2014, pp. 171-172).

Furthermore, the personalised nature of recommendations means that each user has privileged access to their curated information, remaining ignorant to the information that others are interacting with – a point that will be central to conceptualising how recommender systems affect the possibility for deliberation in 3.5.3 and 3.5.4. Citizens' ability for reflective endorsement, then, is hampered by recommender systems as it limits the diversity of opinions they interact with, providing politically-relevant information that affirms the user's beliefs as opposed to encouraging reflection on the authenticity of their volitions. The repetitive and homogenous nature of these recommendations, framed in neutrality, foreclose an individual's space to authentically reflect on their conceptions of the good, leading to less authentic and considered decision-making in the political context.

3.5.2. *Echo Chambers*

Many of these platforms allow users to freely form and join groups, forums, and instant messaging platforms. Beyond simply recommending personalised news coverage to citizens, recommender systems also routinely promote groups to users. Often these communities are based on any number of shared interests, such as fictional universes, video games, musicians, television, movies, celebrities, sharing memes, and cryptocurrencies. Political action, too, has benefitted from online communities as likeminded individuals can overcome geographical boundaries, expanding their group and developing their political thought (Feenberg, 2016). Recommender systems play a significant role in the formation and growth of these associations as they suggest groups to join and multimedia that other statistically similar individuals have interacted with, thus mediating the original introductions between group members. As such, political groups can expand with ease online, traversing continental boundaries, in part through utilising recommendations and purchasing advertising space.

Through recommended systems, it has never been easier to find like-minded people and engage in communities that share conceptions of the good. Unfortunately, at times, this leads to the creation of online *echo chambers*, as polarised digital spaces in which only one viewpoint or belief is

present (Nguyen, 2018; Flaxman et al., 2016; Quattrociocchi et al. 2016; Reviglio, 2019; Elahi et al., 2022). Often, not only do the groups become intolerant of outside opinion, but the beliefs and discourse within them grow more extreme as the belief goes unchallenged. Moreover, recommender systems have a propensity to suggest progressively more extreme content as it is likely to promote engagement (Waterson & Milmo, 2021; Zubrow, 2021). Where the former dynamic of intolerance may not be an intended consequence, the latter more extreme recommendations are a design specification. This is not to say that platforms necessarily intend to polarise groups and foster extreme views but that it is an unchecked by-product of the economic demands of the attention economy. As the novel and surprising nature of sensationalised news, extreme views, and misinformation is often more engaging than moderate reporting and advertising, it thrives within click-through rate-based recommender systems (Vosoughi et al., 2018 Hendriks Vettehen & Kleemans, 2017; Pennycook et al., 2017). Furthermore, the multinational reach of these platforms means they are simultaneously recommending content to many millions of individuals from an equally vast number of sources. The recommended content often goes unchecked for accuracy or extreme views, typically entrusting screening to algorithms or hoping conscientious users will flag it. This perpetuates the harm to individuals' propensity for reflective endorsement as they are being exposed to greater volumes of homogenous (mis)information. Moreover, as their second-order volitions and conception of the good become based on biased sources and misinformation, they less authentically reflect the real-world political context they believe they are responding to, resulting in stances that are irrational and misguided. As a consequence, the group member's individual influences on political decision-making lead to policies, discourse, and voter behaviour that less authentically represent the population and the political context they are responding to.

Carol's Journey is a significant example of how recommender systems can influence one's group memberships, and thus their freedom of thought and conceptions of the good. Carol Smith was a fake profile created by a Facebook researcher for an experiment in 2019-2020 (Zadronzy, 2021). Whilst signing up, Carol indicated an interest in politics, Christianity, Fox News, and Donald Trump. Despite never demonstrating an interest in conspiracy theories, within two days she was recommended to join groups dedicated to QAnon. Although continuing to ignore this content, her feed quickly became full of information that would be posted and promoted by these groups, such as hate speech, conspiracies, and misinformation. These consistent extreme content recommendations are more than a frustrating inconvenience to one's browsing as it has been demonstrated that social media users are prone to emulating behaviour they believe their online friends to be engaged in (Ashton University, 2020). Currently, then, recommender systems risk expanding divisive and ideologically entrenched groups by repetitively exposing users to extreme content and

misinformation, and thus normalising it. Where Carol never joined these groups, she was still exposed to homogenous (mis)information that could have informed her conceptions of the good and second-order volitions. Given that the content was repeatedly recommended with more diverse sources being simultaneously excluded, there was the possibility that recommender systems could close her space for reflective endorsement and undermine the authenticity of her political decision-making. Not only would her beliefs no longer correspond to the real-world sociopolitical context, but the volume of repetitive information would begin to overwhelm her conceptions of the good as opposed to informing them. This forecloses the possibility of reflection, resulting in less authentic beliefs.

3.5.3. Influence of Recommender Systems on Democratic Deliberation

Recommender systems also affect the possibility of deliberation informing political decision-making. Deliberation understands the justification for political action to be more than the aggregation of votes. Rather, deliberative processes make political decisions more justifiable and acceptable in pluralist societies which share irreconcilable conceptions of the good. This is a process of engaging with citizens to deliberate towards the common good, listening to their beliefs, and outlining how a decision is representative of the population. It requires that participants' interests are shaped through deliberation, meaning one cannot dogmatically defend their position as the process aims to incorporate a plurality of beliefs as opposed to reward the most unrelenting. By including more citizens and collectively shaping peoples' conception of the good, the decision becomes more justifiable. Through deliberation, both the content of the decision becomes more justifiable and the political process becomes more inclusive and tolerant of competing worldviews. In this sense, deliberation is a valuable activity both consequentially and socially.

As recommendations stray into the territory of hypernudging, with dynamic and diverse political content being targeted at increasingly specific groups with equally specific messaging, the shared epistemic ground and reference points on which to base deliberation are undermined. The hyper-curated content disseminated by recommender systems is a threat to deliberation as individuals are locked into filter bubbles and echo chambers before interacting with opposing viewpoints. Instead of engaging with a belief and reflecting on the merits of different conceptions of the good, the homogenous information repetitively diffused by recommender systems creates conditions where individuals are more likely to blindly accept or reject political decision-making, with little interest in reflecting on its reasoning or critical engagement. This is because repetitive exposure to information increases the perceived accuracy that we attribute to it (Pennycook et al., 2018). The information disseminated within these filter bubbles and echo chambers is also personalised, meaning that when citizens interact with people who hold different worldviews, their basic information about

the world is often irreconcilable. For political advertising, designers can take advantage of these ideologically entrenched positions by employing messaging that corresponds to the information that citizens have been repetitively exposed to. This replicates Susser's concerns about online manipulation as the nudging becomes hidden, seeking subconscious agreement through exploiting an individual's psychological characteristics.

Moreover, recommender systems undermine collective deliberation as hyper-curated recommendations mean that participants no longer have access to the same information sources. A necessary requirement for deliberation is exchange between individuals, with this being the opportunity to participate and outline their differing conceptions of the good. A prerequisite for these conceptions is that they are rational beliefs about what is valuable for human life (Rawls, 1993). In political decision-making, such conceptions will often be in response to the current and emerging sociopolitical context, expressing what is currently valuable. These beliefs are informed by information about the world that is, in part, disseminated by recommender systems. As deliberation aims to establish how political decision-making contributes to the common good despite the populace having irreconcilable worldviews, effective deliberation implies that individuals and groups are collectively addressing the same phenomena for their responses to be appropriate. If their information about the sociopolitical context is drastically different, political decision-making will seldom seem justified as the response will be inadequate under their different conceptions of the good and understandings of the world. Recommender systems and their tendency to provide individuals with homogenous sources of information, repeatedly portraying the world as being a particular way, jeopardise effective deliberation as individuals and groups are responding to incommensurable sociopolitical contexts. These biased recommendations, framed as neutral and objective, endanger the shared reference points to which deliberation is responding, compromising the possibility of more justified political decision-making.

Beyond obstructing deliberative processes by endangering shared reference points, recommender systems can provide a means of bypassing deliberation altogether. By providing limited discourse on policies and only publicising their reasoning to select groups, political decision-makers can avoid having their views challenged or facts checked. This is particularly exploited when discussing polarised topics – those in most need of deliberation (Goodman et al., 2019). Instead of encouraging deliberation and outlining how a (proposed) policy contributes to the collective good, recommender systems provide a means of limiting the deliberative space to those who already agree. This limits the diversity of opinion, avoiding the requirements of reasonable pluralism. Through hypernudging, then, deliberation can be selectively bypassed, resulting in less justified political decision-making, particularly regarding the most contentious topics.

3.5.4. Deliberation vs Discourse

Through promoting groups for citizens to join and inadvertently creating echo chambers, recommender systems subtly convince users they have engaged with other values and reflected on their beliefs, when, in fact, they have interacted with people who already agree with them. To demonstrate this difference, I will distinguish between *deliberation* and *discourse*. Deliberation, as outlined in 2.2, is a necessary and ideal process by which political decision-making becomes more justifiable through practices of inclusion, listening to and sharing irreconcilable conceptions of the good, and establishing how a decision is conducive to the collective good despite not being in everyone's interest. Deliberation is necessary for more justified political decision-making when the populace holds reasonable yet irreconcilable worldviews. It is through the process of openly engaging with conceptions of the good and participating that a decision becomes more justifiable (Fishkin, 2009, pp. 76-78). Discourse, on the other hand, should be thought of in terms of discussions and communicative exchanges that are not directed at political justification. In this sense, discourse encapsulates most forms of (online) communication.⁶

Deliberation and discourse differ for several reasons. Deliberation requires engaging with and listening to people with irreconcilable worldviews, and adopting an open attitude whereby participants contribute and recognise there will be differences in opinion (Cohen, 1997a, p. 69). Discourse, however, does not have these procedural or attitudinal requirements and is instead any written or spoken communication. This is not to create a hierarchy between discourse and deliberation as more often than not the former is a sufficient form of interaction. Deliberation is simply required for political decision-making as it is a more open and inclusive form of interaction through which choices become more justified in situations of reasonable pluralism.

Discourse is common to the groups promoted by recommender systems. Within these groups, people converse with each other, often discussing the topic around which their group has been formed. Moreover, given that recommender systems base their recommendations on statistical similarities, the members often share concordant beliefs. The interactions within these groups, then, are often discursive as their similarities predispose them to having amenable viewpoints. Despite feeling as if one has exchanged ideas, listened to others, and reflected on their beliefs, these interactions are not deliberative as they do not include the diversity of opinion or openness that is necessary for deliberation – particularly as echo chambers are hostile to outside beliefs (Nguyen, 2018; Flaxman et al., 2016; Quattrociocchi et al. 2016; Reviglio, 2019; Elahi et al., 2022). This is not to

⁶ In this thesis, I refer to 'discourse' as a form of communicative interaction. This is to differentiate between typical forms of communication (discourse) and forms of exchange necessary for justified political action (deliberation). My use of 'discourse' is unrelated to Foucault's (1969, 1982) concept and analysis of discourse.

say that the groups recommender systems promote are meant to be deliberative spaces but that the exchanges between citizens with ideologically entrenched worldviews give a false sense of having reflected on one's opinion and interacted with a plurality of beliefs. Given the fact of reasonable pluralism, there is an awareness that citizens should not purposely limit their interactions to those who share the same worldview as this would be a limited part of the populace. The communication facilitated by recommender systems creates a false sense of having communicated with a diverse range of people as it introduces them to many other users, users who consistently reaffirm each other's beliefs. Consequently, there is not an open exchange of conceptions of the good but the reinforcement of their beliefs, with group members becoming steadily more estranged from the diversity of opinion. The groups disseminated by recommender systems provide an illusion of having openly interacted with a plurality of beliefs, giving a false sense of deliberation.

3.6. Conclusion

In this chapter, I have outlined recommender systems as a form of nudge-based behavioural influence. I then established how recommender systems influence behaviour in the context of political decision-making. This included the benefits of recommender systems, such as distributing information and effectively expanding (political) groups. I then outlined how recommender systems make political decision-making less authentic, before establishing how they can obstruct and bypass deliberation. In chapter 4, I identify solutions, focusing on how more authentic and justifiable political decision-making can be encouraged.

4. Solutions

In chapter 2 of this work, I outlined reflective endorsement as necessary criteria for more authentic political decision-making. Additionally, I argued that deliberation was necessary for more justifiable and acceptable political decision-making. In chapter 3, I explored how recommender systems go beyond traditional nudge-based forms of behavioural influence, demonstrating how this affects reflective endorsement and deliberation – thus affecting political decision-making. I argued that the economic incentives of the attention economy are not conducive to political decision-making within liberal democracies, specifying how maximised engagement inadvertently contributes to less authentic and justified political decisions. In this chapter, I begin to look at possible solutions for limiting the undesirable effects of recommender systems. First, I outline recommender systems as sociotechnical systems that are interacting with the valuable context of political decision-making, therefore establishing that any range of possible solutions needs an equally sociotechnical focus. I then approach reflective endorsement and deliberation, in turn, conceptualising what changes to recommender systems and the social context could foster more authentic and justified political decision-making.

4.1. Sociotechnical Solutions

This paper's critique of recommender systems is not directed at the technology alone. Instead, the focus has been on the relationship between recommender systems and the social context of political decision-making, arguing that the economic incentives of these systems are not conducive to liberal democracies. As such, the focus of this critique is understanding how recommender systems, as partially technical systems, interact and influence the partially social context of political decision-making. To reflect the sociotechnical nature of the proposed problems, the solutions need to be equally sociotechnical.

Adopting a sociotechnical perspective recognises the interrelation and co-evolution between social and technical aspects (Walker et al., 2008, p. 480; Geels, 2005, pp. 364-366). In other words, when technological artefacts (as a range of technologies, infrastructures, and techniques) interact with people, groups, and societies, they influence and shape one another, sometimes in unexpected ways. Employing a sociotechnical perspective includes both learning about these interrelations and deriving methods to optimise the relationship (Walker et al., 2008, p. 480). A *sociotechnical system* can be any practical instantiation of social and technical elements engaged in goal-oriented behaviour (Walker et al., 2008, p. 480). Examples and studies of sociotechnical systems include land-based transportation (Geels, 2005), water systems (Fuenfschilling & Truffer, 2014; Geels, 2005; Konrad et al., 2007), and electricity systems (Verborg & Geels, 2010). These systems are comprised of networks

of technologies, organised social groups, normalised routines and behaviours, legislation, (global) infrastructures, maintenance, and the cultural meanings ascribed to each of these practices (Geels, 2005, p. 365). Cultural meanings go beyond the uses of technology and incorporate the values ascribed to the sociotechnical system. For instance, roads are more than just extensive networks of asphalt, they are imbued with connotations of freedom and possibility (Larkin, 2013). Such sociotechnical systems are seldom static, with new technologies, regulations, social meanings, and behaviours emerging and evolving at the individual, societal, and global level (Larkin, 2013).

Sociotechnical systems, as a classification, have limits. It is an approach for studying the relationship between expansive networks of technologies and society on a broader scale, as opposed to personal relationships. Geel's *multi-level perspective* (2002, 2005, 2012) incorporates studying smaller groups of actors at the *niche level*, but this is typically with reference to how these emergent technologies are designed in response to specific societal problems, eventually influencing and innovating sociotechnical systems more broadly. As such, the relationships between an individual and the technologies they interact with are part of a sociotechnical system, as opposed to an individual system itself. The solutions proposed, then, are broad recommendations that seek to optimise the sociotechnical relationship between recommender systems and political decision-making. Consequently, they may not apply to each individual's relationship with recommender systems.

Recommender systems are inherently sociotechnical because they are comprised of both people and technologies (Ananny, 2015; Bozdag, 2013; Nissenbaum, 2015; Mühlhoff, 2021). Although recommender systems are a form of big data analytics that infers preferences to personalise online experiences, they are systems created by people for use by the public. Moreover, the attention economy model that motivates the design of recommender systems was established by people, responding to a socioeconomic context, and is thus embedded with social values related to maximising profit and growth. The nudge-based techniques embedded within recommender systems also have concrete social effects on the lives of individuals and communities, such as how they spend their time and money, whom they interact with, and the information that informs their conceptions of the good. Recommender systems, then, are inextricably linked with social values, behaviour, cultural meanings, and beliefs.

Political decision-making can also be sociotechnical. The information that contributes to peoples' political beliefs and conceptions of the good is partially diffused by technologies, such as recommender systems. Political campaigns, discussions, and movements are all mediated by technologies, informing not only what people see but how they can interact with it. Individuals and groups can create online polls and petitions to pressure incumbent governments, using the sharing

capabilities of social media sites to spread their message across the nation and globe. Where political decision-making has inherently social elements – such as collective action – technological aspects have transformed the process of engaging politically. Consequently, political decision-making can also be sociotechnical.

4.1.1. The Shortcomings of Wholly Technical Solutions

Where recommender systems are designed by developers within organisations, their technological form is seldom final. The big data analytics underpinning recommender systems constantly analyses users' and the population's behaviour, altering the workings of these systems to more effectively respond to observed trends (Yeung, 2016). In this sense, recommender systems are *adaptive algorithms* that consistently tweak their workings towards the end of maximised engagement and click-through rates (Bhargava & Velasquez, 2020; Chesson, 2018; Schou & Farkas, 2016). In addition to not sufficiently recognising the social context of political decision-making, incorporating merely technical solutions would risk adaptive algorithms gradually removing changes in favour of tweaks that promote engagement. For effective solutions that promote authentic and justifiable political decision-making, reflective endorsement and deliberation need to be incorporated into the socioeconomic priorities which are embedded in adaptive algorithms. This will give greater assurance that the undesirable effects of recommender systems on political decision-making do not re-emerge further down the line.

4.2. Addressing Recommender Systems

Given that both recommender systems and the context of political decision-making are sociotechnical, this chapter provides sociotechnical solutions that reduce the undesirable effects of recommendations on liberal democracies. This includes which technical fixes could be introduced but also what attitudes, behaviours, and deliberative processes should be fostered to make political decision-making more authentic and justifiable.

A central aspect of these solutions is they encourage behaviour that results in more authentic and justifiable political decision-making. This does not intend to be paternalistic or impose strict limits on peoples' actions but encourage behaviour that is conducive to authentic and justified political decision-making. This takes inspiration from libertarian paternalism, as the solutions attempt to improve people's experiences within the context of political decision-making without foreclosing their range of possible actions (Thaler & Sunstein, 2003, 2008: 4-6; Sunstein & Thaler, 2003). Moreover, considering Mitchell's (2004) criticism of libertarian paternalism (see 3.1) and the importance of authentic decision-making, these solutions seek to promote reflection about one's choices as opposed to merely creating choice architectures that redirect someone's subconscious decisions to alternative,

more paternalistic ends. The aim of these solutions, then, is not simply to introduce competing forms of unconscious influence, but to create conditions that encourage the possibility of reflection on one's desires and volitions.

4.3. Reflective Endorsement

Typically, recommender systems suggest content that a user has already demonstrated an interest in, providing news and politically-relevant multimedia that echoes their beliefs. This creates filter bubbles and limits the diversity of opinion available for a user, implicitly lending credibility to the recommended content and hiding the diverse range of other information sources. Moreover, through promoting groups, recommender systems introduce users to others with similar interests, sometimes creating echo chambers, as online spaces in which one viewpoint is constantly repeated. As a result, users' pre-existing beliefs and conceptions of the good are consistently reinforced, limiting their ability to reflect on their conceptions of the good or volitions. The root of this problem, then, is personalisation and how it intersects with authentic political decision-making (Susser et al., 2019, p. 12).

4.3.1. Diversity of Information

One possible solution for aiding freedom of thought is designing novelty and serendipity into recommender systems (Reviglio, 2017; Elahi et al., 2022, p. 109). Serendipity, as a design feature, is a pre-existing area of research, with studies demonstrating its possibility (Campos & Figueiredo, 2002; Niu et al., 2018). Increased serendipity would result in users receiving surprising and diverse information among their typical recommendations, potentially introducing them to alternative viewpoints and news that covers a variety of issues. This information does not have to be opposed to their beliefs and interests, but simply disclose to them reliable and accurate information to facilitate reflection. Particularly as this technical solution is in response to the social context of political decision-making, it merely requires that politically-relevant content be given special consideration. As such, this does not request social media sites to forego personalisation altogether, but to acknowledge the unique nature of the political context and the need for people to reflect on their beliefs for authentic decision-making. For example, in their settings, users could be given a serendipity slider that allows them to choose between content personalisation, generalisation, and randomisation (Reviglio, 2019, p. 158). Nudges, too, could be introduced that remind users of the benefits of diverse information and indicate the possible negative effects associated with hyper-personalisation. From a libertarian paternalist perspective, the default setting would be 'generalisation' meaning users retain the choice to have homogenous recommended content, but disincentivise this option by making it opt-in. However, considering Mitchell's (2004) critique of choice architectures being overly directive towards

paternalistic ends, it is important to consider whether this default would be too imposing. Given the alternatives of (a) 'personalisation' locking individuals into filter bubbles, (b) 'randomisation' making the benefits of recommender systems effectively redundant, and (c) providing no default option and thus imposing a choice on users, possibly resulting in them picking an option at random which may not fulfil their preferences, 'generalisation' is the most liberty-preserving default selection. Not only would users remain free to change this, but this option provides the convenience and freedom of personalised content whilst not deluging them with homogenous content and effectively closing off alternative options.

Introducing more diversity to recommendations also requires social changes. Where sites that employ recommender systems have partially recognised the influence that information has over one's political decision-making, they need to adopt a broader understanding of this process. This requires understanding political decision-making as an ongoing and recurrent social process that extends beyond termly elections. For instance, Facebook (2020) updated their community standards and advertising policies to protect the integrity of elections, reduce voter interference, combat hate speech, and promote more transparent political advertising. Among these policies were a ban on new political advertisements between October 27th and November 3rd (the week before the 2020 US Presidential Election); a strict ban on advertisements that denigrate migrants, immigrants, asylum seekers, and refugees; removal of content that promotes or threatens violence during and after the election; and removing or labelling posts that misrepresent the official election results (Facebook, 2020). Where these policies attempt to limit divisiveness and recognise the influence that misinformation can have over citizens' voting activity, they are primarily directed at official candidates and their behaviour in the build-up and aftermath of the democratic event. The guidelines do not extend to information routinely promoted by recommender systems in the years between elections, a period in which citizens' beliefs are also influenced. Socially, then, sites that use recommender systems to disseminate politically-relevant information need to adopt a more holistic understanding of political decision-making, recognising how preserving authentic decision-making and democratic integrity is not just a termly activity, but a constant process. One that is influenced by recommender systems insofar as they diffuse politically-relevant information to individuals, informing and, at times, limiting their reflective endorsement.

The second social change requires acknowledging the importance of political decision-making as a valuable social activity, one informed and inadvertently undermined by the homogenous information endemic to recommendations. In chapter 2, I outlined the vital role that freedom of thought and information play in authentic political decision-making, helping individuals form and reflect on their conceptions of the good. As recommender systems personalise a user's information

sources, the diversity is reduced, repeatedly recommending information that agrees with their beliefs and closing off the space for reflection. Consequently, recommender systems limit individuals' space for freedom of thought by reinforcing their opinion with homogenous information, at times locking them into filter bubbles (Pariser, 2012). If we accept freedom of thought as necessary for reflective endorsement, then homogenous recommendations that reinforce one's beliefs threaten this process. A precondition for designing novelty and diversity into recommender systems, then, is accepting the positive effect of these values on authentic political decision-making.

Thirdly, and most importantly, these changes require prioritising conditions that are conducive to authentic political decision-making at the expense of maximised engagement and click-through rates. This would require platforms actively recognising and vowing to change their systems, potentially at the expense of their bottom line. Changes, such as incorporating more novelty to the recommendations, require both technical alterations to the goals that adaptive algorithms are internally directed towards and social commitments to facilitate reflective endorsement. This may be in direct conflict with the economic imperatives of the attention economy as concordant information, misinformation, sensationalised news, and conspiracy theories generate greater engagement and revenue (Vosoughi et al., 2018; Hendriks Vettehen & Kleemans, 2017).

Where implementing these solutions may seem unlikely, it is not altogether impossible. The sociotechnical changes are proposed specifically in response to the context of political decision-making, meaning any restrictive requirements for recommender systems would only apply to politically-relevant content, such as news, opinion pieces, and multimedia produced by political figures, commentators, and journalists. Comparatively, recommending forms of entertainment or multimedia produced by content creators – such as influencers, fitness instructors, celebrities etcetera – would not necessarily have to commit to more diverse recommendations as their influence on political decision-making would be less significant.⁷ Moreover, these proposed solutions do not require platforms to abandon profits but merely give social concerns a greater role. The introduction of similar solutions and organisational commitments to address social issues has been demonstrably successful before. For instance, in 2019, when Twitter introduced a global ban on paid political advertising, following pressure from citizens and political decision-makers (BBC News, 2019). Then-

⁷ It is important to recognise that entertainment and politics are often not distinct. The global audiences of many celebrities extend far beyond that which most political figures can dream of. In 2021, for example, Nicki Minaj tweeted an anecdote to over twenty million followers about her cousin refusing COVID-19 vaccinations as it had caused his friend to become impotent (BBC News, 2021). As vaccine hesitancy became heavily politicised during the COVID-19 pandemic, Minaj's tweet would have fallen into the politically-relevant category and therefore subject to these more restrictive limitations. Whether the content of traditionally non-political figures should face limits to the extent it can be promoted by recommender systems is thus dependent on the content of the message as opposed to the profession of the author.

CEO, Jack Dorsey (2019), argued that paying for the reach of your political message has significant ramifications for democracy, prioritising social value over potential revenue. Similarly to the recommendations of this chapter, the ban did not extend to all advertisements but simply those that affect political decision-making.

4.4. *Deliberation in Democracy*

In chapter 2, I argued that deliberation is a necessary requirement for more justifiable political decision-making in a pluralist democracy. By including the populace in deliberative practices, decision-makers engage with citizens, listening to their conceptions of the good, and outline how their policy contributes to the common good despite being irreconcilable with some worldviews. This process shapes peoples' conceptions of the good, becoming more justifiable through their inclusion and recognising why the decision is in the interest of the collective. In chapter 3, I outlined that recommender systems undermine and obstruct the possibility of deliberation. As recommender systems lock people into filter bubbles, they are less able to interact with opposing viewpoints, thus undermining deliberation. Moreover, as argued in 3.5.4, the discourse between group members within echo chambers creates the illusion that users have engaged in deliberation about their beliefs and conceptions of the good, when in reality they have discussed effectively similar beliefs. Finally, the hyper-curated nature of recommendations can subvert deliberation altogether, allowing candidates to attempt to limit discourse about policies to those who already agree with them, preventing fact-checking or being challenged by opposing views. This section responds to each of these issues, providing solutions to reopen the possibility of deliberation.

4.4.1. *vTaiwan*

To consider solutions that encourage deliberation, I will introduce *vTaiwan*, a collection of open-source tools used by the Taiwanese government to aid decision-making, promote consensus, and encourage participation in matters concerning the digital economy (*vTaiwan*, n.d.; Tang, 2019; Horton, 2020). Where *vTaiwan* is neither a recommender system nor a platform created to directly respond to the undesirable aspects of recommender systems, it is designed to mitigate the harms of political division and encourage deliberation and consensus amongst citizens.

One aspect of *vTaiwan* is 'Pol.is', a platform where topics are proposed for debate. Taiwanese citizens can register, voice their opinion, and upvote and downvote other comments. Pol.is allows decision-makers to see clusters of popular opinions, revealing areas of consensus and contention. An important feature is that users cannot reply to comments meaning there is less incentive to troll or write divisive statements. Furthermore, it gamifies the process, encouraging people to draft comments that overcome differences of opinion, thus narrowing down the most central issues for the

public and making repetitive messages with unyielding values pointless (Tang, 2019; Horton, 2020). Once crucial values and points of consensus emerge, key stakeholders, experts, and political decision-makers enter a consultation stage that reflects on the deliberation and responds to the public's questions. The consultation is streamed live on YouTube and Livehouse.in. This outlines rough areas of consensus which is transformed into a draft policy and presented to Parliament. Finally, the government either accept or reject the proposal, providing sufficient reasoning. Since 2016, the Taiwanese government has executed over 80% of vTaiwan's proposals, including regulations for UberX (Tang, 2019; Horton, 2020).

vTaiwan is a significant example of technologies designed to promote deliberation, thus making political decision-making more justifiable. Pol.is is the foundation for open deliberation, encouraging the public to share their conceptions of the good whilst simultaneously discouraging divisiveness, trolling, and entrenched beliefs. The focus on consensus, as opposed to maximised engagement, encourages users to be congenial and open to different ideas. As an official platform, vTaiwan becomes the destination associated with deliberation, thus countering the false sense of deliberation endemic to online group membership. Moreover, vTaiwan supports reflective endorsement by broadcasting a reflection stage with key stakeholders and industry experts which overcomes the spread of misinformation, clearly establishing the current political context. This provides the foundations for more authentic political decision-making as individuals' conceptions of the good and volitions can accurately respond to the current sociopolitical context.

vTaiwan, however, is not without its faults. As citizens show their approval and disapproval of statements, they create a record of their beliefs, thus sparking fears of possible misuse by authoritarian governments. For instance, between 1956-57, Chairman Mao Zedong launched *The Hundred Flowers Campaign* which encouraged citizens to openly express their opinions of the Chinese Communist Party (CCP) (Encyclopaedia Britannica, n.d.). As criticism of the CCP grew, the state cracked down on critics, with many losing their jobs, assigned forced labour, and sent to prison camps. To prevent digitally-mediated deliberation from becoming a punitive force, significant privacy and anonymity safeguards need to be built into the design.

Moreover, the digital divide means that this deliberative space is closed off to people with poor technological literacy or lack the means to participate. Consequently, the popularity of particular beliefs may be misrepresented through higher proportions of wealthy and technologically-enabled participants. The diversity of topics proposed may also be limited, reflecting the values and priorities of those with the means to participate. To ensure deliberation is open and incorporates pluralist conceptions of the good, the state must ensure that the technology and training necessary to

participate are affordable and accessible. Despite these drawbacks, vTaiwan is a promising example of how deliberation can be facilitated online. vTaiwan, then, can act as a form of inspiration for solutions and simultaneously validate their feasibility and effectiveness.

It is important to clarify the differences between recommender systems and vTaiwan. The latter has been designed precisely to aid political decision-making, whereas the former aims to promote new content and engage user attention through any form of online multimedia. Consequently, facilitating deliberation and directly contributing to political decision-making is simply not the purpose of recommender systems. They are, however, part of a network of technologies particularly common to social media, on platforms described as akin to the town square (Zuckerberg, 2019). As a traditional use for such spaces is as an arena of political discussion, the question becomes what inspiration can be taken from vTaiwan to reopen the space for deliberation and reduce the divisiveness of interactions. This is not to divorce recommender systems from their ultimate purpose of curating personalised content but to incorporate solutions that reflect both the specific aim of recommender systems and the many uses of the sites recommending content.

4.4.2. Facilitating More Open Discourse

This section outlines possible recommendations to facilitate more open online discourse. It recognises that deliberation is dependent on the open exchange of worldviews between individuals, meaning the reduction of divisive discourse is a precondition for increasing the possibility of deliberation. These solutions include technical changes to recommender systems, such as introducing more serendipity to promoted content and nudges that slow down peoples' response time and encourage reflection on the content of their response. Particularly, these nudges attempt to make online exchanges more open. The purpose is not for unanimous agreement but to foster conditions in which a diversity of opinions can be freely exchanged. This section, then, demonstrates that the divisive state of online discourse is not necessarily a lost cause. Instead, through sociotechnical changes, facilitating more open discourse is a feasible end.

A simple sociotechnical solution, and one already outlined in 4.3.1, is introducing more novelty and serendipity to recommender systems (Reviglio, 2017; Elahi et al., 2022). This would lead to users interacting with a more diverse range of accurate information sources, providing them with the means to form and interrogate their conceptions of the good. Consequently, when exchanging with others, users will be predisposed to interacting with varied information sources, may have more balanced worldviews, and be more accepting of others' beliefs. However, for open deliberation and the exchange of ideas to become increasingly likely, additional changes are required to disincentivise trolling and intentionally divisive statements. Where vTaiwan does not allow users to reply to each

other and disincentivises comments that do not contribute to the discussion, recommended content typically allows any and all responses.⁸ Of course, each of these sites has community guidelines about hate speech, threats, and acceptable interactions, one look into the comments section of a politically contentious topic suggests they are seldom enforced. Analogously to Susser's approval of nutritional information on food labels (Susser et al., 2019), social media sites could encourage reflection over a user's comments by asking if they really want to use a phrase that has been flagged as potentially harmful or contains terms they have repetitively used. This technological solution aims to encourage reflection that fosters less divisive behaviour.

Similarly, to reduce divisive and hateful comments that obstruct deliberation and the open exchange of ideas, nudges could be employed that ask a user to consider whether their comment is helpful or necessary, thus slowing down their response. Such nudges could remind users that they are interacting with another person, questioning whether they would use similar phrases face-to-face. Recently, Twitter has introduced a similar prompt, asking users to review their potentially harmful and offensive language before submitting, leading to people changing or deleting their responses over 30% of the time (Twitter Safety, 2022; Twitter Engineering, 2022). This demonstrates that the more widespread introduction of nudges to disincentivise trolling is not only feasible but effective. Moreover, these nudges satisfy Mitchell's critique of libertarian paternalism being overly paternalistic as they do not limit an individual's freedom but merely offer a moment to reflect on one's actions and make a more conscious choice. When applied to reflective endorsement, these nudges are akin to prompting a user to have a second-order desire regarding their comment's content (Frankfurt, 1971). If a user's volition is to use that phrasing, they remain free to do so. However, others users may be unable to endorse this action and thus refrain from doing so.

The aim of these nudges is not to prevent honest interactions but encourage users to reflect on the content of their message, slowing down their response time and questioning whether they are being open, helpful, and respecting community guidelines. Where this would not necessarily facilitate deliberation between citizens, it is a technical solution that reopens the possibility of it occurring by fostering conditions that encourage the open exchange of conceptions of the good. The aim of disincentivising divisiveness and trolling, then, is to create conditions that reopen the possibility of deliberation and thus more justified decision-making, as opposed to ensuring that it happens.

⁸ Where Twitter has tried to limit abuse and the possibility of coordinated attacks by allowing users to limit who can respond to their tweets (Peters, 2021), it does not prevent thousands of people from attacking them in the 'quote tweets'.

4.4.2.1. The Social Context of Political Decision-Making

In the current political context, however, nudge-based solutions that encourage reflection on the content of one's comments would be insufficient. Additionally, there would need to be a significant cultural shift regarding the norms of political engagement and the expectations of exchanges and responses to different conceptions of the good. Originally, this would require strong leadership from all sides of the political spectrum (Ryfe, 2005, p. 63). So long as there are candidates constructing culture wars, spreading misinformation, and encouraging divisive and malicious interactions, their supporters will do the same. Furthermore, given the recent successes of breeding cultural division, particularly among right-wing and populist candidates, it is hard to see why political parties would be willing to adopt more moderate discourse. Without a unanimous commitment to encourage open deliberation and renounce ideologically entrenched discourse, it is difficult to see how online interactions will become more open and deliberative. The situation, however, is not necessarily futile. Rather than waiting hopefully for political candidates to commit to more open deliberation, I propose restrictions on the content recommender systems can promote. Namely, I argue that by imposing limitations on the personalisation of recommendations, political candidates will be less able to tailor their messages to particular audiences, thus reopening the possibility of deliberation and fostering a less divisive social context.

4.4.2.2. Targeted Political Advertising

Recommender systems allow for political candidates to target their campaign advertisements to curated groups of individuals, typically those who already share these opinions. Although non-targeted groups who endeavour to look can still find these advertisements, recommender systems facilitate attempts to close off the deliberative space by preventing outside opinion and the exchange of ideas, particularly regarding the most politically contentious topics (Goodman, 2017, p. 19).

One technical solution would be to create an online repository that documents every variation of each candidate's political advertising, thus theoretically providing access to those omitted by recommender systems. However, given that during Trump's 2016 Presidential Campaign there were up to 50000 variations of adverts run each day (Anderson & Hovath, 2017), this database would become so overwhelmed with advertisements that it would be effectively useless. Instead, I propose a more substantial solution – that online political advertising cannot be targeted to specific groups. This would still allow for advertisements to be updated following the audience's reaction, but there would be restrictions on whether different groups can receive personalised messaging. As such, this solution requires a social commitment to change each platform's conditions for political advertising, whilst still allowing campaigners to utilise the technical benefits of recommender systems and big data

analytics. Much like political advertising on billboards, flyers, television, radio, and newspapers, the message would be universal to those exposed to it. This does not mean that candidates and parties could not tailor advertisements to appeal to their core voters, but that they would be unable to use recommender systems to simultaneously disseminate multiple different personalised campaigns that other groups may not have access to. This would facilitate deliberation between political decision-makers and citizens as there would be more stable knowledge of their intentions, making it easier to challenge their beliefs and hold their advertisements to account. This would prevent employing recommender systems as a means of obstructing deliberation and interacting with different conceptions of the good. A further benefit is that these restrictions would facilitate deliberation between citizens as they would be interacting with more stable knowledge about the political context, as opposed to discussing information and policies that others have not had access to. On the individual level, too, this solution would encourage more authentic decision-making as citizens would have equal access to a candidate's policy intentions and thus be able to reflect on whether they are aligned with the conception of the good they wish to pursue. By limiting the use of recommender systems to disseminate personalised political messaging, deliberation becomes more possible as candidates cannot as easily conceal their political intentions, resulting in a greater likelihood of having to justify their beliefs.

I concede that platforms will find this solution to be undesirable for economic reasons. Although Twitter introduced an outright ban on political advertising (Dorsey, 2019), it is still commonplace on platforms such as Facebook, Instagram, YouTube, and Google. It is simply not in the interests of sites that employ recommender systems to limit the sophistication of personalisation. Facebook state their reasons for offering political advertisements are not financial as it makes up less than 1% of its revenue (\$2.2 billion USD from mid-2018 to October 2020) (Canales, 2020) and instead willingly host these advertisements so as not to limit political expression (BBC News, 2019). Primarily, however, the motivation for hosting political advertising is because it is a substantial source of behavioural data which can be applied to make recommender systems more sophisticated (*à la* Yeung's (2016) hypernudging) and to maximise user engagement (Zuboff, 2019, pp. 130-133). The process of recommending content generates vast amounts of behavioural data regarding peoples' interests, beliefs, and motivations, all of which would be sacrificed were the proposed limitations to be adopted.

Looking beyond economic incentives, there are significant moral and political reasons for platforms to adopt more restrictive political advertising policies. As employing recommender systems to disseminate targeted political advertising can allow for candidates to exploit voters' cognitive shortcuts, gaining support while not offering reasoning for their policies or prompting reflection on

their volitions and whether this aligns with individuals' conceptions of the good, they undermine the authenticity of voter behaviour and obstruct deliberation. As reflective endorsement and deliberation are necessary criteria for authentic and justifiable political decision-making, recommender systems have the potential to undermine democratic processes. In chapter 2, I outlined the value of democracy as a system that recognises and protects the equality of each citizen. Consequently, there are serious moral and political reasons to impose restrictions on the targeted political advertising curated by recommender systems. If left unchecked, less authentic and justified political decision-making could compromise democratic processes, the processes by which equality is assured. Platforms, then, have a political and moral responsibility to preserve and promote democracy. Though they may still be reluctant to implement these changes, the bitter pill of lost behavioural data may potentially be sweetened by knowledge that their competitors face similar restrictions meaning they have not lost their advantage.

4.4.3. Recognise Lack of Deliberative Space

As argued in chapter 3, the groups promoted by recommender systems connect individuals who share statistically similar interests. At times, these groups become echo chambers, as spaces in which only one opinion is discussed, consistently becoming more extreme and intolerant of outside beliefs. Furthermore, *Carol's Journey* demonstrates that recommender systems inadvertently promote more extreme and divisive groups as the misinformation and conspiracies discussed are more likely to generate engagement and thus higher click-through rates. A significant consequence of these homogenous echo chambers is that individuals are given a false sense of having reflected on their opinion and deliberated with others. Instead, they engage in discourse where conceptions of the good remain unproblematised, consistently agreeing with each other instead. The proposed solutions highlight the lack of deliberation and encourage reflection on whether users are critically engaging with their beliefs. This looks to create the possibility for deliberation, whilst not imposing it on individuals.

It would be misguided to propose solutions which make the groups promoted by recommender systems more deliberative as, ultimately, they have never been a space for deliberation. Instead, their general purpose is to help users find other like-minded members so they can form communities based on their shared interests. That these groups have come to host echo chambers and misinformation is an unintended consequence of the attention economy. Alternatively, to limit the illusion of engaging in deliberation within these groups, new platforms should be created which facilitate deliberation between citizens. Such platforms could be similar to vTaiwan, becoming the recognised destination for deliberation and the exchange of ideas, with common knowledge that

the products of this have concrete effects on political decision-making. From a sociotechnical perspective, this would influence the cultural meanings ascribed to online groups. By creating dedicated deliberative spaces that inform political decision-making, the groups promoted by recommender systems would become less associated as a destination for exchanging political values and conceptions of the good. Instead, the discursive groups promoted by recommender systems would remain, being a location of comparatively little political consequence, whilst deliberative spaces would host the open exchange of conceptions of the good and inform more justified political decision-making. This recommendation is based on a sociotechnical perspective that recognises the disruptive effect that new technological capabilities can have on social dynamics (Geels, 2005, p. 366). Further sociotechnical systems research into deliberative platforms could seek to explore this possibility, highlighting how current services only facilitate discourse and obstruct deliberation.

This would not impose limits on recommender systems and the groups they promote, but change the uses and cultural meanings associated with them. For example, where it would still be possible for taxi drivers to discuss Uber within their online groups, there would be an awareness that their exchanges could inform policy if they were to contribute to the dedicated deliberative platform, as was the case with vTaiwan (Tang, 2019; Horton 2020). This would disincentivise the repetitive discourse of echo chambers and redirect citizens' attention to engaging with others towards practical solutions. This contributes to more justified political decision-making precisely because citizens have the opportunity to voice their conceptions of the good and witness it inform policy. Consequently, researching and developing deliberative platforms that encourage participation and openly exchanging ideas could change the cultural meanings currently ascribed to online groups. Where the groups promoted by recommender systems would benefit from less divisive and confrontational discourse, creating dedicated deliberative spaces would reopen the possibility of deliberation, resulting in more justifiable decision-making.

Within this further research, responses to the challenges associated with vTaiwan could also be investigated – namely, the political concerns of having your belief associated with an online profile and the digital divide limiting participation. Of the former challenge, more stringent data protection controls could limit the biographical information the state has access to, but this may come at the expense of knowing that all those participating are citizens. Whereas for the latter, ensuring that any deliberative platform requires minimal technological capabilities and know-how would be an important start. Moreover, research into digital deliberative platforms should not be at the expense of non-digital methods of participating. Where accessibility to digital resources should be increased, it should not be imposed upon citizens. As part of this further sociotechnical research, it should explore

how the introduction of deliberative platforms could foster more deliberation in both digital and non-digital settings.

Rather than specific sociotechnical changes to recommender systems and the groups they promote, I have argued for new platforms that could reopen the deliberative space. Taking inspiration from vTaiwan, the creation of these platforms would (a) alter the cultural meaning associated with online groups, rightly making them more akin to spaces of discourse, and (b) incentivise deliberation through experiencing public contributions inform political decision-making. Although recommender systems are undermining deliberation by creating echo chambers that give a false sense of having deliberated and providing a means to limit deliberation, these external sociotechnical solutions can mitigate these harms and reopen the possibility of deliberation.

4.5. Conclusion

This chapter has provided sociotechnical solutions in response to the undesirable effects of recommender systems on political decision-making. It has proposed introducing serendipity for politically-relevant recommendations – such as advertisements, news, and blogs. Accompanying this technical solution were social changes which acknowledge the effect information has on political decision-making, requiring that social factors are given greater consideration in the political context. Taking inspiration from vTaiwan, I proposed nudges to disincentivise trolling and intentionally divisive statements, thus facilitating less combative online discourse. I conceded that the effects of disincentivising divisive discourse are unlikely to be effective if there is not a significant cultural shift in the norms of exchange from political leaders. To encourage more deliberation, or at least reduce the possibility of limiting access to one's policy intentions, I propose restrictions on targeted advertising. Particularly, I argue that political candidates should not be able to use recommender systems to promote advertisements with personalised messaging. By creating conditions where citizens have universal access to advertising, it will be easier to challenge a candidate's views and engage in deliberation. I recognise that this solution would face resistance from platforms that employ recommender systems as restricting targeted political advertising would sacrifice vast amounts of behavioural data, thus limiting the effectiveness of recommender systems. Ultimately, I argue the moral and political harms associated with undermining democratic processes are so great that platforms have a responsibility to implement these changes. Finally, to change the cultural meanings associated with the groups promoted by recommender systems, I argued for further research into creating dedicated deliberative platforms, similar to vTaiwan. I outline that this sociotechnical research should focus on creating platforms that encourage deliberation, thus revealing social media groups as locations of discourse. This would not impose deliberation on citizens but provide a means

of exchanging pluralist conceptions of the good away from the filter bubbles and echo chambers endemic to recommender systems.

5. Final Thoughts

5.1. Overview

This thesis has explored the relationship between recommender systems and political decision-making. Particularly, it has asked whether recommender systems, in their current form, facilitate authentic and justified political decision-making. I explored this question with reference to the effects of recommender systems on reflective endorsement and deliberation.

In chapter 1, I outlined recommender systems as a highly profitable network of technologies, ubiquitous in digital spaces. I explained that recommendations are generated through big data analytics and based on users' behavioural data in conjunction with the data of thousands of other similar users. I argued that recommender systems require philosophical consideration as they structure our relationship with valuable activities, such as the arts, how we spend our leisure time, news, information, etcetera. Moreover, precisely because recommender systems attempt to influence a user's behaviour and political beliefs, it is important to consider the extent of this influence, particularly given the importance of the political context.

In chapter 2, I focused on the context of political decision-making. This chapter explored political decision-making within liberal democracies, outlining that democracy ensures a degree of equality amongst citizens. Here, I explored two necessary criteria for authentic and justifiable political decision-making. First, I outlined reflective endorsement as a requirement for authentic political decision-making. This requires reflecting on and endorsing one's volitions and conception of the good, a process that is dependent on freedom of thought and access to reliable, accurate information. Secondly, taking inspiration from deliberative democracy, I introduced deliberation as necessary criteria for more justified political decision-making. Through deliberation, individuals share their conceptions of the good and engage in the open exchange of ideas, witnessing how their beliefs influence the final decision. Where this decision may be irreconcilable with their conception of the good, through participation citizens begin to recognise how it contributes to the common good.

Chapter 3 began with a deeper exploration of the behavioural influence exhibited by recommender systems, grounding this in nudging (Thaler & Sunstein, 2003, 2008; Sunstein & Thaler, 2003). As big data analytics informs recommender systems, their influence goes beyond traditional nudging techniques, at times becoming *hypernudges* (Yeung, 2016) or forms of online manipulation (Susser et al., 2019). I then assessed the relationship between recommender systems, as an advanced form of behavioural influence, and the context of political decision-making. I argued that the homogenous nature of recommendations limits an individual's propensity for free thought by locking them into filter bubbles and echo chambers, thus undermining the authenticity of their beliefs. I then

demonstrated that the groups promoted by recommender systems and the communication within them are often divisive, inadvertently prioritising the dissemination of misinformation and conspiracies.

Furthermore, as recommender systems seek statistically similar individuals with amenable interests and beliefs, the conversations within groups are typically discursive instead of deliberative. Given that similar political beliefs are repetitively exchanged, individuals develop the illusion of having deliberated and listened to outside opinion when in reality they have reinforced each other's beliefs. Finally, recommender systems provide a means of attempting to close off the deliberative space through disseminating targeted advertising to those who have been identified to agree, thus reducing the possibility of having to justify one's position or engaging with different conceptions of the good. Ultimately, I argue that undermining reflective endorsement and deliberation are unintended consequences of the economic demands of the attention economy. It is not that the platforms employing recommender systems are actively trying to affect political decision-making, but that pursuing maximised engagement does not encourage reflective thought and the open exchange of ideas.

Chapter 4 offered sociotechnical solutions to the undesirable effects that recommender systems have on political decision-making. To encourage reflective endorsement, I proposed incorporating serendipity and novelty into recommender systems, meaning users would be exposed to more diverse information sources, reducing the undesirable effects of filter bubbles and echo chambers. This technical solution should be complemented with social and cultural changes within platforms that recognise the centrality of information for authentic political decision-making, giving social concerns greater thought when democratic processes are concerned.

To foster more deliberation and thus more justified political decision-making, the solutions were twofold. The first of these facilitated more open discourse in online spaces. This included adopting nudges to disincentivise trolling and divisive statements by asking users to reflect on the content of their statement when it contains potentially combative phrasing. However, I argued that as long as political figures are creating culture wars and encouraging divisiveness, then the influence of nudging will be limited. Consequently, to encourage more open deliberation, I proposed restrictions on the use of recommender systems to disseminate targeted political advertising. If political messaging was accessible to all citizens, candidates would be less able to conceal their conceptions of the good behind targeted messaging and would have to justify their views to others, thus engaging with different beliefs. Although platforms may be unwilling to adopt these changes, I argued the risk of undermining political decision-making created stringent moral requirements for less targeting.

The second category of solutions looked to reopen the deliberative space, taking inspiration from vTaiwan. By creating dedicated deliberative spaces in which citizens can voice their opinion, listen to other conceptions of the good, and participate in decision-making processes, citizens can recognise how a decision contributes to the common good and thus makes it more justifiable. Rather than imposing limits on recommender systems, this solution, adopting a sociotechnical systems perspective, looks to change the cultural meanings associated with the groups disseminated by recommender systems. Through creating dedicated deliberative spaces that inform political decision-making, the groups promoted by recommender systems would become an area for discourse as opposed to one associated with interacting with opposing beliefs.

5.2. Recommendations

As the economic demands driving recommender systems result in homogenous and ideologically entrenched recommendations, consequently harming the authenticity of one's beliefs and obstructing the possibility of deliberation, there should be temporary constraints imposed on these systems. For politically-relevant recommendations, such as political advertisements, news, opinion pieces, blogs, and podcasts, there should be optional limits on the amount of personalised content one is exposed to. For instance, rather than someone's feed quickly becoming inundated with homogenous and divisive recommendations, as was the case with *Carol's Journey* (Zadronzy, 2021), platforms should ensure that users can experience more diversity. Prior to further research into the solutions, these limits should apply to politically-relevant recommendations to reduce their effect on individuals' reflective thought and openness to exchange. This does not necessarily mean platforms have to provide ideologically antithetical suggestions but simply offer apolitical recommendations that reflect their other interests, such as entertainment, art, or sports. The root of these undesirable consequences is personalisation directed at maximised engagement, with this often resulting in recommending misinformation, ideologically congenial news, conspiracies, and divisive content. Whilst the provided solutions are justified with further research, the personalisation manifest in recommender systems should be redirected to promote content that reflects the diversity of one's interests, instead of topics which maximise screen time.

5.3. Limitations

This work has limitations. Although incorporating quantitative evidence to justify arguments, such as the profitability of recommender systems (Statista, 2022) and the effectiveness of serendipity (Campos & Figueiredo, 2002; Niu et al., 2018; Reviglio, 2019), the research largely employed philosophical and political theory. To develop greater justification for the solutions, there would need to be further research into how serendipity can affect one's propensity for reflective endorsement, the extent of

the relationship between targeted advertising and adopting divisive behaviour, and whether dedicated deliberative platforms lead to participants finding a decision to be more justifiable when it is irreconcilable with their beliefs.

Should you be convinced by the arguments and solutions, the most significant limitation of this work is its lack of influence. As outlined in chapters 1 and 3, platforms that employ recommender systems are subject to regular critiques regarding their business practices and lack of commitment to ethical values. Many of these platforms offer an ever-growing range of services with the undesirable and unintended consequences of recommender systems maximising engagement being one small aspect. Each of these services faces regular criticism and potential legislative requirements but often introduces small changes to appease regulators. As such, one further ethical criticism is unlikely to have much influence. Moreover, considering the profitability of these platforms, there is a reluctance to respond to ethical concerns, meaning the concrete influence of any proposed social changes will be limited.

All that being said, as demonstrated by vTaiwan, Twitter removing political advertising, and Facebook introducing more stringent political advertising restrictions around the 2020 US Presidential Election, online platforms and recommender systems can incorporate values that extend beyond maximised engagement and profit. In the cases of Twitter and Facebook, platforms recognised how their services could be used to harm the integrity of democratic processes or cause mass unrest and chose to introduce measures that reduced this risk. Each of their decisions would have been at the expense of revenue and behavioural data but they chose to prioritise social values, demonstrating that further improvements are possible. Furthermore, vTaiwan demonstrates that we are not dependent on Big Tech corporations to introduce changes that reduce the harms of disseminating homogenous and inaccurate information. Citizens and governments can strive to create new platforms that prioritise social values, such as incorporating pluralist beliefs in open exchanges that directly contribute to more justifiable political decision-making.

Moreover, as 4.1 outlined, recommender systems are sociotechnical systems and ultimately under our control. Adopting a sociotechnical systems perspective reveals how social and technical aspects influence one another, providing a means of uncovering the ways these relationships can be optimised to prioritise social values. Although the introduction of these solutions may seem unlikely, recommender systems are ultimately sociotechnical and therefore under our control. Sociotechnical systems can evolve in unexpected ways but the introduction of new capabilities can redirect these systems towards more desirable ends. Recommender systems and their effects on political decision-

making, then, remain under our control and can feasibly be redirected to promote conditions that are conducive to more authentic and justifiable political decision-making.

5.4. Further research

As proposed in 4.2.4 and alluded to in the limitations, there should be dedicated sociotechnical research into the relationship between recommender systems and political decision-making, exploring how deliberative platforms could foster participation and directly respond to the discursive exchange found in the groups promoted by recommender systems. This research does not seek to make communication within current groups redundant but to identify the nudges and attitudes that encourage reflection and steer exchange away from divisive and hateful statements. Additionally, further research into a dedicated deliberative platform should assess whether such services make decisions more justifiable to participants, particularly those whose beliefs are irreconcilable with the final decision. This could lend further justification to the creation of deliberative platforms and create more diverse forms of solution generation.

There should also be research into the effects of recommender systems on the authenticity and justifiability of political decision-making in other contexts. This paper has predominately focused on the United Kingdom but could be extended to other countries or regions. Particularly, further research could identify countries that are highly dependent on social media sites for accessing information and assess the effects of recommender systems on their beliefs and exchanges.

Where this paper focused on reflective endorsement and deliberation criteria as necessary criteria for authentic and justifiable political decision-making in liberal democracies, further research could explore the influence of recommender systems on other political criteria. For example, the relationship between recommender systems and rights such as freedom of association and freedom of expression could warrant further exploration. Particularly, this could explore whether recommender systems affect the authenticity of one's groups memberships or if regularly interacting with recommended content influences how one expresses their conceptions of the good.

5.5. Final Statement

This paper has demonstrated how recommending repetitive information and homogenous groups, facilitating divisive exchanges, promoting misinformation, and providing limited access to a candidate's advertising can undermine authentic freedom of thought and justified political decision-making. Where the effects of recommender systems on reflective endorsement and deliberation are unintended consequences of the economic incentives of the attention economy, their potential to undermine authentic and justified political decision-making cannot be ignored. Introducing solutions

that reflect the significance of the political context is vital to ensure that any undesirable consequences can be avoided. This does not require abandoning the attention economy business model altogether but merely giving social concerns greater concern when recommending politically-relevant multimedia. By proposing sociotechnical solutions that could foster new technological capabilities and change current social practices, recommender systems could encourage more authentic and justified political decision-making.

Word count: 19955

Bibliography

- Alabdulwahhab, F. A. (2018). Web 3.0: The Decentralized Web Blockchain networks and Protocol Innovation. *2018 1st International Conference on Computer Applications & Information Security (ICCAIS)*. <https://doi.org/10.1109/cais.2018.8441990>
- Ananny, M. (2015). Toward an Ethics of Algorithms. *Science, Technology, & Human Values*, 41(1), 93–117. <https://doi.org/10.1177/0162243915606523>
- Anderson, B., & Horvath, B. (2017, February 12). *The Rise of the Weaponized AI Propaganda Machine - Scout: Science Fiction + Journalism*. Medium. Retrieved 25 July 2022, from <https://medium.com/join-scout/the-rise-of-the-weaponized-ai-propaganda-machine-86dac61668b>
- Aston University. (2020, February 7). *Social media users 'copy' friends' eating habits*. ScienceDaily. Retrieved 25 July 2022, from <https://www.sciencedaily.com/releases/2020/02/200207074715.htm>
- Barrett, N., & Palumbo, D. (2021, July 4). *What's happened to the 40 new hospitals pledge?* BBC News. Retrieved 26 July 2022, from <https://www.bbc.com/news/59372348>
- BBC News. (2019, October 31). *Twitter to ban all political advertising*. Retrieved 24 July 2022, from <https://www.bbc.com/news/world-us-canada-50243306>
- BBC News. (2021, September 16). *Nicki Minaj swollen testicles claim 'wasted' Trinidad health officials' time*. Retrieved 26 July 2022, from <https://www.bbc.co.uk/news/newsbeat-58571353>
- Bhargava, V. R., & Velasquez, M. (2020). Ethics of the Attention Economy: The Problem of Social Media Addiction. *Business Ethics Quarterly*, 1–39. <https://doi.org/10.1017/beq.2020.32>
- Bohman, J., & Rehg, W. (1997). Introduction. In J. Bohman & W. Rehg (Eds.), *Deliberative Democracy: Essays on Reason and Politics* (pp. ix–xxx). The MIT Press.

- Bosker, B. (2016, November). *Addicted to Your iPhone? You're Not Alone*. The Atlantic. Retrieved 25 July 2022, from <https://www.theatlantic.com/magazine/archive/2016/11/the-binge-breaker/501122/>
- Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209–227. <https://doi.org/10.1007/s10676-013-9321-6>
- Broussard, M. (2018). *Artificial Unintelligence: How Computers Misunderstand the World* (The MIT Press) (Illustrated ed.). The MIT Press.
- Cadwalladr, C. (2016, December 4). *Google, democracy and the truth about internet search*. The Guardian. Retrieved 27 July 2022, from <https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook>
- Cadwalladr, C. (2018, March 18). *'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower*. The Guardian. Retrieved 27 July 2022, from <https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump>
- Campos, J., & Figueiredo, A. D. D. (2002). Programming for Serendipity. *Proceedings of the 2002 AAAI Fall Symposium on Chance Discovery - The Discovery and Management of Chance Events*, AAAI Technical Report FS-02-01. <https://doi.org/10.2139/ssrn.1385402>
- Canales, K. (2020, October 28). *Mark Zuckerberg said Facebook makes a 'relatively small' amount from political advertising. The company has made \$2.2 billion from political ads since mid-2018*. Business Insider. Retrieved 25 July 2022, from <https://www.businessinsider.com/zuckerberg-facebook-political-ad-revenue-2020-10?international=true&r=US&IR=T>
- Chessen, M. (2019). *The MADCOM Future: How Artificial Intelligence Will Enhance Computational Propaganda, Reprogram Human Culture, and Threaten Democracy . . . and What can be*

- Done About It [E-book]. In R. V. Yampolskiy (Ed.), *Artificial Intelligence Safety and Security* (pp. 127–144). Taylor & Francis Group. <https://doi.org/10.1201/9781351251389>
- Coeckelbergh, M. (2020). *AI Ethics*. Amsterdam University Press.
- Cohen, J. (1997a). Deliberation and Democratic Legitimacy. In J. Bohman & W. Rehg (Eds.), *Deliberative Democracy: Essays on Reason and Politics* (pp. 67–92). The MIT Press.
- Cohen, J. (1997b). Procedure and Substance in Deliberative Democracy. In J. Bohman & W. Rehg (Eds.), *Deliberative Democracy: Essays on Reason and Politics* (pp. 407–438). The MIT Press.
- Daly, C. (2010). *An Introduction to Philosophical Methods*. Amsterdam University Press.
- Department of Health and Social Care. (2021, August 1). *New Hospital Programme Communications Playbook* [Press release]. <https://nursingnotes.co.uk/wp-content/uploads/2021/08/NHP-Communications-Playbook- August-2021 -Redacted.pdf>
- Dorsey, J. [jack]. (2019, October 30). *A final note. This isn't about free expression. This is about paying for reach. And paying to increase the reach* [Tweet]. Twitter. <https://twitter.com/jack/status/1189634377057067008>
- Dot Everyone. (2018, April). *People, Power and Technology: The 2018 Digital Understanding Report*. https://doteveryone.org.uk/wp-content/uploads/2019/07/Doteveryone_PeoplePowerTechDigitalUnderstanding2018.pdf
- Elahi, M., Jannach, D., Skjærven, L., Knudsen, E., Sjøvaag, H., Tolonen, K., Holmstad, Y., Pipkin, I., Throndsen, E., Stenbom, A., Fiskerud, E., Oesch, A., Vredenberg, L., & Trattner, C. (2021). Towards responsible media recommendation. *AI and Ethics*, 2(1), 103–114. <https://doi.org/10.1007/s43681-021-00107-7>
- Encyclopedia Britannica. (n.d.). *Hundred Flowers Campaign | Chinese history*. Retrieved 24 July 2022, from <https://www.britannica.com/event/Hundred-Flowers-Campaign>
- Facebook. (2020). *Facebook's Policies for Elections and Voting: What You Need to Know*. Retrieved 24 July 2022, from <https://about.fb.com/wp-content/uploads/2020/10/Facebooks-Policies-for-Elections-and-Voting.pdf>

- Fishkin, J. S. (2009). *When the People Speak: Deliberative Democracy and Public Consultation*. Oxford University Press.
- Fishkin, J. S. (2016). Deliberative Democracy. *Emerging Trends in the Social and Behavioral Sciences*, 1–16. <https://doi.org/10.1002/9781118900772.etrds0405>
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly*, 80(S1), 298–320.
<https://doi.org/10.1093/pog/nfw006>
- Foucault, M. (1971). Orders of discourse. *Social Science Information*, 10(2), 7–30.
<https://doi.org/10.1177/053901847101000201>
- Foucault, M. (1982). *The Archaeology of Knowledge and The Discourse on Language* (Reprint ed.). Vintage. Original work published 1969
- Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68(1), 5–20. <https://doi.org/10.2307/2024717>
- Fuchs, C., Hofkirchner, W., Schafranek, M., Raffl, C., Sandoval, M., & Bichler, R. (2010a). Theoretical Foundations of the Web: Cognition, Communication, and Co-Operation. Towards an Understanding of Web 1.0, 2.0, 3.0. *Future Internet*, 2(1), 41–59.
<https://doi.org/10.3390/fi2010041>
- Fuchs, C., Hofkirchner, W., Schafranek, M., Raffl, C., Sandoval, M., & Bichler, R. (2010b). Theoretical Foundations of the Web: Cognition, Communication, and Co-Operation. Towards an Understanding of Web 1.0, 2.0, 3.0. *Future Internet*, 2(1), 41–59.
<https://doi.org/10.3390/fi2010041>
- Fuenfschilling, L., & Truffer, B. (2014). The structuration of socio-technical regimes—Conceptual foundations from institutional theory. *Research Policy*, 43(4), 772–791.
<https://doi.org/10.1016/j.respol.2013.10.010>
- Garcin, F., Faltings, B., Donatsch, O., Alazzawi, A., Bruttin, C., & Huber, A. (2014). Offline and online evaluation of news recommender systems at swissinfo.ch. *Proceedings of the 8th ACM*

Conference on Recommender Systems - RecSys '14.

<https://doi.org/10.1145/2645710.2645745>

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy*, 31(8–9), 1257–1274.

[https://doi.org/10.1016/s0048-7333\(02\)00062-8](https://doi.org/10.1016/s0048-7333(02)00062-8)

Geels, F. W. (2005). Co-evolution of technology and society: The transition in water supply and personal hygiene in the netherlands (1850–1930)—a case study in multi-level perspective.

Technology in Society, 27(3), 363–397. <https://doi.org/10.1016/j.techsoc.2005.04.008>

Geels, F. W. (2012). A socio-technical analysis of low-carbon transitions: introducing the multi-level perspective into transport studies. *Journal of Transport Geography*, 24, 471–482.

<https://doi.org/10.1016/j.jtrangeo.2012.01.021>

Gillespie, T. (2014). The Relevance of Algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 167–193). The MIT Press.

Goldstein, D. G., Johnson, E. J., Herrmann, A., & Heitmann, M. (2008, December). Nudge Your Customers Toward Better Choices. *Harvard Business Review*, 82(12). Retrieved 25 July 2022, from

https://www.researchgate.net/publication/262042571_Nudge_Your_Customers_Toward_Better_Choices

Goodman, E., Labo, S., Moore, M., & Tambini, D. (2017, March). *The new political campaigning* (Media Policy Brief 19). Media Policy Project, London School of Economics and Political Science.

https://eprints.lse.ac.uk/71945/7/LSE%20MPP%20Policy%20Brief%2019%20-%20The%20new%20political%20campaigning_final.pdf

Griffiths, M. D. (2018). Adolescent social networking: How do social media operators facilitate

habitual use? *Education and Health*, 36(3), 66–69. <http://irep.ntu.ac.uk/id/eprint/35779/>

- Habermas, J. (1991). *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. MIT Press.
- Habermas, J. (1996). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (W. Rehg, Trans.). The MIT Press.
- Hendriks Vettehen, P., & Kleemans, M. (2017). Proving the Obvious? What Sensationalism Contributes to the Time Spent on News Video. *Electronic News*, 12(2), 113–127.
<https://doi.org/10.1177/1931243117739947>
- Henschke, A. (2018). *Ethics in an Age of Surveillance: Personal Information and Virtual Identities* (Reprint ed.). Cambridge University Press.
- Hern, A. (2018, April 10). *How to check whether Facebook shared your data with Cambridge Analytica*. The Guardian. Retrieved 25 July 2022, from
<https://www.theguardian.com/technology/2018/apr/10/facebook-notify-users-data-harvested-cambridge-analytica>
- Hill, K. (2012, February 16). *How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did*. Forbes. Retrieved 26 July 2022, from
<https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>
- Horton, C. (2020, April 2). *The simple but ingenious system Taiwan uses to crowdsource its laws*. MIT Technology Review. Retrieved 24 July 2022, from
<https://www.technologyreview.com/2018/08/21/240284/the-simple-but-ingenious-system-taiwan-uses-to-crowdsource-its-laws/>
- Isaak, J., & Hanna, M. J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. *Computer*, 51(8), 56–59. <https://doi.org/10.1109/mc.2018.3191268>
- Jackson, F. (2000). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Clarendon Press.
- Kahneman, D. (2012). *Thinking, Fast and Slow*. PENGUIN.

- Khan, C. (2021, October 29). *Is my phone listening to me? We ask the expert*. The Guardian.
Retrieved 26 July 2022, from <https://www.theguardian.com/lifeandstyle/2021/oct/29/is-my-phone-listening-to-me-we-ask-the-expert>
- Kirshenbaum, E., Forman, G., & Dugan, M. (2012). A Live Comparison of Methods for Personalized Article Recommendation at Forbes.com. *Machine Learning and Knowledge Discovery in Databases*, 51–66. https://doi.org/10.1007/978-3-642-33486-3_4
- Konrad, K., Truffer, B., & Voß, J. P. (2008). Multi-regime dynamics in the analysis of sectoral transformation potentials: evidence from German utility sectors. *Journal of Cleaner Production*, 16(11), 1190–1202. <https://doi.org/10.1016/j.jclepro.2007.08.014>
- Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge University Press.
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15), 5802–5805. <https://doi.org/10.1073/pnas.1218772110>
- Kroese, F. M., Marchiori, D. R., & de Ridder, D. T. D. (2015). Nudging healthy food choices: a field experiment at the train station. *Journal of Public Health*, 38(2), e133–e137.
<https://doi.org/10.1093/pubmed/fdv096>
- Larkin, B. (2013). The Politics and Poetics of Infrastructure. *Annual Review of Anthropology*, 42(1), 327–343. <https://doi.org/10.1146/annurev-anthro-092412-155522>
- Locke, J. (1983). *A Letter Concerning Toleration* (J. H. Tully, Ed.; UK ed.). Hackett Publishing Company, Inc.
(Original work published 1689)
- Luckerson, V. (2015, July 9). *Here's How Facebook's News Feed Actually Works*. TIME.Com. Retrieved 25 July 2022, from <https://time.com/collection-post/3950525/facebook-news-feed-algorithm/#:%7E:text=Instead%20of%20assuming%20that%20all,engineers%20are%20constantly%20tinkering%20with>

- Macnish, K., & Galliot, J. (2020). An Introduction to Big Data and Democracy. In K. Macnish & J. Galliot (Eds.), *Big Data and Democracy* (pp. 1–12). Edinburgh University Press.
- McKnight, D. (2013). *Murdoch's Politics: How One Man's Thirst For Wealth and Power Shapes our World*. Pluto Press.
- Mehdi Afsar, M., Crump, T., & Far, B. (in press). Reinforcement Learning based Recommender Systems: A Survey. *ACM Computing Surveys*.
- Merrill, J. B., & Oremus, W. (2021, October 26). *Five points for anger, one for a 'like': How Facebook's formula fostered rage and misinformation*. Washington Post. Retrieved 25 July 2022, from <https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/>
- Mill, J. S. (1985). *On Liberty* (G. Himmelfarb, Ed.). Penguin Classics.
(Original work published 1859).
- Mitchell, A., Jurkowitz, M., Oliphant, B. J., & Shearer, E. (2020, July 30). *Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable*. Pew Research Center. Retrieved 25 July 2022, from <https://www.pewresearch.org/journalism/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/>
- Mitchell, G. (2004). Libertarian Paternalism is an Oxymoron. *FSU College of Law, Public Law Research Paper No. 136; FSU College of Law, Law and Economics Paper No. 05–02*, 1–42.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=615562
- Mittelstadt, B. (2016). Auditing for Transparency in Content Personalization Systems. *International Journal of Communication*, 10, 4991–5002. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/89/2016/10/mittelstadt.pdf>
- Mühlhoff, R. (2021). Predictive privacy: towards an applied ethics of data analytics. *Ethics and Information Technology*, 23(4), 675–690. <https://doi.org/10.1007/s10676-021-09606-x>

- Nguyen, C. T. (2018). ECHO CHAMBERS AND EPISTEMIC BUBBLES. *Episteme*, 17(2), 141–161.
<https://doi.org/10.1017/epi.2018.32>
- Nissenbaum, H. (2015). Respecting Context to Protect Privacy: Why Meaning Matters. *Science and Engineering Ethics*, 24(3), 831–852. <https://doi.org/10.1007/s11948-015-9674-9>
- Niu, X., Abbas, F., Maher, M. L., & Grace, K. (2018). Surprise Me If You Can. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*.
<https://doi.org/10.1145/3173574.3173597>
- Nix, A. (2016, March 7). *How big data got the better of Donald Trump*. Campaign. Retrieved 25 July 2022, from <https://www.campaignlive.co.uk/article/big-data-better-donald-trump/1383209>
- O’Neil, C. (2017). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin.
- Pariser, E. (2012). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin.
- Parnaby, L. (2022, June 19). *Youngsters trust politicians less than influencers to give truthful news – poll*. The Independent. Retrieved 25 July 2022, from
<https://www.independent.co.uk/news/uk/survation-tiktok-facebook-youtube-twitter-b2104669.html>
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880.
<https://doi.org/10.1037/xge0000465>
- Peters, J. (2021, July 13). *Twitter will let you change who can reply to a tweet after you post it*. The Verge. Retrieved 27 July 2022, from
<https://www.theverge.com/2021/7/13/22575998/twitter-tweet-change-who-can-reply-after-you-post>
- Purcell, K., & Rainie, L. (2014, December 8). *Americans Feel Better Informed Thanks to the Internet*. Pew Research Center. Retrieved 30 July 2022, from
<https://www.pewresearch.org/internet/2014/12/08/better-informed/>

- Quattrociocchi, W., Scala, A., & Sunstein, C. R. (2016). Echo Chambers on Facebook. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2795110>
- Rajtmajer, S., & Susser, D. (2020). Automated influence and the challenge of cognitive security. *Proceedings of the 7th Symposium on Hot Topics in the Science of Security*, 1–9. <https://doi.org/10.1145/3384217.3385615>
- Rawls, J. (1993). *Political Liberalism*. Columbia University Press.
- Rawls, J. (1999). *A Theory of Justice* (Rev. ed.). The Belknap Press of Harvard University Press
Cambridge.
- Reviglio, U. (2019). Serendipity as an emerging design principle of the infosphere: challenges and opportunities. *Ethics and Information Technology*, 21(2), 151–166. <https://doi.org/10.1007/s10676-018-9496-y>
- Rosner, J. A. (2000). Reflective Endorsement and the Self: A Response to Arpaly and Schroeder. *Philosophical Studies*, 101(1), 107–112. <https://doi.org/10.1023/a:1026454224017>
- Rudman, R., & Bruwer, R. (2016). Defining Web 3.0: opportunities and challenges. *The Electronic Library*, 34(1), 132–154. <https://doi.org/10.1108/el-08-2014-0140>
- Ryfe, D. M. (2005). DOES DELIBERATIVE DEMOCRACY WORK? *Annual Review of Political Science*, 8(1), 49–71. <https://doi.org/10.1146/annurev.polisci.8.032904.154633>
- Saunders, J. (2020). Dark Advertising and the Democratic Process. In K. Macnish & J. Galliot (Eds.), *Big Data and Democracy* (pp. 73–88). Edinburgh University Press.
- Scanlon, T. (1972). A Theory of Freedom of Expression. *Philosophy & Public Affairs*, 1(2), 204–226. <https://www.jstor.org/stable/2264971>
- Schmidt, A. T., & Engelen, B. (2020). The ethics of nudging: An overview. *Philosophy Compass*, 15(4). <https://doi.org/10.1111/phc3.12658>
- Schou, J., & Farkas, J. (2016). Algorithms, Interfaces, and the Circulation of Information: Interrogating the Epistemological Challenges of Facebook. *KOME*, 4(1). <https://doi.org/10.17646/kome.2016.13>

- Seaver, N. (2018). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, 24(4), 421–436. <https://doi.org/10.1177/1359183518820366>
- Shearer, E. (2021, January 12). *More than eight-in-ten Americans get news from digital devices*. Pew Research Center. Retrieved 25 July 2022, from <https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/>
- Smith, J., McCarthy, J. D., McPhail, C., & Augustyn, B. (2001). From Protest to Agenda Building: Description Bias in Media Coverage of Protest Events in Washington, D.C. *Social Forces*, 79(4), 1397–1423. <https://www.jstor.org/stable/2675477>
- Statista. (2021, October 7). *Platforms used to consume news in the UK 2021, by nation*. <https://www.statista.com/statistics/278519/platforms-used-for-news-in-the-united-kingdom-uk/#:%7E:text=According%20to%20a%20recent%20survey,or%20downloaded%20news%20content%20online>
- Statista. (2022, May 2). *Global Facebook advertising revenue 2017–2026*. Retrieved 26 July 2022, from <https://www.statista.com/statistics/544001/facebooks-advertising-revenue-worldwide-usa/#:%7E:text=In%202021%2C%20Facebook%20generated%20nearly,of%20the%20global%20ad%20revenue>.
- Sunstein, C. R., & Thaler, R. H. (2003). Libertarian Paternalism Is Not An Oxymoron. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.405940>
- Susser, D. (2019). Invisible Influence. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 403–408. <https://doi.org/10.1145/3306618.3314286>
- Susser, D., & Grimaldi, V. (2021). Measuring Automated Influence: Between Empirical Evidence and Ethical Values. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 242–253. <https://doi.org/10.1145/3461702.3462532>

- Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, 8(2). <https://doi.org/10.14763/2019.2.1410>
- Tang, A. (2019, March 22). *Inside Taiwan's new digital democracy*. The Economist. Retrieved 24 July 2022, from <https://www.economist.com/open-future/2019/03/12/inside-taiwans-new-digital-democracy>
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian Paternalism. *The American Economic Review*, 93(2), 175–179. <https://www.jstor.org/stable/3132220>
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press.
- The Conservative and Unionist Party. (2019). *The Conservative and Unionist Party Manifesto 2019*. Conservatives. Retrieved 26 July 2022, from https://assets-global.website-files.com/5da42e2cae7ebd3f8bde353c/5dda924905da587992a064ba_Conservative%202019%20Manifesto.pdf
- Turilli, M., & Floridi, L. (2009). The ethics of information transparency. *Ethics and Information Technology*, 11(2), 105–112. <https://doi.org/10.1007/s10676-009-9187-9>
- Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- Twitter Safety. (2022, June 1). *How Twitter is nudging users to have healthier conversations*. Twitter. Retrieved 25 July 2022, from <https://blog.twitter.com/common-thread/en/topics/stories/2022/how-twitter-is-nudging-users-healthier-conversations>
- United Nations. (1948). *Universal Declaration of Human Rights*. Retrieved 26 July 2022, from <https://www.un.org/en/about-us/universal-declaration-of-human-rights>
- United Nations. (1966, December 16). *International Covenant on Civil and Political Rights*. OHCHR. Retrieved 27 July 2022, from <https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

- van den Hoven, J., & Rooksby, E. (2009). Distributive Justice and the Value of Information: A (Broadly) Rawlsian Approach. In J. van den Hoven & J. Weckert (Eds.), *Information Technology and Moral Philosophy* (pp. 376–396). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511498725.019>
- Verbong, G., & Geels, F. W. (2010). Exploring sustainability transitions in the electricity sector with socio-technical pathways. *Technological Forecasting and Social Change*, 77(8), 1214–1221.
<https://doi.org/10.1016/j.techfore.2010.04.008>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- vTaiwan. (n.d.). *about vTaiwan*. Retrieved 24 July 2022, from <https://vtaiwan.tw/intro/>
- Walker, G. H., Stanton, N. A., Salmon, P. M., & Jenkins, D. P. (2008). A review of sociotechnical systems theory: a classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science*, 9(6), 479–499. <https://doi.org/10.1080/14639220701635470>
- Walker, P., & Campbell, D. (2020, October 2). *Most of Boris Johnson's promised 40 new hospitals will not be totally new*. The Guardian. Retrieved 26 July 2022, from <https://www.theguardian.com/society/2020/oct/02/johnsons-37bn-for-40-new-hospitals-in-england>
- Ward, K. (2018). Social networks, the 2016 US presidential election, and Kantian ethics: applying the categorical imperative to Cambridge Analytica's behavioral microtargeting. *Journal of Media Ethics*, 33(3), 133–148. <https://doi.org/10.1080/23736992.2018.1477047>
- Waterson, J., & Milmo, D. (2021, October 25). *Facebook whistleblower Frances Haugen calls for urgent external regulation*. The Guardian. Retrieved 25 July 2022, from <https://www.theguardian.com/technology/2021/oct/25/facebook-whistleblower-frances-haugen-calls-for-urgent-external-regulation>
- Williams, J. (2018). *Stand Out of Our Light*. Cambridge University Press.

- Yang, K., & Fratamico, L. (2022, June 9). *Prompts change actions in the moment*. Twitter Engineering. Retrieved 25 July 2022, from https://blog.twitter.com/engineering/en_us/topics/insights/2022/reconsidering-tweets
- Yeung, K. (2016). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, 20(1), 118–136. <https://doi.org/10.1080/1369118x.2016.1186713>
- Zadrozny, B. (2021, October 22). *What Facebook knew about how it radicalized users*. NBC News. Retrieved 25 July 2022, from <https://www.nbcnews.com/tech/tech-news/facebook-knew-radicalized-users-rcna3581>
- Zuboff, S. (2015). Big other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology*, 30(1), 75–89. <https://doi.org/10.1057/jit.2015.5>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for the Future at the New Frontier of Power*. Profile Books Ltd.
- Zubrow, K. (2021, October 4). *Facebook whistleblower says company incentivizes 'angry, polarizing, divisive content' - 60 Minutes*. CBS News. Retrieved 25 July 2022, from <https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-60-minutes-polarizing-divisive-content/>
- Zuckerberg, M. (2019, March). *A Privacy-Focused Vision for Social Networking*. Facebook. Retrieved 25 July 2022, from <https://www.facebook.com/notes/2420600258234172/>